Journal of Semantics 27: 271–305
doi:10.1093/jos/ffq006
Advance Access publication March 18, 2010

Comparative and Superlative Quantifiers: Pragmatic Effects of Comparison Type

CHRIS CUMMINS AND NAPOLEON KATSOS
University of Cambridge

Abstract

It has historically been assumed that comparative ('more than', 'fewer/less than') and superlative ('at most', 'at least') quantifiers can be semantically analysed in accordance with their core logical–mathematical properties. However, recent theoretical and experimental work has cast doubt on the validity of this assumption. Geurts & Nouwen (2007) have claimed that superlative quantifiers possess an additional modal component in their semantics that is absent from comparative quantifiers and that this accounts for the previously neglected differences in usage and interpretation between the two types of quantifier that they identify. Their semantically modal hypothesis has received additional support from empirical investigations. In this article, we further corroborate that superlative quantifiers have additional modal interpretations. However, we propose an alternative analysis, whereby these quantifiers possess the semantics postulated by the classical model and the additional aspects of meaning arise as a consequence of psychological complexity and pragmatic implicature. We explain how this model is consistent with the existing empirical findings. Additionally, we present the findings of four novel experiments that support our model above the semantically modal account.

1 INTRODUCTION

There is an ever-burgeoning literature in semantics and pragmatics dealing with the interpretation of number terms. At first sight, the natural numbers appear ideal candidates for formalization using classical approaches borrowed from mathematical logic. However, this approach turns out to be fraught with difficulties, which continue to be debated at great length.

Until recently, it has been assumed that certain categories of numerically quantified expressions can be formalized in a fairly intuitive way. This article deals with two classes of such expressions, comparative and superlative quantifiers. Comparative quantifiers are those of the form 'more than' and 'fewer/less than' and have

traditionally been regarded as equivalent to the mathematical symbols $>$ and $<$, respectively. Superlative quantifiers are those of the form 'at most' and 'at least' and have traditionally been regarded as equivalent to the symbols $\leqslant$ and $\geqslant$, respectively. Restricting our attention to cardinalities, it follows from this approach that superlative and comparative quantifiers are interdefinable, as in the following examples.

(1) John has at most two cars ↔ John has fewer than three cars.
(2) Kelly has at least three children ↔ Kelly has more than two children.

This view is challenged by Geurts and Nouwen (2007), who argue that equivalence does not typically hold between comparative and superlative quantifiers. They observe four differences in the usage and interpretation of these types of quantifier. First, they note that superlative quantifiers admit a specific construal, absent from the comparative quantifier. Secondly, they identify differences in the patterns of inference that arise from the putatively equivalent sentences. Thirdly, they observe distributional differences between comparative and superlative forms. And finally, they claim that certain usages of comparative quantifiers give rise to ambiguity that does not follow from the superlative 'equivalent'. To address these issues, they develop a proposal in which comparative and superlative quantifiers differ in modality.

Geurts and Nouwen's proposal gives rise to several empirically testable predictions, notably that the superlative quantifiers will be mastered more slowly by acquirers, that they will be disfavoured in processing and that they will give rise to different reasoning patterns. These proposals are investigated by Geurts *et al*. (2010), in a series of experiments. Broadly, the predictions are borne out, and thus, Geurts and Nouwen's account is favoured by comparison with the classical approach.

In this article, we develop an alternative proposal to that offered by Geurts and Nouwen (2007). Rather than proposing a modal component to the semantics of superlative quantifiers, we propose that there is a fundamental difference in complexity between expressions conveying $\leqslant$ and $\geqslant$ and those conveying $<$ and $>$, which we will argue arises from the disjunctive nature of non-strict comparison. We will argue, following Büring (2007), that the use of superlative quantifiers triggers an implicature. We show that the classical model of quantifier semantics, augmented with this distinction, gives similar predictions to those made by Geurts and Nouwen, and is consistent with the data

they use to argue against the classical model. We further demonstrate that the data obtained by Geurts *et al*. (2010) are compatible with our account. We then conduct additional experiments that demonstrate the availability of inferences predicted to be unavailable by Geurts and Nouwen and the acceptability of statements they predict to be unacceptable. We show that these data are compatible with our augmented classical model. Finally, we argue that on the basis of these data, as well as considerations of parsimony and acquirability, the augmented classical model should be preferred to the semantically modal account.

2 PROBLEMS WITH THE CLASSICAL VIEW OF COMPARATIVE AND SUPERLATIVE QUANTIFIERS

As outlined above, the classical view of comparative and superlative quantifiers holds them to be interdefinable in a systematic way: 'at most n' = 'fewer than $n + 1$', 'at least n' = 'more than $n - 1$'. In addition to providing us with an elegant formal treatment of these entities, this account conforms with our naive intuitions about the truth conditions of numerically quantified statements. These intuitions require that 'John has at most two children' is false only in cases where the cardinality of the set {John's children} is 3 or more, and 'Kelly has at least three cars' is false only in cases where the cardinality of the set {Kelly's cars} is 2 or less.

However, Geurts and Nouwen (2007) identify a number of areas in which this account is unsatisfactory. They are suspicious of interdefinability on the grounds that it implies that one set of quantifiers is entirely redundant, given the existence of the other.[1] They also specify a number of additional objections to the analysis, as described in the following paragraphs.

A key objection is that the inference patterns arising from the superlative quantifiers differ from those admitted by the comparative forms. Geurts and Nouwen argue that a sentence such as (3a) gives rise to the inference (3b) but not to the inference (3c), despite (3b) and (3c) being semantically identical on the classical view.

[1] Arguably the quantifiers are not, even on the classical view, interdefinable on non-discrete sets such as are used in measurements: 'more than 2 metres' cannot be equivalent to 'at least x metres' for any x. This could license the existence of both types of quantifier, although by itself it does not explain why both are actually used in discrete cases. In this article, we follow Geurts and Nouwen (2007) in focusing on these discrete cases but note that our argument will have implications for the analysis of the more general continuous case.

(3) a. Dave had (exactly) three Martinis.
b. Dave had more than two Martinis.
c. Dave had at least three Martinis.

They support their intuitions by verifying this claim experimentally. Hence, it appears that the meanings of comparative and superlative quantifiers systematically differ in some profound way.

In addition to these, Geurts and Nouwen collate data suggesting that the distribution of comparative and superlative quantifiers differs systematically. They observe that the argument of superlative quantifiers can have a specific construal, which is not licensed by the theoretically equivalent comparative quantifiers. This provides a referent for the 'namely' clause in (4a), which is unacceptable in (4b).

(4) a. There are at most two people who have that authority, namely the Queen and the Prime Minister.
b.*There are fewer than three people who have that authority, namely the Queen and the Prime Minister.

Geurts and Nouwen (2007: 537f) also note that the superlative quantifiers have a wider range than their comparative counterparts, citing the following examples.

(5) a. Betty had three Martinis at most / *fewer than.
b. At least / *More than Betty had three Martinis.

They note that some contexts permit the comparative but not the superlative quantifier, though the prohibition is less clear-cut, as in (6).

(6) Betty didn't have ?at least / more than three Martinis.

Finally, they claim that sentences with comparative quantifiers are sometimes ambiguous in a way that those with superlative quantifiers are not. In particular, they contrast (7a) and (7b).

(7) a. You may have at most two beers.
b. You may have fewer than three beers.

Geurts and Nouwen argue that the latter (comparative) does not necessarily rule out the possibility that the addressee may have more than three beers; it merely grants permission to have a smaller number of beers. Although they admit that 'this construal may seem far-fetched' (2007: 539), they contend that this does represent a difference in

meaning between these two sentences that is not captured by the classical account.

In summary, there are reasons arising from both the interpretation and distribution of these quantifiers to support the contention that the classical view of their meaning is inadequate. In the following section, we discuss the specific proposal outlined by Geurts and Nouwen (2007) for dealing with these puzzles.

3 THE SEMANTICALLY MODAL ACCOUNT OF SUPERLATIVE QUANTIFIER MEANING

Geurts and Nouwen (2007) propose an account of quantifier meaning in which superlative quantifiers have a modal component of meaning. Specifically, they consider example (8).

