Proceedings of SALT 27: 680–699, 2017

Pragmatic effects of *more than* and *at least* in incremental interpretation*

Stavroula Alexandropoulou
Utrecht University

Jakub Dotlačil
University of Amsterdam

Rick Nouwen
Utrecht University

Abstract We investigate the incremental interpretation of comparative and superlative numeral modifiers by manipulating the speaker's epistemic state in an eye-tracking reading experiment. The results reveal a different processing profile for two types of numeral modifiers. We take this difference to point to a difference in the source and nature of the attested effects (e.g., Quantity- vs. Manner-based pragmatic reasoning). Our findings inform the existing theoretical landscape, invalidating a number of accounts of speaker ignorance effects with numeral modifiers and giving support to Quantity-based accounts of such effects with superlative modifiers.

Keywords: numeral modifiers, speaker ignorance effects, reading eye-tracking, implicature

1 Introduction

For a long time in the literature on modified numerals, theorists were concerned with the contrast illustrated in (1). They have been arguing that infelicity arises in (1b), and not in (1a), because *at least* as opposed to *more than* signals uncertainty or ignorance on the part of the speaker, which clashes with the speaker competence and knowledgeability conveyed by the continuation sentence. As a consequence, the assumption that comparative modifiers do not give rise to speaker ignorance effects was beyond question by a large number of studies (Geurts & Nouwen 2007; Büring 2008; Cummins & Katsos 2010; Nouwen 2010; Coppock & Brochhagen 2013b; Kennedy 2015).

(1) a. Olga drank more than three margaritas. In fact, she drank five.

* We thank Ilse Angenent, Sofia Bimpikou, Dominique Blok, Heidi Klockmann, Sophia Manika, Yaron McNabb, Teodora Mihoc, Iris Mulders, Myrto Pantazi, Maartje Schulpen, Elli Tourtouri. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP/2007-2013)/ERC Grant Agreement no. 313502 (first & third author), and from the NWO Veni Grant 257.80.005 (second author).

b. Olga drank at least four margaritas. # In fact, she drank five.

This contrast between comparative and superlative numeral modifiers is also manifested in various offline experiments probing speaker ignorance effects using stimuli comparable to (1), but different methodologies (Geurts & Nouwen 2007; Cummins & Katsos 2010; Geurts, Katsos, Cummins, Moons & Noordman 2010; Coppock & Brochhagen 2013a). Although there were already hints implicit in a few earlier theoretical studies (e.g., Fox & Hackl 2006; Mayr 2013), only recently studies have explicitly argued that comparative modifiers, too, can give rise to speaker ignorance, once we shift our attention away from a paradigm similar to that in (1) and consider what an utterance with a numeral modifier can be an answer to (Mayr & Meyer 2014; Westera & Brasoveanu 2014; Ciardelli, Coppock & Roelofsen 2017). More specifically, these studies observe that when an utterance with a comparative modifier is used as an answer to a *how many* question, see (2), it signals speaker ignorance. The idea is that a *how many* question asks for a precise answer and answering with *more than* implies that the speaker is unable to give the precise answer she is asked for, hence she must be ignorant about the precise quantity under discussion.

(2) A: How many margaritas did Olga drink yesterday?
B: She drank more than three margaritas.

Interestingly, there is no complete consensus among the aforementioned authors as to the status and strength of speaker ignorance implications of comparative modifiers. Westera & Brasoveanu (2014) argue that B's utterance in (2) gives rise to speaker ignorance to the same extent as its superlative counterpart, while Ciardelli et al. (2017) claim that superlative modifiers trigger more robust speaker ignorance implications than comparative modifiers.

Westera & Brasoveanu (2014) ran two experiments, each combining a validity judgement task that targets ignorance interpretations and a self-paced reading task in order to look into the incremental interpretation of utterances with comparative and superlative modifiers. They do find that both numeral modifiers trigger speaker ignorance when they are used as an answer to a *how many* question as opposed to other types of questions. However, while their first offline task shows a difference between the two numeral modifiers in ignorance ratings with respect to the *how many* question, this difference does not reach significance in the second offline task. The picture is similar in their online experiments. The first online task reveals interesting effects, which could possibly be connected to the offline data obtained in the first offline task. These effects, though, are not replicated by the second online experiment, which has a better-defined design. Therefore, although Westera & Brasoveanu's (2014) offline experiments do show that *more than* triggers ignorance implications with a *how many* question as opposed to other types of questions under discussion,

their online experimental investigation delivers overall inconclusive results.

We maintain that looking into the incremental interpretation of comparative and superlative numeral modifiers can shed interesting light on the differences and similarities of comparative and superlative modifiers as to speaker ignorance implications and their strength. In the present paper, we directly probe ignorance effects and the time course thereof with a different design from Westera & Brasoveanu's (2014), measuring what happens in real time when interpreting a superlative or a comparative modified numeral in a context with a *how many* question under discussion and where the information state of the speaker is manipulated (knowledgeable vs. ignorant). We do so by means of an eye-tracking reading experiment aiming to obtain insight into the status of ignorance effects of comparative and of superlative modifiers. The results directly inform the theoretical accounts and the relevant debate.

The next section (section 2) presents our study in detail and section 3 concludes.

2 Current study

The current study was conducted in Dutch and consists of two acceptability questionnaires and an eye-tracking reading experiment.

2.1 Design

In order to directly examine and access speaker ignorance effects of *at least* and *more than* in incremental interpretation, we manipulated the speaker's information state. That is, we manipulated:

(i) the numeral modifier in the target sentence (NM manipulation): *meer dan* 'more than' vs. *minstens* 'at least',

(ii) the speaker's epistemic state as set up by the context before the target sentence with the numeral modifier (Context manipulation): *speaker ignorance* vs. *speaker authority* context (cf. McNabb & Penka's (2015) similar manipulation in their second experiment). There was also a third Context condition (*speaker indifference* context), which we will not be concerned with in the present paper.