(8) Betty drank at least four highballs.

They ascribe to (8) a semantic formula that can be glossed as 'the speaker is certain that there is a group of four highballs each of which was drunk by Betty, and considers it possible that Betty drank more than four highballs' (552).

For the corresponding 'at most' sentence, 'Betty drank at most four highballs', they obtain the analysis that 'it grants the possibility that Betty had four highballs, and it excludes the possibility that she had more than four' (ibid.).

They also propose that comparative and superlative quantifiers differ in argument type. Their conjecture is that superlative quantifiers accept arguments of any Boolean type, that is, both propositional and predicative arguments. In this way, they analyse 'At least it isn't raining' as an assertion 'that the speaker is sure it isn't raining, and that he considers it possible that something "better" than non-raining might be the case, as well' (ibid.).

Using their account, Geurts and Nouwen are able to give solutions to all the problems they previously discussed. From the assumption that comparative and superlative quantifiers are not interdefinable, it follows that there is no redundancy in the system. The more substantive issues raised in the preceding section may be resolved as follows.

- The inference patterns involving superlative quantifiers are a subset of those arising from comparative quantifiers because of clashes of

modality. 'At most *n*' does not imply 'at most *n* + 1', 'exactly *n*' does not imply 'at most *n*' or 'at least *n*' and so on.

- The specific construal of the argument of the superlative quantifier is possible because its argument (e.g. 'two people') may be parsed as an existential quantifier. This is not legitimate in the comparative case because the comparative quantifier does not accept non-predicative arguments.
- The validity of superlative quantifiers in a broader range of contexts than comparative quantifiers stems from their ability to accept additional argument types. At the same time, the inappropriateness of superlative quantifiers in other contexts—for example, under the scope of negation—reflects their modal semantic content.
- The ambiguity in the fourth case does not arise in the modal case due to factors involving the semantic combination of modal expressions—Geurts and Nouwen refer to this as modal concord. Put simply, their view appears to be that the superlative quantifier enters unambiguously into a concord reading with the preceding modal, as these both express possibility.

In sum, Geurts and Nouwen's modal theory of superlative quantifier meaning accounts well for the observed findings, modulo some concerns about the treatment of superlative quantifiers in conditional environments (to which we return later). In addition, their theory gives rise to empirically testable predictions. In the following section, we review the work done on investigating these predictions.

4 EMPIRICAL INVESTIGATION OF QUANTIFIER MEANING

Geurts and Nouwen (2007) subject some of their intuitions about the inference patterns arising from these quantifiers to empirical investigation. However, their theory gives rise to a broader range of predictions that are also susceptible to testing by experimental means. Geurts *et al*. (2010) argue that three particular predictions arise from the modal view of superlative quantifier usage: (i) that superlative quantifiers give rise to different inference patterns to comparative quantifiers; (ii) that superlative quantifiers should be harder to learn than comparatives, on the basis of their additional semantic complexity; and (iii) that superlative quantifiers should be harder to process than comparatives, for the same reason.

In this section we summarize the work done by Geurts *et al*. to investigate these predictions, as well as the findings of other research bearing upon these questions.[2]

4.1 *Inference patterns arising from comparative and superlative quantifiers*

Supporting their intuitions, Geurts and Nouwen (2007) performed a pencil-and-paper experiment in which they asked participants to decide whether certain implications were valid. Broadly, their participants concurred that (in Dutch) 'Beryl had three sherries' implied both 'Beryl had more than two sherries' and 'Beryl had fewer than five sherries'. By contrast, 78% of their participants rejected the implication 'Beryl had at most four sherries', and only about half accepted 'Beryl had at least three sherries'.

This technique was employed for a wider range of premise–conclusion pairs by Geurts *et al*. (2010), again in Dutch. Instead of using 'fewer than five' and 'at most four' as Geurts and Nouwen did, Geurts *et al*. used 'fewer than four' (accepted 93% of the time) and 'at most three' (accepted 61% of the time).[3] They also added three further pairs of conditions: 'at most two' → 'at most three' (14% acceptance) v. 'fewer than three' → 'fewer than four' (71% acceptance), 'at least three' → 'three' (50% acceptance) v. 'at most three' → 'three' (18% acceptance) and 'three or four' → 'at least three' (96%) v. 'two or three' → 'at most three' (93%). As these latter two comparisons do not contrast superlative with comparative quantifiers, but instead explore the effect of entailment direction, we shall not discuss them further here.

In summary, these data constitute additional evidence for the non-equivalence of comparative and superlative quantifiers in a reasoning context. Valid arguments involving comparative quantifiers seem to fail, in the opinion of the majority, when recast using superlative quantifiers that are 'classically' equivalent. Most strikingly, the inference from 'at most two' → 'at most three' succeeds in only 14% of cases, where the putatively equivalent 'fewer than three' → 'fewer than four' achieves 71% acceptance. This concurs with the prediction arising from Geurts and Nouwen's account and thus offers it empirical support.

[2] These experiments were conducted using native-speaker participants and in English unless otherwise stated.

[3] This controls for informativeness and makes it more legitimate to compare upward- and downward-entailing quantifiers.

As Geurts and Nouwen (2007) discuss, there are issues concerning the interpretation of the bare numerals in these items, which might mean 'exactly *n*' or 'at least *n*'. This distinction is critical to the validity of several of the inferences under test; for instance, 'at least three' → 'three' is false under the first reading of 'three' but is tautologous under the second. Geurts *et al*. (2010) address this in a follow-up task using 'exactly three', obtaining similar results; however, it could be argued that this modification draws the participants' attention to the under-informativeness of the non-exact statement and could bias them towards its rejection. Nevertheless, the distinction between comparative and superlative quantifiers is compellingly supported by this experiment.

4.2 *Delay in acquisition of superlative quantifiers*

Prior to Geurts and Nouwen's (2007) prediction, the acquisition of comparative and superlative quantifiers had already been compared by Musolino (2004: 26–8). In his experiment, participants are given a selection of cards with zero to four objects on them and asked to select those with 'exactly 2', 'at least/most 2' or 'more than 2'. Adults performed at or near ceiling in all conditions. However, while children (aged 4–5) were 100% accurate on 'exactly 2' and 88% accurate on 'more than', they performed at chance on the superlative quantifiers. By asking the child participants about their understanding of these terms, Musolino demonstrated that their poor performance on the superlative quantifiers was rooted in a profound lack of understanding of these quantifiers' meanings.

Geurts *et al*. (2010) further develop this line of enquiry, using a different experimental protocol. Their experiment involved presenting participants with a set of six boxes, some of which each contained a toy of a certain kind. They asked the participants to make the situation match the sentence they were about to hear, either by adding toys, removing toys or leaving the boxes as they were. The test sentences were of the form 'Q of the boxes have a toy', where Q was a numerical quantifier; the subjects were adults and children aged 11 years.

Adults performed at 100% in all conditions, while the children's performance ranged from 97% on 'more than three' to 42% on 'at most three'. 'More than' and 'at least' were easier than 'fewer than' and 'at most', respectively, and comparative quantifiers were privileged over superlative quantifiers.

Assuming that these snapshots of development are a reasonable depiction of stages in the process of acquiring comparative and

superlative quantifiers, then this experimental evidence also supports Geurts and Nouwen's hypothesis. It appears that children first master 'more than' and much later develop an understanding of the corresponding superlative quantifier 'at least'. Similarly, 'fewer than' is followed at a distance by 'at most', which has typically still not been mastered at the age of 11. Therefore, we can conclude that comparative quantifiers are indeed mastered earlier than superlative quantifiers, as Geurts and Nouwen's theory predicts.

In passing, we note that these experimental data are also coherent with input frequencies, if we take general corpora to be indicative of these. It is clear that comparative quantifiers are used substantially more frequently than superlative quantifiers: for instance, the British National Corpus (BNC) gives the frequency of their occurrence with the numerals 1–20 (in digital form) as follows.

- More than: 1243
- Less than: 698
- Fewer than: 86
- At least: 812
- At most: 4[4]

However, this is clearly not the whole story, as we have to account for these frequency trends themselves. From a modal perspective, these could be argued to arise from the difference in core meaning. Later in this article, we will be arguing that the infrequency of superlative quantifiers stems from their complexity. In either case, we do not commit to a view as to whether the order of acquisition is causally modulated by frequency, although this is not implausible.

4.3 *Delay in processing of superlative quantifiers*

Geurts *et al*. (2010) also test the online processing of comparative and superlative quantifiers by adult participants, as we discussed earlier. Recall that participants were presented with a sentence 'There are Q As' or 'There are Q Bs' and then a display in which some numbers of the letter A or B are present. They were asked to press a button to indicate whether the sentence is true or false of the situation displayed. Reading times and decision times were measured. Within the decision

[4] A reviewer of an earlier draft of this article pointed out that 'at most' is much closer to the others in frequency in a Google search, with around 30 million hits as against 450 million for 'at least'. In the BNC (2007), 'at most' occurs 478 times in all, and 'at least' 25 090 times. However, many of the Google instances of 'at most' are locative + quantifier, for example, 114 000 'at most places', 90 800 'at most stores' and numerous instances of 'at most of'. The same holds for the BNC, which is why we restrict the statistics reported to the cases of collocations with numerals.

times, there were effects of both entailment direction and quantifier type, just as for the acquisition experiment: 'at most' was the slowest to be verified, 'more than' the quickest. For the reading times, there were no significant effects. This was argued to support the hypothesis of complexity; the task involving deeper processing gave rise to evident delays in the superlative case, as predicted from Geurts and Nouwen's account.

It is also worth remarking upon the correlation between complexity/processing difficulty and age of acquisition for these quantifiers, as demonstrated by the previous experiment. The data from this experiment support a view in which the quantifiers that are acquired later are more difficult to process. Also, although the reading times are statistically not significantly different, they can be seen to pattern with the decision times, though they vary to a much less pronounced effect.

4.4 *Interim summary*

We have seen how Geurts *et al.*'s (2010) experiments bear out the predictions made by Geurts and Nouwen (2007) and thus constitute evidence in favour of the semantically modal account over the classical model. In the following section, we spell out an alternative proposal that captures the aspects of the interpretation of superlative quantifiers that Geurts and Nouwen highlighted but with a different division of labour between semantics and pragmatics than they envisioned. We further examine how this can account for the empirical data discussed above.

5 A PRAGMATIC ACCOUNT OF SUPERLATIVE QUANTIFIER MEANING

In discussing the classical model (attributed to Barwise & Cooper 1981), both Geurts and Nouwen (2007) and Geurts *et al.* (2010) have drawn no distinction between the comparative and superlative forms. From a formal point of view, this is correct; the classical view predicts that 'at most n' is true in exactly the same set of cases for which 'fewer than $n + 1$' is true. In this approach, comparative and superlative quantifiers are also equal in mathematical complexity; each corresponds to one symbol, be it $>$, $<$, $\leqslant$ or $\geqslant$. This amounts to the observation that each quantifier maps a pair of arguments to a truth-value.

However, here we wish to consider the possibility that the operators $>$ and $<$ are not equivalent to the operators $\leqslant$ and $\geqslant$ because the latter

pair can each be regarded as disjunctions, at some non-linguistic level of representation. We suggest that $\leqslant$ and $\geqslant$ are represented as '$= n$ or $< n$' and '$= n$ or $> n$', respectively. Furthermore, as per the classical model, we wish to propose that the operators $\leqslant$ and $\geqslant$ provide the semantics of natural language 'at most' and 'at least', just as the operators $<$ and $>$ provide the semantics of the comparative quantifiers. On this proposal, not only $\leqslant$ and $\geqslant$ but also natural language 'at least' and 'at most' are treated as disjunctions. The latter part of this claim is consistent with Büring's (2007) analysis whereby 'at least *n*' is interpreted as 'exactly *n* or more than *n*' and 'at most *n*' is interpreted as 'exactly *n* or fewer than *n*'. This in turn extends the semantic account proposed by Krifka (1999).

This proposal has three important implications. First, in accordance with Büring's analysis, the use of a superlative quantifier, whose meaning is a disjunction, gives rise to a quantity implicature, as a consequence of Grice's (1975) first maxim of quantity. 'At least *n*' conveys semantically that $\geqslant n$ holds and implicates that the speaker does not know (or is not at liberty to say) whether it is the case that $= n$ or $> n$ holds. Similarly, 'at most *n*' conveys semantically that $\leqslant n$ holds and implicates that the speaker does not know (or is not at liberty to say) whether it is the case that $= n$ or $< n$ holds. This is the classical clausal implicature associated with disjunction (e.g. Horn 1972), whereby by asserting that 'p or q' the speaker implicates that (s)he is not in a position to make a stronger statement, such as asserting that 'p' or that 'q'.[5] This proposal is distinct from that of Geurts and Nouwen (2007) as no specific notion of modality is stipulated as part of the semantics of superlative quantifiers. Indeed, no semantic difference is proposed between comparative and superlative quantifiers (other than the numerical difference in their arguments); the difference in modality is now captured as a pragmatic inference.

A second consequence of this proposal is that both the natural language expressions 'at least *n*' and 'at most *n*' and the logical operators $\leqslant$ and $\geqslant$ may be more difficult to process at a psychological level than

[5] As observed by Büring (2007), this proposal assumes that the implicature is triggered by a property of the expression—namely that it is disjunctive—rather than the expression itself. This is a departure from standard Gricean analyses, in which implicatures are derivable from what is said. If correct, this proposal requires a thorough reconsideration of the notion of implicature. However, note that the implicature we wish to derive could also be based on the manner maxim: a speaker must have a reason to use a complex expression (e.g. 'at most *n*', according to Büring's analysis in which this is disjunctive) when a simpler one is available (e.g. 'fewer than $n + 1$', which is analysed as a simplex proposition). In this case, a good reason to use 'at most *n*' is to highlight that 'exactly *n*' is a likely possibility. With this in mind, we leave the question of the precise nature of the implicature involved to future study.

the natural language expressions 'more than n' and 'fewer than n' and the logical operators $<$ and $>$, respectively. This is because, at a psychological level, the disjunction '$= n$ or $> n$' could be more complex than either of the disjuncts '$= n$' or '$> n$' on their own. We would expect the difference in psychological complexity to manifest itself in a usage preference for the less complex option, when both are available.

A third consequence of this proposal is that any differences in usage and meaning between superlative and comparative quantifiers are not attached to the specific forms 'at most' and 'at least' but rather that they arise because of the pragmatic implicatures and psychological complexity associated with expressions whose underlying meaning is a disjunction. Thus, our account predicts that all forms that are semantically expressible as $\leqslant$ and $\geqslant$ will pattern together and behave distinctively from the forms expressible as $<$ or $>$. This is coherent with the account offered by Nouwen (2010): his class A modifiers are those traditionally expressible as $<$ or $>$, and his class B modifiers those expressible as $\leqslant$ and $\geqslant$. However, unlike Nouwen, we do not posit a more elaborate semantic difference between these classes of modifier. We will follow Nouwen (2008) in referring to $<$ and $>$ as *strict comparison* and $\leqslant$ and $\geqslant$ as *non-strict comparison*.

Thus, we propose that by treating 'at most' and 'at least' as disjunctions, we can identify two grounds on which superlative and comparative quantifiers differ, namely the clausal implicature and the psychological complexity associated with the former. In section 6, we shall see how these differences can be invoked to explain extant experimental findings on the difference between comparative and superlative quantifiers.

In the remainder of this article, we will first present an experimental justification of the analysis of superlative quantifiers as psychologically complex and specifically disjunctive. We will discuss the explanation of extant experimental findings in terms of this proposal. We will then exhibit evidence in support of the classical view of quantifiers, augmented by this disjunctive analysis (and consequent implicatures), against the semantically modal account of Geurts and Nouwen (2007).

6 EXPERIMENT 1—EMPIRICAL INVESTIGATION OF THE COMPLEXITY OF NON-STRICT COMPARISON

To the best of our knowledge, the distinction between strict and non-strict comparison has not been studied in detail within the domain of the psychology of mathematics. However, we posit that non-strict comparison is more complex, and specifically disjunctive, on the following grounds.