Thus, the NM and Context factors were manipulated in a 2×3 design. Below we give an example of a test item in all relevant conditions. The texts consisted of an intro followed by the context sentence with the Context manipulation, which was in turn followed by the target sentence with the NM manipulation; some of the texts would end with an outro.

(3) **Example item**

Intro

Wesley heeft zijn eigen zaak waar hij met veel plezier tatoeages zet. Het is er meestal erg druk en hij probeert elke dag acht mensen te tatoeëren. Deze donderdag was hij ook weer hard aan het werk.

'Wesley runs his own tattoo parlor, which he enjoys a lot. It's usually very busy and he tries to tattoo eight customers per day. Last Thursday, he was very busy.'

Context sentence

a. **Speaker ignorance context**

Ik weet niet precies hoe het met de drukte zat, maar ik heb wel een idee.

'I don't know exactly how busy it got, but I have got an impression.'

b. **Speaker authority context**

Ik weet precies hoe het met de drukte zat en daarom zal ik je erover vertellen.

'I know exactly how busy it got, and I'll tell you about it.'

Target sentence

Wesley heeft die dag ***meer dan / minstens*** tien mensen met veel oog voor detail getatoeëerd.
Wesley has that day **more than / at least** ten people with much eye for detail tattooed.

'That day, Wesley tattooed more than/at least ten people with a real eye for detail.'

As is also evident in the example item, the speaker ignorance contexts provide a cue for drawing an ignorance implicature, while authority speaker contexts do not provide a cue to any particular inference and merely convey that the speaker is knowledgeable. In the latter case, the context sentence is compatible with the core meaning of both numeral modifiers, i.e., with the number in question being ≥ 10 for the *at least* condition and ≥ 11 for the *more than* condition of the example item above. This context setup is similar in spirit to that used by Breheny, Katsos & Williams (2006), and Bergen & Grodner (2012) (see also Politzer-Ahles & Fiorentino 2013, a.o.). To illustrate with Breheny et al.'s (2006) study, they aimed to directly measure what happens in real time when people interpret a scalar expression in two different types of context: i.e., contexts that bias comprehenders toward drawing the relevant scalar implicature versus contexts that do not have such a bias, but are rather merely compatible with the basic, lower-bound meaning of the scalar expression.

They did so by means of a self-paced reading task and tested the Greek equivalents of the scalar expressions *or* and *some* (in different experiments). Below we exemplify with *some*, breaking down a translated example item from Breheny et al. (2006) into the context part with the context manipulation and the target sentence (the continuation sentence is omitted).

(4) **Example item** (adapted from Breheny et al. 2006)

Context with bias: Mary asked John whether he intended to host all of his relatives in his tiny apartment. John replied that

Context without bias: Mary was surprised to see John cleaning his apartment and she asked the reason why. John told her that

Target sentence: he intended to host some of his relatives.

The **context with bias** provides a cue for drawing the *not all* scalar implicature, while no such cue is provided by the **context without bias**, hence fewer or no implicatures are expected to arise in the latter condition compared to the former one. By analogy, in our experiment a speaker ignorance implicature is expected to arise in the target sentence in the ignorance contexts, which have an ignorance bias, while no such implicature is expected in the authority contexts, because they do not have any bias and they are even contradictory with ignorance.

One could argue that speaker ignorance is entailed by the specific ignorance context and that this together with the core meaning of the numeral modifiers is sufficient for an ignorance interpretation of the target sentence, hence no implicature is necessary to be generated in the target sentence. However, there is experimental evidence by Mendia (2016) showing that the speaker ignorance implicature triggered by *at least* is not just the reading *the speaker lacks any beliefs about numbers in a certain range*, say, the range $[10, ...)$ for the example item, but rather the stronger reading that *the speaker considers it possible that the number under discussion is equal to 10 and she considers it possible that it is greater than 10*. Contrary to the former reading, the latter is not entailed and, thus, it has to be derived by pragmatic means. Although there is no similar finding for the type of the ignorance implicature *more than* is associated with, there is evidence that, in contrast to the rest of the values, the minimum value of the relevant range (11 in the example item) is a necessary alternative fed into the pragmatic mechanism *more than* involves (see Alexandropoulou, Dotlačil, McNabb & Nouwen 2015 on variation/scalar implicatures of modified numerals). Glossing over the (derivation) details, this would result in the stronger reading given above for *at least*'s speaker ignorance.

Lastly, our texts were constructed in such a way so as to implicitly introduce a certain quantity under discussion.

2.2 Predictions

2.2.1 Predictions of theoretical accounts

The theoretical accounts that derive speaker ignorance only with *at least* and not with *more than* (Geurts & Nouwen 2007; Büring 2008; Nouwen 2010; Coppock & Brochhagen 2013b; Kennedy 2015) would predict a difference between comparative items and superlative items; especially those accounts deriving obligatory ignorance with unembedded occurrences of *at least* (i.e., Geurts & Nouwen 2007; Nouwen 2010; Coppock & Brochhagen 2013b). While the target sentence with *more than* is semantically compatible both with a speaker ignorance and a speaker authority context, this is not the case for the superlative condition. The core meaning of *at least* is compatible with the ignorance and the authority context, but its obligatory ignorance implication—be it an entailment or an implicature—is compatible with the ignorance context, but clashes with the speaker authority context. Thus, these accounts predict incoherence to arise between the authority context sentence and the target sentence with *at least*. No such incoherence is predicted by Büring's (2008) or Kennedy's (2015) Quantity-based account of ignorance, as they do not derive obligatory ignorance effects with *at least*, so given our context setup ignorance is expected to arise only in the ignorance context sentence.