(a) Assuming that the basic means of comparison between two quantities are 'more/less than' and 'equal to', $>$ and $<$ each correspond to a single simplex operation of comparison. $\leqslant$ and $\geqslant$ are each obtained by the disjunction of two simplex operations and are thus greater in complexity.

(b) The operators $>$ and $<$ are customarily glossed as 'greater than' and 'less than', while $\leqslant$ and $\geqslant$ are customarily glossed as 'less than or equal to' and 'greater than or equal to'. The absence of a non-disjunctive expression for these operators in common parlance suggests that they are naturally regarded as complex.

The above observations suggest that the claim that superlative quantifiers are disjunctive might be plausible but fall far short of making a convincing case for it. However, it is also possible to investigate this claim by adapting a technique applied to comparative and superlative quantifiers by Geurts *et al.* (2010).

In their experiment, Geurts *et al.* (2010) presented participants with a sentence on a computer screen of the form 'There are Q N Xs', where Q denotes a quantifier of the form 'exactly', 'at least', 'at most', 'more than' or 'fewer than'; N denotes a number; and X denotes a letter (either A or B). The participants were instructed to press a key once they had read and understood the sentence. They were then shown a display consisting of some number of instances of the relevant letter, either A or B, and instructed to indicate whether the preceding sentence had been true or false of this situation, by pressing the appropriate key. Over a series of 38 trials, the response times for each participant in each condition were measured and analysed.

Geurts *et al.* (2010) demonstrated a processing preference for comparative over superlative quantifiers, by demonstrating that the former gave rise to a shorter response time. If non-strict comparison is processed as a disjunction, we expect that comparisons involving strict comparison will give rise to longer response times than strict comparisons, even in the absence of any linguistically relevant content (such as comparative or superlative quantifiers). By contrast, if both strict and non-strict comparisons are equally demanding, there should be no significant effect of comparison type.

6.1 *Methodology*

In our experiment, we replicated the methodology of Geurts *et al.* (2010), as described above, with the following change. In place of sentences of the form 'There are Q N Xs', our participants read

statements of the form 'X ? N', where X denotes a letter (either A or B), N denotes a number and ? denotes a symbol (either $=$, $>$, $<$, $\leqslant$ or $\geqslant$). Each of the 38 items used by Geurts *et al*. was translated into this form (6 for equality, 8 for each of the other conditions). The display of letters proceeded as in Geurts *et al*.'s experiment, and participants were given the same instructions (substituting the word 'statement' for 'sentence').

The experiment was administered to 20 subjects, all members of the University of Cambridge, aged between 20 and 36 years. A total of 16 of the subjects were female. The results of two participants were excluded from the following analysis, one for a high error rate ($18/38 = 47\%$, compared with $30/684 = 4.4\%$ for the 18 participants analysed) and one for a slow mean response time (more than twice the mean for the 18 participants analysed). In addition, following Geurts *et al*. (2010), we excluded incorrect responses from consideration. Also, in order to minimize the effect of outliers arising through lapses in concentration, we further trimmed the data by removing from consideration any responses that exceeded [mean + 2 standard deviation (SD)] for the individual participant.[6]

6.2 *Results*

The participants' performance across the test conditions was as follows.

Condition	No. of observations	Mean response time in ms (SD)
$=$	105	982 (314)
$>$	132	1007 (369)
$<$	123	1061 (354)
$\geqslant$	139	1110 (466)
$\leqslant$	130	1131 (384)

To compare all four test conditions at once, a linear regression was performed in *R* (*R* Development Core Team 2008), using two predictor variables—'direction of entailment', which was set to +0.5

[6] Note that this is a conservative manipulation in that, under the hypothesis that there are no differences between conditions, it should apply across conditions equally; while under the hypothesis that the conditions differ, it tends to suppress those differences.

for > and ⩾ and −0.5 for < and ⩽, and 'comparison type', which was set to +0.5 for ⩾ and ⩽ and −0.5 for > and <. This analysis failed to reach significance for direction of entailment ($t = -1.08$, $P = 0.14$) but a significant effect was obtained for comparison type ($t = 2.48$, $P = 0.013$). That is, the conditions of non-strict comparison gave rise to significantly longer response times than their strict comparison counterparts.

Further pairwise comparisons were performed between the upward- and downward-entailing conditions, and between those of strict and non-strict comparison. ⩾ differed significantly from > (Student's *t*-test, $t = 2.01$, df = 269, $P < 0.05$ two tailed), while ⩽ v. < narrowly failed to reach significance ($t = 1.51$, df = 251, $P = 0.066$ one tailed).

6.3 *Discussion*

The findings of this experiment closely parallel those of Geurts *et al.* (2010). Specifically, the condition of equality is numerically the fastest to be verified. Of the remaining four conditions, the ⩽ condition is the slowest and the > condition is the fastest to be verified. There is a statistically significant main effect of non-strict comparison, supporting the hypothesis that non-strict comparison is more complex than strict comparison.

In this experiment, the effect of entailment direction was marginally non-significant, but the pattern resembled that of Geurts *et al.*, with < faster than > and ⩽ faster than <. However, neither our proposal nor the semantically modal account of Geurts and Nouwen (2007) makes clear predictions about the effect of entailment direction, so we propose to leave this matter aside in the remainder of this article.

In sum, the pattern across the five conditions is highly similar to that obtained in Geurts *et al.*'s (2010) third experiment, suggesting that the substitution of sentences involving comparative and superlative quantifiers with statements involving mathematical operators has not materially influenced the outcome.

One possible explanation of this is that participants are converting experimental items between these two formats—the mathematical expression and the numerical quantifier—in order to perform the task. If this is occurring, the time taken in one format should be a function of that taken in the other. This interpretation is, however, problematic for Geurts *et al.*'s explanation of their findings within the modal proposal of Geurts and Nouwen (2007). We reason as follows. If the sentences with numerical quantifiers are being converted into mathematical expressions, it appears that the modality present in the

superlative quantifiers is being jettisoned during this process (unless we make the ambitious and novel claim that the mathematical operators $\leqslant$ and $\geqslant$ also possess modal semantic content). Hence, it is no longer clear why superlative quantifiers should give rise to longer decision times than comparative quantifiers, especially if we make the reasonable supposition that this conversion occurs during the reading phase rather than the decision phase. Conversely, if the statements containing the operators $\leqslant$ and $\geqslant$ are being converted into sentences of the form 'at most' and 'at least', they appear to be acquiring modal content—on Geurts *et al.*'s account—which they previously lacked. That is, the conversion is not a semantically appropriate one.

The observation that response times in our study are faster than those in Geurts *et al.*'s study suggests that the conversion is more likely to be proceeding from linguistic to mathematical expressions than vice versa. If this is the case, then the modal semantic explanation for the processing delay is not viable, unless the translation from modal sentence to non-modal statement is happening during the decision phase. However, we should note that variation between participants' response times is considerable, and it is possible that the faster response times merely reflect higher performance by this set of subjects, by comparison with those who performed Geurts *et al.*'s version of the task.

For the moment, therefore, we do not wish to commit to a view on whether the above descriptions correctly characterize the participants' process in performing this task. It suffices for our purpose to draw the conclusion—licensed by the same arguments put forward by Geurts *et al.* (2010)—that non-strict comparison is more complex than strict comparison. This supports the hypothesis that non-strict comparison is treated as disjunctive. In what follows, we explore the consequences of adding this into the classical model of numerical quantifier semantics.

7 CONSEQUENCES OF THE COMPLEXITY OF NON-STRICT COMPARISON

Having presented empirical evidence for the claim that non-strict comparison is psychologically more complex than strict comparison, and argued that this is consistent with the idea that non-strict comparative forms such as superlative quantifiers are treated as disjunctions, we now consider the consequences of this for the system of numerical quantifiers in general.

One such consequence might be that the reasoning patterns arising from superlative quantifiers are not identical to those arising from

comparative quantifiers, in that the superlative quantifiers obstruct the generation of logically valid inferences. This follows because the disjunction gives rise to an implicature, as set out in Büring (2007). Recall that, for instance, 'at most' can be analysed as 'less than or equal to' and hence the use of 'at most' in declarative contexts gives rise to the implicatures that 'less than' and 'equal to' are both possibilities. The latter implicature disrupts the inferential process explored in experiment 1 of Geurts *et al*. (2010), as described below.

Consider the case discussed in section 4.1, where participants are asked whether the implication 'Beryl had at most three sherries' → 'Beryl had at most four sherries' is legitimate. According to the classical account, this should be acceptable, but it is rejected by the vast majority of participants. By contrast, in our disjunctive account of 'at most', the consequent 'Beryl had at most four sherries' gives rise to the implicature that it is possible, from the speaker's point of view, that Beryl had exactly four sherries. This directly contradicts the antecedent, 'Beryl had at most three sherries'. This has the potential to block participants' acceptance of the implication under test.

Note that this explanation is similar in character to that proposed by Geurts and Nouwen (2007). In their account, the implication 'at most n' → 'at most $n - 1$' fails because of a contradiction at the level of semantics, in that 'at most n' is held to encompass the explicit possibility of 'exactly n'. Our account differs only in that this explicit possibility arises through pragmatics rather than semantics; specifically, through the implicature derived from the use of the complex disjunctive non-strict comparison.

So far we have seen how the classical account, augmented with a more sophisticated analysis of non-strict comparison, can answer some of the criticisms developed by Geurts and Nouwen (2007) and coheres with their observations about the inference pattern, as supported empirically by Geurts *et al*. (2010). However, we also need to demonstrate that this account is equally compatible with the other experimental data discussed earlier.

Geurts *et al*.'s (2010) second experiment bore out the prediction that the comparative quantifiers are mastered earlier in acquisition than their superlative counterparts. Here, once again, our enhanced version of the classical account makes the same prediction as the semantically modal account of Geurts and Nouwen (2007). As we argue that superlative quantifiers possess a more complex meaning than comparatives, in that they are represented disjunctively, we also predict that superlative quantifiers are disfavoured in acquisition. Hence, this experiment does not distinguish between the two competing proposals.

The third experiment conducted by Geurts *et al*. (2010) shows faster verification times for comparative over superlative quantifiers, as discussed in section 6 of this article. We demonstrated in that section that the same applies for strict v. non-strict comparison, in the absence of the specific linguistic constructs under investigation (comparative and superlative quantifiers). Given this finding, we would predict that the same pattern should be replicated whenever strict and non-strict comparisons are in competition. Therefore, our proposal is compatible with the finding that comparative quantifiers are verified faster than superlative quantifiers. Moreover, we might argue that this proposal is more parsimonious than the semantically modal account of Geurts and Nouwen (2007), in that it explains the outcomes both of Geurts *et al*.'s (2010) third experiment, for the quantifier case, and our first experiment (section 6), for the comparison case. We have already concluded that the superior performance on strict rather than non-strict comparison cannot readily be attributed to a semantic difference between comparative and superlative quantifiers, as participants do not appear to be invoking these quantifiers in their verification process.

Hence, with reference to the major lines of argument made by Geurts and Nouwen (2007), and the experimental evidence adduced in support of this proposal by Geurts *et al*. (2010), we have seen that our augmented version of the classical account is equally well supported as the semantically modal account. In the following sections, we present data from a further set of experiments, first replicating Geurts *et al*.'s findings on the inference judgment task and then obtaining support for our augmented classical account over and above Geurts and Nouwen's proposal.

8 EXPERIMENT 2—JUDGMENTS OF LOGICAL INFERENCE PATTERNS

In our next experiment, we set out to replicate the findings of Geurts *et al*. (2010) with respect to the inference judgment task. We performed this experiment in order to verify that the patterns observed for Dutch quantifiers were also to be found in English. We also administered a post-test questionnaire to ascertain whether participants had explicit knowledge of any differences between the two types of quantifier.

8.1 *Method*

A total of 15 adult participants were recruited, all of whom were students at the University of Cambridge. None had any university-level

background in mathematics or logic. They were informed that they would see a series of pages, each with two sentences written on them, and that they should circle the answer 'yes' if the first sentence implied the second and 'no' if it did not. Three sentence pairs were used for each of 12 conditions (36 sentence pairs in all); these included the first eight conditions tested by Geurts *et al.* (2010) and four additional 'false' conditions as controls. The order of these sentence pairs was randomized and the same order used for each participant.

As a post-test, participants were asked to write a brief explanation of why they had answered the way they did, for the first instance of each of the 12 conditions.

8.2 *Results*

The following table presents the acceptance rates for the implication judgment task and compares these rates with those obtained by Geurts *et al.* (for the eight conditions common to both studies).

1st quantifier	2nd quantifier	Acceptance (%)	Geurts *et al.* acceptance (%)
3	At least 3	62	50
3	More than 2	100	100
3	At most 3	42	61
3	Fewer than 4	84	93
At most 2	At most 3	2	14
Fewer than 3	Fewer than 4	64	71
At least 3	3	90	50
At most 3	3	26	18
More than 3	Fewer than 3	9	N/A
Fewer than 3	More than 3	0	N/A
3	Fewer than 3	13	N/A
3	More than 3	0	N/A

The metalinguistic judgments elicited from the questionnaire were generally uninformative. These exhibited a degree of uncertainty about some of the conditions but provided no clear indications of any explicit awareness of modality or other semantic effects. We shall not discuss these further in what follows.

8.3 *Discussion*

The results for the conditions common to both this experiment and Geurts *et al.*'s (2010) were generally very similar, with the exceptions of the 'three' → 'at most three' and, most strikingly, the 'at least three' → 'three' pairs. We tentatively attribute this to item effects, arising from the potential ambiguity of the bare numeral.[7] Rejection rates for the semantically incorrect control conditions were generally at or near ceiling.

Crucially for our purposes, this experiment constitutes a replication of Geurts *et al.*'s (2010) study as far as the comparative and superlative quantifiers are concerned. We can see that, in English as in Dutch, performance on comparative quantifiers exceeds that on superlatives—judged by traditional standards of logical correctness—in all the cases for which we have comparable data. Thus, we conclude that, as expected, this is not an effect specific to Dutch.

9 EXPERIMENT 3—COMPATIBILITY JUDGMENTS OF NUMERICALLY QUANTIFIED EXPRESSIONS

Following Geurts *et al.*'s analysis, the rationale for the failure of 'at most two' → 'at most three' in experiment 2 was as follows. On the modal view, 'at most three' admits the possibility of '(exactly) three'. 'At most two' uncontroversially excludes this possibility. Therefore, as the possibility of '(exactly) three' cannot follow from 'at most two', the implication fails.

Following this line of analysis, it is predicted that two such sentences will be judged as logically contradictory. By contrast, the augmented classical view predicts that the two sentences will be logically compatible but pragmatically infelicitous when juxtaposed. We investigated this issue using a method introduced by Katsos (2007: chapter 3; see Katsos 2008 for a review) with the aim of capturing the difference between logical contradiction and pragmatic infelicity for the case of scalar implicature.

[7] Note that 'at least three' → 'three' is valid if and only if the latter 'three' is interpreted as existential rather than precise, in which case it is a tautology. In our experiment, the consequent sentences in this case were 'Anna wrote 3 letters', 'There are 3 cities on the map' and 'Steve owns 3 suits', all of which could plausibly be existential statements. For items such as 'Dave has 3 children', we might expect lower rates of acceptance.

In this paradigm, participants are presented with statements and they are asked to give coherence judgments on a scale. Under this methodology, we predicted that semantically self-contradictory statements would be judged as incoherent, while statements that were pragmatically self-contradictory (i.e. in which an implicature is explicitly revised) would be judged more coherent, and statements with neither type of self-contradiction would be judged more coherent still. The modal view of superlative quantifiers, holding that 'at most *n*' semantically conveys the possibility of 'exactly *n*', predicts that a statement containing 'at most *n*' and 'exactly $n - 1$' (with reference to the same entities) should pattern with the semantically self-contradictory statements. Our proposal predicts that a statement containing 'at most *n*' and 'exactly $n - 1$' should pattern either with the pragmatically self-contradictory statements (if an implicature is generated) or with the non-self-contradictory statements (if it is not).