According to Westera & Brasoveanu (2014), speaker ignorance implications should arise to the same extent in the target sentence with *more than* and the target sentence with *at least* when following an ignorance context sentence as opposed to an authority context sentence. This is expected because in our contexts there is always a certain (precise) quantity under discussion, which corresponds to the *how many* question under discussion that Westera & Brasoveanu (2014) take as the right context for triggering homogeneous ignorance implications with superlative and comparative numeral modifiers. Recall that this homogeneity was only confirmed by the findings of their second offline experiment, as their first experiment showed higher ignorance ratings in the superlative than in the comparative condition when there was a(n explicit) *how many* question under discussion. As ignorance implications are argued to be context-sensitive on this account, no ignorance is expected to arise in the target sentence in the authority context condition, where the context sentence directly contradicts speaker ignorance.

Ciardelli et al. (2017) derive speaker ignorance implicatures both with superlative and comparative numeral modifiers with a *how many* question under discussion, but argue that the implicatures are stronger with superlative modifiers as compared to comparative modifiers in such a context. Specifically, they derive them as a type of a Quality implicature for superlatives and as a standard Quantity implicature for comparatives, basing their difference in strength on the assumption that observing

a Quality maxim is more urgent than observing a Quantity maxim. Thus, given our context setup (with a certain quantity being discussed), speaker ignorance implicatures are expected to arise in the target sentence in both NM conditions in the ignorance condition, but will be of a more robust nature with *minstens* than with *meer dan* due to their different source. Moreover, the ignorance implicatures with superlatives are taken to be obligatory because they are tied to a certain Quality maxim by virtue of the particular semantics of superlative modifiers on this account and this maxim roughly requires that the speaker entertains multiple possibilities in her information state with respect to the quantity under discussion. This means that the target sentence with *minstens* will be incompatible with the authority context sentences, because those reveal a totally knowledgeable speaker with respect to the quantity under consideration.

2.2.2 Processing predictions

There are no explicit predictions as far as the processing of numeral modifiers and their incremental interpretation are concerned. The vast majority of the existing accounts, including the aforementioned ones, consist in theoretical proposals with no direct processing implications. Nonetheless, there exist a few proposals incorporating a part that relates to processing, capturing the difference between superlative and comparative modifiers. Cummins & Katsos (2010), for example, claim that superlative modifiers are *more difficult to process at a psychological level* (Cummins & Katsos 2010: 279) compared to comparative modifiers, because the former have a disjunctive meaning (e.g., = n or > n for *at least n*), which is psychologically more complex than either of the disjuncts. Geurts et al. (2010) too, who embrace Geurts & Nouwen's (2007) modal account of superlative modifiers, maintain that because of their modal semantics superlative modifiers are harder to process compared to comparatives, which do not have such a semantics.

Geurts et al. (2010) justify their processing statement by means of the findings of an online verification task, where participants have to indicate whether a sentence with a numeral modifier (superlative/comparative/*exactly*) is true of a situation where up to four identical X's were presented on a screen. Superlative modifiers as opposed to comparative modifiers (or the baseline *exactly* condition) were found to delay verification judgements of the relevant sentences, though no difference was found in the (total) reading times of those sentences. Although the authors conclude that their findings show that superlative modifiers are (semantically) more complex, we can draw no firm conclusion with respect to their online interpretation. That is, one cannot be sure what the delay in decision times should be associated with, especially given that this does not manifest itself in the reported reading times of the relevant sentences. Thus, it is very unclear why the modal interpretation of

superlative modifiers would only occur in the decision times affecting them.

Cummins & Katsos (2010) make a similar point and in order to justify their own processing assumption they conduct a similar experiment making the following modifications: they replace the numeral modifiers (*exactly, more than, less than, at least, at most*) with comparison operators ($=, >, <, \geq, \leq$), where the latter two operators are disjunctive consisting of the simpler operators ($=, >$, and $=, <$, respectively). They find the same results as Geurts et al.: the disjunctive operators delay the decision procedure compared to the respective simple operators, similarly to superlative modifiers in Geurts et al. 2010 as compared to the comparative and *exactly* conditions. Cummins and Katsos conclude that the observed effect has nothing to do with the supposed lexical modality of superlative modifiers, but it should rather be due to their disjunctive interpretation, which they do share with the operators $\geq$ and $\leq$. However, we should be cautious with such a conclusion. The mean decision times for the disjunctive conditions in Cummins & Katsos 2010 ($\geq$: 1110ms, $\leq$: 1131ms) are way shorter than the corresponding superlative conditions in Geurts et al. 2010 (*at least*: 1559ms, *at most*: 1982ms), while the difference between the $=$ and *exactly* baseline conditions is not that big (982ms and 1114ms, respectively). Even if we assume that (part of) the difference is due to Geurts et al.'s (2010) participants converting the linguistic expressions to the mathematical expressions or just due to processing spillover because of the length difference, still it seems that a bigger complexity is possibly associated with superlative modifiers compared to the mathematical operators. All in all, we infer that although the two accounts in question include a processing profile on top of their theoretical one, the (real-time) processing evidence reported lend them very little support.

Neither can Westera & Brasoveanu's (2014) findings be used in favor of the processing proposal of the accounts in question. In their first self-paced reading task, Westera & Brasoveanu (2014) find that participants slow down when reading the region of the numeral that is being modified and at subsequent regions, but they do not find a difference between superlative and comparative modifiers, as Cummins & Katsos (2010) and Geurts et al. (2010) would predict. They only find such a difference (at the expense of the superlative condition) five regions after the numeral modifier. Westera & Brasoveanu (2014), on the other hand, take their effects to be due to the costly on-line calculation of speaker ignorance inferences via pragmatic reasoning or possible silent intonational effects during reading. However, as we mentioned above, in a second similar experiment Westera and Brasoveanu failed to replicate what they found about the incremental interpretation of sentences with numeral modifiers. To conclude, the little processing predictions available in the literature of numeral modifiers seem to be disfavored by any existing conclusive experimental evidence.