9.1 *Methodology*

A total of 20 participants were recruited, all members of the University of Cambridge, in the age range 20–36 years. Fourteen were female. Participants were presented with a pair of sentences linked by the word 'specifically', as in the following example, where Q denotes a quantifier and *n* and *m* denote numbers.

(9) Jean has Q *n* houses. Specifically, she has exactly *m* houses.

They were asked to give a judgment on the coherence of the utterance, rating it on a Likert scale ranging from 5 ('coherent') to −5 ('incoherent').

Two types of control items were included. There were items using 'In fact', rather than 'Specifically', partially to disguise the goal of the experiment and partially to test whether participants' judgments would differ if the second sentence could be interpreted as a weakening of the speaker's commitment to the proposition originally expressed. There were also items in which the quantifier and numeral in the first sentence were replaced with the quantifier 'some', and the numeral in the second sentence was replaced with 'none of', 'half of' or 'all of'.[8] This tested the participants' response to violations of both semantic and pragmatic contradiction, as discussed above.

In total, 78 items were used. The value of *n* was varied over the range 3–5. The results are summarized below.

[8] The sentences were chosen in such a way as to license the partitive usage.

9.2 *Results*

The following table shows the mean judgment of coherence, and the corresponding SDs, in each of the experimental conditions.

	Quantifier in first sentence	Quantifier in second sentence	'Specifically' condition		'In fact' condition		Coherent?
			Mean	SD	Mean	SD	
1	At most *n*	Exactly $n - 1$	1.58	2.57	1.87	2.53	?
2	At most *n*	Exactly *n*	1.90	2.31	1.25	2.60	?
3	At most *n*	Exactly $n + 1$	−4.08	2.34	−4.05	2.10	No
4	At least *n*	Exactly $n - 1$	−4.48	1.50	−4.27	1.88	No
5	At least *n*	Exactly *n*	1.28	2.50	1.33	2.56	?
6	At least *n*	Exactly $n + 1$	1.95	2.53	2.55	2.16	?
7	More than *n*	Exactly $n - 1$	−4.70	0.93	−4.28	1.92	No
8	More than *n*	Exactly $n + 1$	3.10	2.18	3.20	1.90	Yes
9	Fewer than *n*	Exactly $n - 1$	3.08	1.80	3.13	1.93	Yes
10	Fewer than *n*	Exactly $n + 1$	−4.75	0.73	−4.23	2.05	No
11	Some	None	−4.60	1.14	−4.07	1.97	No
12	Some	Half	3.08	2.23	3.35	2.00	Yes
13	Some	All	−1.08	3.13	0.22	3.10	?

The global mean across all conditions was −0.64 and the SD 3.88.

The semantically self-contradictory cases achieved low ratings in both conditions. The semantically uncontroversially non-contradictory cases (8, 9 and 12) achieved high ratings. The 'in fact' condition was regarded as more coherent than the 'specifically' condition for 12 out of the 13 cases (significant by the sign test, $P < 0.01$). We will discuss the 'specifically' condition for the following cases.

Our first critical case was condition 1, 'at most n ... exactly $n - 1$'. Comparing condition 1 with condition 2, 'at most n ... exactly n', a Student's t-test gives us $t = 0.72$, df = 118, $P > 0.1$. Hence, there is no significant difference between these conditions. Comparing condition 1 with condition 3, the semantically false control 'at most n ... exactly $n + 1$', we obtain $t = 12.6$, df = 118, $P < 0.01$. Thus, condition 1 is highly significantly better than the relevant semantically false control.

For the 'at least' case, we compared condition 5 with condition 6, and obtained $t = 1.46$, df = 118, $P > 0.1$. Again, there is no significant difference between these cases. Both are highly significantly ($p < 0.01$) more acceptable than the semantically false control condition 4.

Comparing the superlative with the comparative quantifiers, we find that the semantically true comparative conditions (8 and 9) yield significantly higher results than the corresponding superlative conditions (again, $P < 0.01$ for all comparisons).

For the control 'some' conditions, further pairwise comparisons reveal a significant preference for 'some ... half' over 'some ... all' and 'some ... all' over 'some ... none' ($P < 0.01$ for each). 'Some ... half' also outperforms conditions 1, 2, 5 and 6, which in turn outperform 'some ... all' (all $P < 0.01$).

9.3 *Discussion*

Broadly, the trends in both conditions are clear. The statements in which pairs of sentences are semantically compatible, on a classical semantic view, are systematically judged to be coherent. Those in which the sentences are semantically contradictory are judged incoherent. This trend plays out over both the 'in fact' and the 'specifically' cases, although understandably participants appear to show slightly more leniency to self-contradictory utterances in the 'in fact' case. The results for the comparative cases license the assumption that this test is diagnostic for semantic coherence.

For our purposes, the critical data are those in which the first sentence contains a superlative quantifier. In these cases, there is once

again a clear-cut division between the cases that are semantically self-contradictory and those that are not. Crucially, given a first sentence containing 'at most *n*', participants accept both a continuation 'exactly *n*' and 'exactly *n* – 1', with no statistically significant preference for one or the other.

This fails to confirm the predictions arising from the semantically modal account of superlative quantifiers. On this account, the superlative quantifier encodes the possibility of exact equality, a possibility which in our critical cases is then denied by the second sentence. Nevertheless, the participants accept the relevant utterances as clearly coherent rather than incoherent. Under the hypothesis of modality in the semantics, the lack of significant difference between conditions 1 and 2 indicates that revising possibility to certainty is equally incoherent to revising possibility to impossibility, which appears implausible. Notably, revising possibility to impossibility is judged significantly more coherent than revising 'some' to 'all', which is a case of pragmatic self-contradiction. We therefore consider these data incompatible with the hypothesis that 'at most' semantically encodes the possibility of equality.

According to our competing proposal, the superlative quantifiers give rise to modal interpretations due to pragmatic implicature. Specifically, we are proposing that the 'at most *n*' first sentence gives rise to an implicature that 'exactly *n*' is possible. Therefore, we would predict that the 'at most *n* . . . *n* – 1' case should be more comparable to the 'some . . . all' case, where such incoherence as exists arises from the second sentence contradicting an implicature of the first, than to the semantically contradictory cases. This is indeed the case in our data: 'some . . . all' is judged significantly more acceptable than 'some . . . none' and the other contradictory cases, as does 'at most *n* . . . *n* – 1'. Both are judged less acceptable than the non-contradictory cases with the comparative quantifiers (that do not trigger an implicature), as we predicted.

Furthermore, our account also suggests that 'at most *n*' should give rise to an implicature that 'fewer than *n*' is possible, which in turn predicts that condition 2 will yield lower ratings than the comparative conditions and similar ratings to those elicited in condition 1. Our results suggest that this is indeed the case.

Again, we note that these forms are judged to be significantly more coherent than the 'some . . . all' case, even though we are arguing that these are both self-contradictory at a pragmatic level. However, there is a sizeable gulf between these and the control semantically self-contradictory cases. Put simply, the non-theory-critical cases present

a clear pattern: the semantically self-contradictory cases are rated low, the pragmatically self-contradictory cases towards the middle of the scale and the non-self-contradictory cases are rated higher. If we measure the superlative quantifier cases against this scale, we see them clearly as pragmatically self-contradictory or fully non-self-contradictory. Either way, this contradicts the semantically modal account. Within our pragmatic account, we would suggest that the intermediate status of these results indicates implicature being generated, but less reliably or robustly than in the paradigm case of 'some' ('but not all').

In summary, the coherence judgments elicited in this experiment point to a clear division between semantically self-contradictory and non-self-contradictory utterances, with pragmatically self-contradictory utterances occupying the middle ground. The case of 'at most n ... exactly $n - 1$', which should be semantically self-contradictory on the semantically modal account, appears to be either pragmatically self-contradictory or fully coherent, in accordance with our augmented classical account. In fact, as predicted by our account, it patterns similarly to 'at most n ... exactly n'. Therefore, we conclude that this experiment favours the augmented classical over the semantically modal account.