Despite the absence of a concrete processing profile for superlative modifiers in

the literature (due to the limited availability of successful predictions and definitive empirical evidence), we can inherit some processing predictions from a previous experiment we carried out (see Alexandropoulou 2018)[1], which was very similar to the present one. In that experiment, we set out to investigate from scratch the time course (i) of interpreting the superlative modifier *at least* (in Dutch) in a similarly manipulated context, and (ii) of accessing speaker ignorance inferences. Crucially, we maintained a neutral approach to the processing of the superlative modifier. That is to say, we considered all possible processing implications that the predictions of the theoretical accounts we discussed in section 2.2.1 could have. We considered both the possibility that the online interpretation (of a felicitous occurrence) of the modified numeral may be associated with a processing difficulty and the possibility that it may be rapidly generated. We furthermore considered the possibility that incompatibility of a speaker ignorance interpretation with the context, as predicted by a number of accounts (e.g., Geurts & Nouwen 2007; Nouwen 2010; Coppock & Brochhagen 2013b; Ciardelli et al. 2017) for the use of unembedded *at least* in authority contexts, may (or may not) disrupt reading. We found that readers were more likely to re-read the numeral phrase, which followed the superlative modifier and where the interpretation of the whole modified numeral is completed, when they were in an ignorance as opposed to an authority context. We took this finding to indicate that the online interpretation of an *at least* modified numeral is associated with a processing cost and that readers return to the region in question in order to derive the ignorance implicature *the speaker considers it possible that exactly ten people were tattooed by Wesley and she considers it possible that Wesley tattooed more than ten people*, here given for the superlative condition of the target sentence in (3). We further concluded that our findings are in favor of the Quantity-based accounts such as Büring 2008, Kennedy 2015, and Schwarz 2016, which derive this type of speaker ignorance.

Taking into account our previous findings and the theoretical predictions we discussed in section 2.2.1, let us now consider in detail what the processing predictions for the present experiment could be. According to the theoretical accounts that derive speaker ignorance only with superlative and not with comparative modifiers (Geurts & Nouwen 2007; Büring 2008; Cummins & Katsos 2010; Nouwen 2010; Coppock & Brochhagen 2013b; Kennedy 2015), we would predict a different processing profile for the two numeral modifiers. Specifically, Büring (2008) and Kennedy (2015), who derive speaker ignorance as a Quantity implicature, predict a replication of the effect of our previous experiment for *at least*, manifested itself as a difference between authority items and ignorance items in the superlative condition. Cummins & Katsos (2010), who take superlative modifiers to be psychologically more complex because

1 Our previous experiment has also been reported in a paper that appeared in SALT 26 proceedings as Alexandropoulou, Dotlačil & Nouwen (2016).

of their disjunctive status, predict a processing penalty for superlative items at the modified numeral phrase across Context conditions. Geurts & Nouwen (2007), who take ignorance to follow from the lexical semantics of superlative modifiers, predict a contradiction of *at least*-target sentences with the authority context sentences, which is expected to disrupt comprehension and cause a processing difficulty (interaction of authority and superlative). The same is expected given Nouwen (2010), who derives speaker ignorance via the combinatorics of a last-resort insertion of a silent existential modal with an unembedded *at least*. Perhaps a smaller disruption is expected if his last-resort strategy is assumed to incur some extra processing cost. Next, Coppock & Brochhagen (2013b), who derive an obligatory type of speaker ignorance via a certain Quality maxim (similar to that assumed by Ciardelli et al. 2017), predict an ignorance implicature to arise across Context conditions. So this is expected to induce a processing penalty given our previous finding, which might be bigger in the authority condition, because there the Quality maxim is violated, as it requires that the speaker considers multiple epistemic possibilities.

Lastly, while Westera & Brasoveanu (2014) predict a similar processing profile for the superlative and the comparative modifier, and thus an effect as in our previous experiment (no interaction), Ciardelli et al. (2017) make different predictions for the two modifiers given the *how many* question under discussion. In fact, their predictions are very similar to those by Coppock & Brochhagen (2013b). The ignorance implicature is more robust and obligatory with *at least* than with *more than*, hence a larger effect is expected in the ignorance condition for *at least* (interaction of superlative and ignorance). However, as such an implicature is obligatory and arises in the authority condition too, which is not the case for *more than*, a violation of the relevant Quality maxim will occur at the *at least*-modified numeral, possibly resulting in a main effect of the superlative condition or in an even greater effect of the superlative in the authority condition, if we assume that violating a fundamental pragmatic maxim (i.e., Quality) is costlier than obeying it.

2.3 Questionnaires

We conducted an offline study in order to test the strength of the contexts in terms of knowledgeability. More precisely, by means of a questionnaire we asked native speakers of Dutch ($N = 32$) to judge the author's (of the texts) knowledgeability of the quantity being discussed. Participants were given 39 test items like that in (3) (Latin square design) intermixed with 36 filler items (75 trials in total), with the difference that the modified numeral in the target sentence was replaced with the hash symbol (#), as in (5). They were said that # stands for a quantity and were next asked to indicate how likely it is that the author knows the precise quantity. They had to do so on a Likert scale from 1 to 7, where 1 stands for "highly unlikely" and 7 for

"highly likely". The author of the ignorance contexts was expected to receive low likelihood ratings, below the middle value of the scale (i.e., 4), while the author of the authority contexts was expected to be judged as being highly likely to know the precise quantity. The filler items were completely different from the test items and contained no specification as far as the author's knowledgeability state is concerned.

(5) **Example item: Target sentence**
Wesley heeft die dag # mensen met veel oog voor detail getatoeëerd.
Wesley has that day # people with much eye for detail tattooed.
'That day, Wesley tattooed # people with a real eye for detail.'

The author in the ignorance contexts was judged to be (highly) significantly less likely to be knowledgeable of the precise quantity compared to the author in the authority contexts ($t = -9.022$, $p < .0001$). Specifically, the latter condition scored very close to the right end of the relevant scale ($mean = 6.050$), while the former condition obtained a mean likelihood rating lower than the middle value of the Likert scale ($mean = 3.411$), as expected. Therefore, our authority contexts are strong knowledgeability contexts with respect to the precise quantity in question, while ignorance contexts are poor(er) knowledgeability contexts.

Perhaps one would expect that the ignorance Context condition would get an even lower mean score, however, the fact that the second half of the ignorance context sentences would state that the author does have an impression or suspicion or some thoughts with respect to the number in question possibly justifies the actual mean score of that condition.

Importantly, before moving to the eye-tracking experiment, we would like to point out that as in our previous study we also ran an offline experiment ($N = 18$) that tested the compatibility of the target sentence (with the NM manipulation) with the preceding authority/ignorance context, and we found the following: the authority condition scored high on the 1-7 coherence scale, with $mean = 5.509$, while no difference was found between the two NM conditions (*meer dan*: $mean = 5.376$, *minstens*: $mean = 5.641$, $p = .456$). This already goes against those accounts that predicted an incoherence for *at least* in the authority condition either due to a contradiction (Geurts & Nouwen 2007; Nouwen 2010) or due to the violation of an inescapable pragmatic maxim (Coppock & Brochhagen 2013b; Ciardelli et al. 2017).

2.4 Eye-tracking reading experiment

2.4.1 Participants

Forty native speakers of Dutch participated in the eye-tracking experiment, who were recruited from the UiL OTS participant database. The data of thirty-eight

(4 male, mean age: 23.66, age range: 18–42) of them were further considered, as two participants did not complete the experiment due to technical problems during testing. All participants were paid for their participation in the experiment. Lastly, participants had normal or corrected to normal vision and were naive as to the purpose of the study.

2.4.2 Materials

Of the 39 items included in the questionnaires, three were excluded from further testing because of poor coherence between the relevant target sentences and the preceding context sentences (low mean coherence ratings overall). Our test items appeared in six conditions (3 Context $\times$ 2 NM). They were rotated through twelve lists in a Latin square design and every list had six items per condition. Every participant saw only one list with 112 trials in total: four practice trials, thirty-six test trials, and seventy-two filler trials. The trials of every list were randomly ordered for each participant. Fifty-two comprehension statements about the story narrated in the texts were included too, to control for whether participants pay attention to the texts they are reading. Participants had to indicate whether the statement was true or false given the text they had just read. These comprehension statements targeted information that was irrelevant to the manipulations of the experiment. Twenty-one (out of the 36) experimental items were followed by such a comprehension statement, while the remaining 31 comprehension statements were follow-ups of the filler items. The target sentence in 22 of the experimental items was not the last sentence of the text, but was followed by an outro sentence, such as *He also has two employees in his service that have learned the art of tattooing from him* for the example item in (3).

2.4.3 Procedure

Participants were seated in a comfortable chair in a sound-treated booth and the distance of their head to the screen was 55–70 cm. Their eye movements were recorded with an EyeLink 1000 eye tracker in remote mode (using a target sticker), sampling at 500Hz. The stimuli were presented on a 17-inch Acer AL1717 monitor and a three-button button box was provided to participants for answering the comprehension questions or moving on.

Participants first read the instructions, where they were informed that they would be presented with short stories, each consisting in a short paragraph. After the story, a comprehension question about the text just presented would occasionally appear, which they had to answer by using the button box (right button for YES and left button for NO). Participants had to press the third, middle, button in order to

Region 3	Region 4	Region 5	Region 6	Region 7	Region 8	Region 9
Wesley	has	that day	*more than* / *at least*	ten people	with a real eye for detail	tattooed.

Table 1 Eye-tracking experiment: Regions of target sentence in (3).

progress. After reading the instructions, the calibration procedure with nine fixation points would start. After a successful calibration, participants would move on to the practice block, where they read four practice items and answered comprehension questions that two of them were associated with. Then the experiment block began.

2.4.4 Results

Of the 38 participants one participant was excluded prior to analysis because she answered 63% of the comprehension statements correctly, while the remaining 37 participants scored 77% or higher. The data of all 37 participants were included in the statistical analyses. Data points of an item seen by 11 participants were excluded because it contained a typo in the context sentence (.86% of the observations). Another 1.12% of the overall observations was excluded because of problematic trials (i.e., trials with few or no fixations due to track losses or participants' sloppy reading). For all measures, fixations of 0 *ms* were excluded from the analysis. We log-transformed the remaining reading-time data to approximate normality.

We split the texts into regions for the purpose of the analyses. The intro (see (3)) constituted Region 1, while the speaker context was Region 2. The target sentence was divided into individual regions as illustrated in Table 1 for the example item in (3). The outro of the texts was excluded from the region partitioning, as not all test items contained an outro sentence. For each of the regions we analyzed seven eye-tracking measures: (i) *first-pass time* (sum of fixations in a region before exited for the first time to any direction), (ii) *right-bounded time* (sum of fixations in a region before exited for the first time to the right), (iii) *regression path duration* (sum of fixations since first fixation in a region until it is exited to the right for the first time), (iv) *probability of regression* (the probability of doing a first-pass regression from a given region), (v) *total reading time* (sum of all fixation durations in a region), (vi) *re-reading time* (sum of fixations in a region excl. first-pass fixations), and (vii) *probability of re-reading* (the probability of entering a region more than once).

We conducted linear mixed-effects regression analyses on the log-transformed reading-time data and mixed-effects logistic regression analyses on the categorical regression probability and re-reading probability measures using the `lme4` package (Bates, Mächler, Bolker & Walker 2015) in `R`. All analyses included three predictor variables, Context (authority / ignorance / indifference context), NM (*more than* vs.