10 EXPERIMENT 4—INFERENCE PATTERNS IN A CONDITIONAL CONTEXT

The results of experiment 3 raise the question of whether the inference 'at most two' to 'at most three', rejected in experiment 2, is available under the right conditions. In experiment 4, we aimed to elicit this inference, using a conditional context.

Intuitively, it seems that the superlative quantifiers behave like comparative quantifiers under the scope of conditionals, for example, 'If Berta had at least/most three drinks ...'. It was noted by Geurts and Nouwen (2007) that their semantically modal account is unsatisfactory with regard to these and certain other contexts. In this experiment, we investigated whether hearers interpret an utterance such as (10) as a commitment on the part of the speaker to the corresponding proposition (11).

(10) 'If Berta has had at most three drinks, she is fit to drive. Berta has had at most two drinks.'

(11) Berta is fit to drive.

This conclusion appears to be licensed by the inference from 'Berta has had at most two drinks' to 'Berta has had at most three drinks'. If

the semantically modal account is correct, this inference should not be available and the conclusion cannot be drawn. If the classical account is correct, this inference is available and the conclusion can be drawn unless it is blocked by an implicature.

Similarly, the conclusion would be licensed by an inference from 'If Berta has had at most three drinks, she is fit to drive' to 'If Berta has had at most two drinks, she is fit to drive'. If the semantically modal account is correct, this inference is not available either, as a modal interpretation of either 'at most three' in the antecedent or 'at most two' in the consequent would render it false. If the classical account is correct, then again this inference is available and the conclusion can be drawn unless it is blocked by an implicature.

10.1 *Methodology*

Eight participants were recruited and were given the test as a questionnaire. There were 14 items in all, three instances of the critical 'at most two'/'at most three' case and one instance of each of the other 11 conditions used in experiment 1. Participants were presented with an utterance of the type given above and then asked, 'Does the speaker believe that . . . ?' They were invited to respond 'yes', 'no' or 'don't know'. 'No' and 'don't know' responses were both treated as negative with regard to the implication being investigated.

10.2 *Results*

23 of 24 (96%) responses in the 'at most two'/'at most three' case were positive, agreeing that the inference was valid. Responses in the other conditions also patterned with the mathematically valid outcomes.

10.3 *Discussion*

Under the conditions of this experiment, it does appear that the inference from 'at most two' to 'at most three' (or the inference from 'if . . . at most three' to 'if . . . at most two') goes through, contrary to the predictions we might expect to draw from the modal hypothesis. This is, however, consistent with the classical account.

This outcome is also consistent with our augmented classical account in which 'at most' is considered disjunctive ('less than or equal to') because the implicature arising from this is clearly different in a conditional environment to that arising in a declarative environment.

For instance, the utterance (12) does not give rise to the pragmatic interpretation (13).

(12) 'If Berta has had at most three drinks, she is fit to drive.'

(13) If it is possible that Berta has had at most three drinks and certain that she has had no more than three drinks, then she is fit to drive.

Why not? Because (13) is a weakening of the original semantics of the utterances, namely (14).

(14) If it is certain that Berta has had no more than three drinks, she is fit to drive.

Therefore, deriving the implicature gives the hearer no additional information. Even if the hearer derives the implicature, this will not supersede the existing semantic content but stands alongside it. Therefore, the pragmatic interpretation of the utterance remains (14), rather than the weaker (13). Rather than triggering a pragmatic enrichment, it seems intuitive to propose that the signal sent by using the 'at most' formulation in this case is to draw attention to the upper bound [see Nouwen (2010) for a semantically oriented account along similar lines].

To verify this observation, one can check the parallelism with the paradigmatic case of a possibility implicature arising from disjunction. The declarative 'There is beef or pork on the menu' appears to convey the implicature that either is possible. By contrast, this implicature is not available in a conditional context, such as 'If there is beef or pork on the menu, Max will be happy'. This is presumably because the pragmatic 'enrichment' of possibility, when applied to the latter utterance, in fact makes it less informative, in that it imposes an extra condition (possibility of each conjunct) that would have to be satisfied before the conclusion could be drawn. Hence, this implicature does not go through under the scope of a conditional.

It might be argued that the materials used in this experiment are slightly awkward from a pragmatic standpoint. A more naturalistic formulation would be along the lines of (15).

(15) Anyone who has had at most three drinks is fit to drive. Berta has had at most two drinks.

However, in this case, it might conceivably be argued that the inferential process is not clear, as it requires an additional step. The additional step might either be 'Berta has had at most three drinks' or 'Anyone who has had at most two drinks is fit to drive'. We do not believe that anything hinges on this distinction, but we elected to use

the methodology described above in order to allay any concerns arising from this uncertainty.

The result of this experiment is not entirely surprising in the light of Geurts and Nouwen's disclaimer about the legitimacy of the modal interpretation of superlative quantifiers. The results can be explained by assuming that the superlative quantifier under the scope of the conditional has a purely classical meaning: it is not necessary to assume further that the superlative quantifier in the logical antecedent ('Berta has had at most two drinks') also lacks modality.

However, there does not appear to be any specific proposal to account for the absence of this type of semantic modal meaning within the conditional environment, apart from the classical view, which denies the existence of the modal meaning *in toto*. It is not impossible to condition on modals, although it is rarely necessary and examples of it are consequently somewhat contrived (e.g. 'If I think the project *might* succeed, I'll fund it for a trial period only').

For this reason, we consider that the outcome of this experiment lends support to an augmented classical account, in which superlative quantifiers are analysed as disjunctions, over the Geurts and Nouwen (2007) account in which these quantifiers possess modal semantics.

These results also motivate us to consider whether the same kind of inference might be available in non-conditional contexts, for which Geurts and Nouwen do not identify a problem with their account. This is explored in the following task.

11 EXPERIMENT 5—JUDGMENTS OF LOGICAL INFERENCE PATTERNS IN FELICITOUS CONTEXTS

Experiment 2 demonstrated the unavailability of certain classically correct inference patterns involving superlative quantifiers. Under the semantically modal account, this arises for semantic reasons. Under the augmented classical account, we have argued that this stems from the implicatures that arise from consequents containing superlative quantifiers.

In experiment 5, we attempt to differentiate these predictions by embedding the reasoning task in experiment 2 within a theoretically more complex task in which the superlative quantifiers are licensed by the context. For example, we ask participants to judge whether (16) implies (17).

(16) Anne has three children but Brian has at most two children.

(17) Anne and Brian each have at most three children.

Assuming that this sentence is correctly analysed as predicating 'having at most three children' of the two individuals 'Anne' and 'Brian', we argue that acceptance of this implication requires acceptance that 'at most two' implies 'at most three' (in addition to acceptance that 'three' implies 'at most three'). Recall that on the semantically modal account this is not available, so the inference should fail. On the augmented classical account, it is available, so the inference should succeed unless it is blocked by implicature.

11.1 *Methodology*

A total of 20 participants were recruited, members of the University of Cambridge, in the age bracket 20–36 years. 14 were female. The methodology of experiment 2 was replicated using 32 pairs of sentences of the type discussed above, using a range of numerical quantifiers. The numerals and other aspects of the sentential content were varied between conditions.

11.2 *Results*

Acceptance rates for the test conditions (those for which multiple items were tested) were as follows.

	First sentence	Second sentence	Implication acceptance (%)
1	n . . . at least $n + 1$	At least $n + 1$	0
2	n . . . at least $n + 1$	At least n	88
3	n . . .at most $n - 1$	At most n	68
4	n . . .at most $n - 1$	At most $n - 1$	7
5	n . . .fewer than n	Fewer than n	3
6	n . . .fewer than n	Fewer than $n + 1$	100
7	n . . .more than n	More than n	3
8	n . . .more than n	More than $n - 1$	95

In addition, results for the filler items were in accordance with semantic expectations. Across all semantically uncontroversial

conditions (i.e. conditions that are not theory-critical for either approach under discussion),[9] participants were correct on 488 out of 500 items (97.6%).