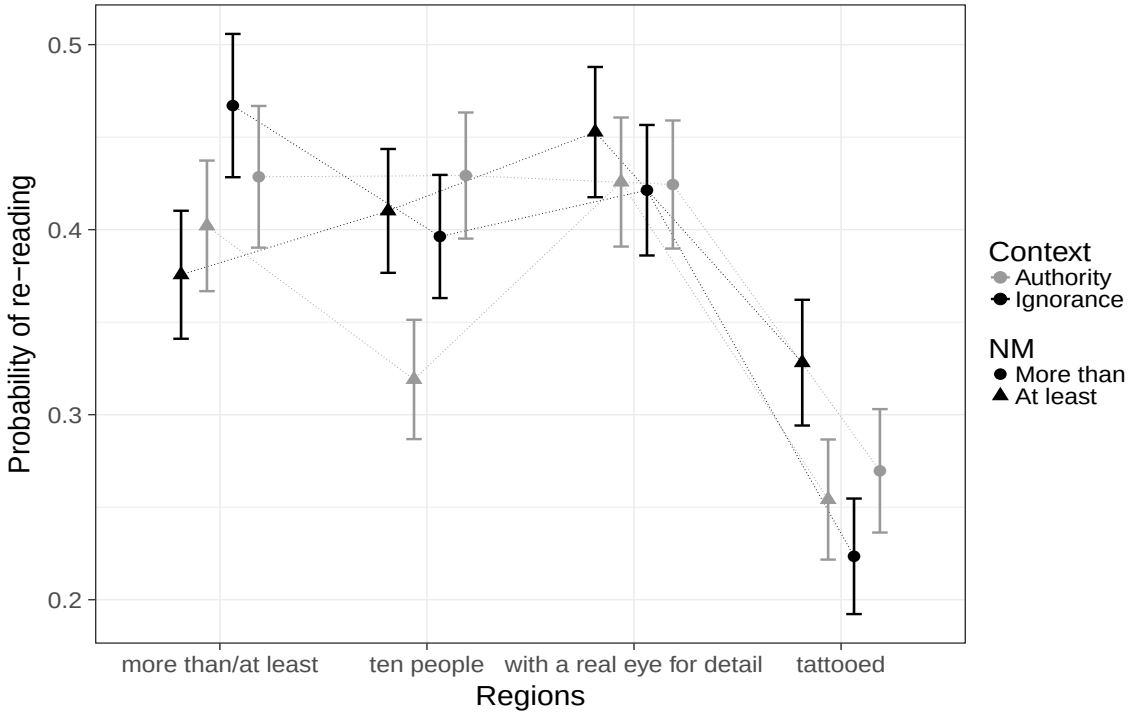

Figure 1 Probability of re-reading regions of the target sentence per condition.

at least) and their interaction. The variables were treatment-coded, with authority and *more than* as baselines. The analyses also included random intercepts and slopes for participants and items for all fixed effects. We applied backward model selection for random effects (Barr, Levy, Scheepers & Tily 2013) and in the following we will be presenting the output of the model with the maximal random-effect structure that converged and had the best fit. *P*-values are calculated based on Satterthwaite's approximations using the *ImerTest* package of R.

No significant main effects or interaction of Context and NM factors on any measure were observed at Region 1 (Intro) up to Region 6 (*more than/at least*). A significant positive interaction was attested between the Context and NM factors at Region 7 (*ten people*) in re-reading probability ($\beta = .601,\ SE = .299,\ z = 2.010,\ p < .05$): the difference between the ignorance and authority Context conditions was bigger in the superlative than in the comparative condition, as illustrated by Figure 1. This indicates that readers were more likely to re-read the critical region of the numeral phrase in an ignorance context as opposed to an authority context in the superlative condition than in the comparative condition.

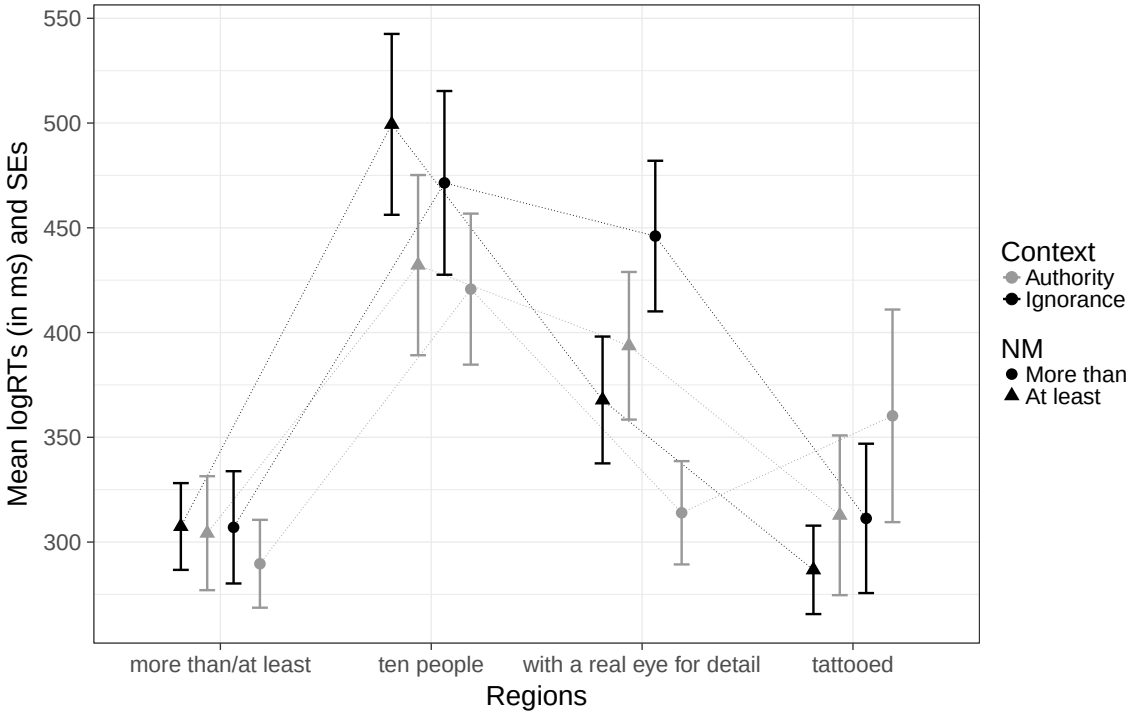

Figure 2 Re-reading times for regions of target sentence per condition.

Furthermore, as also manifested by Figure 1, the statistical analysis revealed a significant negative effect of NM on the probability of re-reading in the same region (Region 7: *ten people*), showing that readers were more likely to re-read that region when it followed *more than* as opposed to *at least* in an authority context. A post-hoc analysis for the separate modifiers showed that there was only an effect of Context in the superlative condition ($\beta = .442,\ SE = .215,\ z = 2.055,\ p < .05$): the numeral phrase of ignorance items was more likely to be re-read than that of authority items when it was modified by *at least*. This constitutes a replication of the effect of our first experiment (see Alexandropoulou 2018). No effect on any other measure was found to be significant in Region 7, except for a marginal positive effect of Context in re-reading times ($\beta = .204,\ SE = .104,\ t = 1.958,\ p = .052$).

In Region 8 (*with a real eye for detail*), which is a spillover region, following the modified numeral phrase, there was a positive effect of Context ($\beta = .272,\ SE = .107,\ t = 2.551,\ p < .05$) and a negative interaction of NM and Context factors ($\beta = -.456,\ SE = .180,\ t = -2.531,\ p < .05$) on re-reading times. When readers re-read the prepositional phrase in Region 8 (see how often they read it, compared to

other regions of the target sentence, in Figure 1) they exhibit a slowdown if they are in an ignorance context as opposed to an authority context and this slowdown is bigger in comparative items than in superlative items (see Figure 2). No significant effect on the other measures was found in Region 8, besides a marginally significant effect of ignorance Context on regression path duration ($\beta = .154,\ SE = .087,\ t = 1.760,\ p = .083$) and on total reading times ($\beta = .110,\ SE = .058,\ t = 1.892,\ p = .063$). These effects suggest that readers tend to spend more time overall reading Region 8 of comparative items, but also when first reading this region including possible fixations to previous regions, in the ignorance condition rather than in the authority condition.

Finally, Region 9 (*tattooed*), the last region of the target sentence, exhibited only a marginal interaction effect between NM and Context on re-reading probability ($\beta = .670,\ SE = .354,\ z = 1.890,\ p = .059$), such that the probability of re-reading that region in an ignorance context as compared to an authority context is higher in the superlative than in the comparative condition. This difference is visible in Figure 1 too. In the next section, we discuss the results of the present experiment and evaluate the predictions presented in section 2.2.2 against all our results.

2.4.5 Discussion

As in our previous experiment (Alexandropoulou 2018), we find effects on the probability of re-reading the critical Region 7 (*ten people*). The significant interaction between Context and NM attested in this measure and region (also marginal in Region 9), i.e., the difference between ignorance and authority items being greater in the superlative condition than in the comparative one, goes against a number of accounts. First, it goes against the semantic accounts of ignorance with superlative modifiers that derive no ignorance with comparatives, such as Geurts & Nouwen 2007, according to which we predicted a disruptive effect in a critical region of the modified numeral phrase in the authority condition of superlative items due to the resulting contradiction. It is also at odds with Nouwen 2010, according to which a similar contradiction, and thus disruptive effect, was predicted for superlative items in the authority condition or a processing penalty for superlative items overall as compared to comparative items, if we assume that Nouwen's last-resort strategy that is responsible for speaker ignorance is an effortful process. Recall that the results of our compatibility pretest were also in conflict with the predictions of these two accounts. Moreover, this interaction effect invalidates Westera & Brasoveanu's (2014) prediction of a similar processing profile between the comparative and the superlative condition given our context setup. It could further be in favor of Coppock & Brochhagen 2013b or Ciardelli et al. 2017, according to which we predicted a difference between superlative and comparatives, as long as we assume that obeying the respective Quality maxims that derive speaker ignorance is more demanding

in terms of processing resources than violating those maxims, as is predicted for superlative items in the authority condition. Note, however, that our compatibility pretest revealed high coherence ratings for the target sentence with an authority context sentence, regardless of the NM type, indicating that no contradiction relation holds between the two sentences. The attested interaction seems to be in support of Büring (2008) and Kennedy (2015), if we assume that the difference between the comparative and the superlative condition is to be associated with something that makes the latter modifiers more complex, such as the availability and derivation of speaker ignorance implicatures via standard Quantity-based reasoning.

Furthermore, in the same measure and region, we also found a difference between comparative and superlative items in the baseline authority context condition at the expense of the former (negative effect of superlative condition). This again is unexpected given accounts that derive obligatory and robust ignorance effects with superlative modifiers and no or less robust ignorance effects with comparative modifiers (Geurts & Nouwen 2007; Nouwen 2010; Coppock & Brochhagen 2013b; Ciardelli et al. 2017), as, if anything, we would expect the relevant difference to be in the opposite direction. But what could have caused this difference, given that the meaning of a comparative target sentence and of a superlative target sentence is equivalent in the authority condition? One could argue that the effect in question could possibly be associated with the fact that, all else being equal, *at least*, but not *more than*, states and does not exclude the minimum value *exactly n* of the relevant range. This might seem more appropriate, since *at least* draws attention to a number that would make the statement true, unlike *more than*.

Importantly, the fact that superlative items were found to be more likely to be re-read in the ignorance condition than in the authority one, is a replication of our finding in our previous experiment (Alexandropoulou 2018). It is also against Cummins & Katsos (2010), who predicted that the superlative modifier will be psychologically as complex and taxing in the two Context conditions. This finding is further taken to reflect that readers revisit the region of *ten people* in order to derive via an implicature the reading *the speaker considers it possible that Wesley tattooed exactly ten people and she considers it possible that he tattooed more than ten people*, which is not entailed by the context. Overall, our results so far are in accordance with Büring's (2008) and Kennedy's (2015) Quantity-based accounts of speaker ignorance effects of superlative modifiers, but against Cummins & Katsos (2010). They are moreover against Geurts & Nouwen's (2007), Nouwen's (2010), Coppock & Brochhagen's (2013b), and Ciardelli et al.'s (2017) accounts that derive obligatory and robust ignorance implications with *at least* as well as Westera & Brasoveanu (2014) who predict a uniform behavior for comparatives and superlatives.

How could we explain, though, the effects in Region 8 (*with a real eye for detail*), which come to confirm the different processing profile displayed by comparative and

superlative modifiers? Readers slow down when re-reading Region 8 of comparative items in the ignorance condition and tend to slow down in regression path and total reading times, too. This is unexpected considering the predictions we put forth given the existing accounts of modified numerals. Perhaps less so given Ciardelli et al. (2017), if we assume that such a finding is possibly the signal of the online derivation of a weaker or different speaker ignorance implication with *more than* as compared to *at least* that kicks in late(r) during incremental interpretation. Another explanation could be that these effects occurring after the critical regions (6 and 7) have been read reflect that comprehenders had an afterthought: *"why did the speaker not use 'minstens' (at least), which is a better cue to speaker ignorance?"*. Such a reasoning results in a Manner implicature, apparently interfering with online processing. On this interpretation, the effect in question seems to be compatible with the theories predicting no ignorance implication with comparative modifiers (Geurts & Nouwen 2007; Büring 2008; Cummins & Katsos 2010; Nouwen 2010; Coppock & Brochhagen 2013b; Kennedy 2015).

3 Conclusion

We investigated the incremental interpretation of numeral modifiers by manipulating the speaker's epistemic state in an eye-tracking reading study. Our results revealed a different processing profile for superlative and comparative modifiers: while a processing penalty is detected at the numeral phrase following the superlative modifier when used by an ignorant speaker (in late measures), a processing penalty occurs at a later point of a sentence with a comparative modifier uttered by an ignorant speaker (marginal in early and significant in late measures). We take our findings to suggest that pragmatic reasoning is attested online: in the form of a non-obligatory Quantity-based ignorance implicature with superlative modifiers, and possibly as a Manner implicature or as an ignorance implicature of a different nature with comparative modifiers.

References

Alexandropoulou, S., Jakub Dotlačil & Rick Nouwen. 2016. *At least* ignorance inferences come at a processing cost: Support from eye movements. In Mary Moroney, Carol-Rose Little, Jacob Collard & Dan Burgdorf (eds.), *Semantics and Linguistic Theory (SALT)* 26, 795–813.

Alexandropoulou, Stavroula. 2018. *On the pragmatics of numeral modifiers: The availability and time course of variation, ignorance, and indifference inferences*: Utrecht University PhD dissertation.

Alexandropoulou, Stavroula, Jakub Dotlačil, Yaron McNabb & Rick Nouwen. 2015.

Pragmatic inferences with numeral modifiers: Novel experimental data. In Sarah D'Antonio, Mary Moroney & Carol Rose Little (eds.), *Semantics and Linguistic Theory (SALT)* 25, 533–549.

Barr, Dale J., Roger Levy, Christoph Scheepers & Harry J. Tily. 2013. Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68(3). 255–278.

Bates, Douglas, Martin Mächler, Ben Bolker & Steve Walker. 2015. Fitting linear mixed-effects models using lme4. *Journal of Statistical Software* 67(1). 1–48.

Bergen, L. & D. J. Grodner. 2012. Speaker knowledge influences the comprehension of pragmatic inferences. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 38(5). 1450–1460.

Breheny, Richard, Napoleon Katsos & John Williams. 2006. Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition* 100(3). 434–463.

Büring, Daniel. 2008. The least *at least* can do. In Charles B. Chang & Hannah J. Haynie (eds.), *West Coast Conference on Formal Linguistics (WCCFL)* 26, 114–120.

Ciardelli, Ivano, Liz Coppock & Floris Roelofsen. 2017. Implicatures of modified numerals: Quality or Quantity? In *Sinn und Bedeutung* 21.

Coppock, E. & T. Brochhagen. 2013a. Diagnosing truth, interactive sincerity, and depictive sincerity. In Todd Snider, Brendan Jackson & Tanya Matthews (eds.), *Semantics and Linguistic Theory (SALT)* 23, 358–375.

Coppock, Elizabeth & Thomas Brochhagen. 2013b. 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.

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

Geurts, Bart, Napoleon Katsos, Chris Cummins, Jonas Moons & Leo Noordman. 2010. Scalar quantifiers: Logic, acquisition and processing. *Language and Cognitive Processes* 25(1). 130–48.

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

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.

Mayr, Clemens. 2013. Implicatures of modified numerals. In Ivano Caponigro & Carlo Cecchetto (eds.), *From grammar to meaning: The spontaneous logicality of language*, 139–171. Cambridge: Cambridge University Press.

Mayr, Clemens & Marie-Christine Meyer. 2014. More than at least. *Two Days At*

Least Workshop, Utrecht University.

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

Mendia, Jon Ander. 2016. Reasoning with partial orders: Restrictions on ignorance inferences of superlative modifiers. In Mary Moroney, Carol-Rose Little, Jacob Collard & Dan Burgdorf (eds.), *Semantics and Linguistic Theory (SALT)* 26, 489–505.

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.

Politzer-Ahles, S. & R. Fiorentino. 2013. The realization of scalar inferences: Context sensitivity without processing cost. *PloS ONE* 8(5).

Schwarz, Bernhard. 2016. Consistency preservation in quantity implicature: The case of at least. *Semantics and Pragmatics* 9(1). 1–47.

Westera, Matthijs & Adrian Brasoveanu. 2014. Ignorance in context: The interaction of modified numerals and QUDs. In Todd Snider, Sarah D'Antonio & Mia Weigand (eds.), *Semantics and Linguistic Theory (SALT)* 24, 414–431.

Stavroula Alexandropoulou
Trans 10
NL-3512JK Utrecht
stavroulaale@gmail.com

Jakub Dotlačil
Science Park 107
NL-1098XG Amsterdam
j.dotlacil@gmail.com

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