Crucially, the acceptance rate for the theory-critical implication '*n* ... at most $n - 1$' → 'at most *n*' items is significantly above 50% (41/60, $P < 0.01$ binomial), that is, neither floor nor ceiling.

11.3 *Discussion*

This finding contradicts the prediction of the semantically modal account of superlative quantifier meaning. According to the modal view, as expounded in the interpretation of experiment 1 of Geurts *et al*. (2010), the implication from 'at most $n - 1$' to 'at most *n*' is unavailable. By contrast, the results of this experiment indicate that this implication is available in certain declarative contexts, albeit less reliably than the corresponding implication with comparative quantifiers.

Contrastingly, these data are compatible with our alternative proposal in which 'at most' can be analysed as a disjunction and give rise to possible implicatures. According to this proposal, 'at most three' serves the function of drawing particular attention to the possibility that 'exactly three' holds and has the potential to give rise to an implicature to that effect. However, in this situation, the implicature is not entirely inappropriate because the possibility of equality holds for one of the conjuncts, namely Anne [according to the implicature arising from (21)]. By contrast, on the semantically modal account, this possibility is part of the semantics of 'at most three' and should thus be required of both conjuncts.

To put it another way, the majority of our participants are able to reason from 'at most two' to 'at most three' when the implicatures arising from 'at most' can be accommodated without a contradiction arising. By contrast, on the semantically modal account, this should not be possible, as the modality renders 'at most three' inherently contradictory of 'at most two'.

Is there a way to reconcile these data with the modal hypothesis? We do not think so. To do this, we would have to accept that the modal semantics of 'at most three' also does not apply to the individual conjuncts of (17), as repeated below.

(17) Anne and Brian each have at most three children.

Instead, we would have to hold that the modal possibility applies only to some member of the set {Anne, Brian} and is not binding upon

[9] This includes conditions 1 and 4–8 in the table above.

the other. In this example, that amounts to saying that (17) is true of a situation in which Anne might have three children, even if Brian cannot possibly have that many.

While this analysis is theoretically viable, we find it unsatisfactory. If the above manipulation is accessible to speakers, they should also attest the truth of such statements as (18a), or under the right circumstances even (18b) and (18c), and so on.

(18) a. Anne and Brian each might have three children.
b. Anne and Brian each might be a young woman.
c. Anne and Brian each might be the only woman named Anne in the village.

As these are intuitively false statements, we submit that 'each' does not admit this analysis and therefore that experiment 5 gives evidence against the semantically modal account and in favour of the augmented classical account.

12 GENERAL DISCUSSION

Geurts *et al*. (2010) demonstrate empirical support for the hypothesis that the superlative quantifiers possess a modal component to their semantics, as proposed by Geurts and Nouwen (2007). In this article, we have robustly corroborated the finding that superlative quantifiers possess a modal component. We further proposed that the appropriate locus of the modality is in the pragmatics—and that the differences in the use of superlative quantifiers are due to this implicature as well as the psychological complexity of non-strict comparison relative to strict comparison. In order to document that non-strict comparison is more complex than strict comparison because it is disjunctive, we reported fresh experimental evidence on mathematical quantifier processing. We then argued that the classical account augmented by a disjunctive analysis for the non-strict comparison of the superlative case is equally compatible with the pre-existing empirical data as the semantically modal account. In a further series of experiments, we have elicited data to suggest that our proposal is in fact more satisfactory than Geurts and Nouwen's semantically modal account of superlative quantification.

Again we must mention that Geurts and Nouwen (2007) are aware of their theory's limitations in its present form, noting in particular the difficulties posed by the conditional context, for example, 'If Berta has had at most three drinks . . .'. In experiment 4, we provide a practical demonstration of this lacuna and show how the non-modal superlative

quantifier enters into logical relations. However, in experiments 3 and 5, we go further. In experiment 3, we show that participants treat the 'exactly *n* is possible' meaning of superlative quantifiers (with particular reference to 'at most', which gives rise to the clearest predictions) as a pragmatic readily revisable inference rather than as part of the logical meaning of the expression. In experiment 5, we show that the superlative quantifiers systematically appear to lack modal semantics but that they give rise to implicatures that are similar but—in certain contexts—distinguishable from this modal semantic meaning.

We contend that the pattern of results observed across our set of experiments coheres with the notion that the semantic meaning of superlative quantifiers is fundamentally the classical meaning. The 'modal' meaning can instead be analysed as a possibility implicature arising from the use of a superlative quantifier with its disjunctive meaning. This means that there are indeed differences in the inferences that are licensed by superlative and comparative quantifiers, but they follow from the presence or absence of this implicature. Moreover, our demonstration that the online performance preferences for comparative over superlative quantifiers are matched by an online performance preference for strict over non-strict comparison is exactly what one would expect if non-strict comparison is more complex (because it is disjunctive) than strict comparison in general.

Under this analysis, we might argue that the communicative effect of 'at most *n*' is to draw particular attention to the number *n* and the possibility of equality, neither of which are emphasized by the comparative alternative 'fewer than $n + 1$'. In declarative contexts, the implicature that the superlative quantifier conveys is very similar to the semantic meaning proposed by Geurts and Nouwen (2007). In downward-entailing contexts, by contrast, the implicature does not arise for standard pragmatic reasons, whereas the semantic meaning proposed by Geurts and Nouwen cannot be explained away in such a principled fashion.

Corpus data support the notion that superlative quantifiers achieve this effect. In the BNC, for instance, tokens of 'at least 20' outnumber those of 'more than 19' by 110 to 6. However, tokens of 'more than 20' outnumber 'at least 21' by 357 to 23. Similar patterns can be observed for round numbers flanked by non-round numbers, both in the BNC and in other corpora. This is a surprising observation within a modal semantic account of superlative quantifier meaning, as it appears to suggest that the use of superlative quantifiers is sometimes motivated by something other than the wish to express modality. Our augmented classical account is compatible with this notion, given that the use of

a round number licenses the use of a superlative quantifier, as the value being called attention to is a salient one. By contrast, imputing modality to superlative quantifiers curtails the expressive power of the system, by constraining our choice of number unless we are indifferent to whether or not we use a modal expression.

This would explain why, in our experiment 5, the logical implication from 'at most two' to 'at most three' is accepted, while it was rejected in experiment 2; in experiment 5, 'at most' is used (in the consequent) in conjunction with a salient number as determined by the preceding utterance. Under these circumstances, the superlative quantifier in the consequent is licensed and no implicature arises. In the absence of this implicature, the superlative quantifier possesses purely classical meaning and the logically correct inference is drawn.

This proposal has implications for the analysis of other expressions. For example, consider the distribution of 'not more than' and 'not fewer/less than'. It could be argued that these are modal, but it seems more plausible, compositionally speaking, to count them as classically semantic comparative quantifiers. The relevant semantic difference between these and their non-negated counterparts is that while 'more/fewer than *n*' excludes the possibility of *n*, 'not more/fewer than *n*' admits this possibility. It is simply this difference that appears to underlie the following differences in distribution.

(19) a. *Fewer than three people have that authority, namely the Queen and the Prime Minister.
b. ?Not fewer than two people have that authority, namely the Queen and the Prime Minister.
c. Not more than two people have that authority, namely the Queen and the Prime Minister.

(20) a. *Wilma danced with fewer/less than every second man who asked her.
b. ? Wilma danced with not less than every second man who asked her.
c. ? Wilma danced with not more than every second man who asked her.

In (19b) and (19c), explicit reference to 'two people' licenses the 'namely' continuation; in (20b) and (20c), the possibility of equality seems to license the use of the comparative quantifier in a context in which it is otherwise understood to be forbidden. In both cases, the difference resides in classical semantic properties of the quantifiers.

More could be said about this, and indeed its relation to the 'no more/fewer than' construction discussed by Nouwen (2008), but

like Geurts and Nouwen (2007: 37f), we take the view that the definitive answer to the question of superlative and comparative quantifier distribution will arise from a more general theory of comparison.

Acknowledgements

We thank Bart Geurts, Rick Nouwen and two anonymous reviewers of the *Journal of Semantics* for providing substantially helpful comments on the manuscript. Chris Cummins acknowledges the support of a University of Cambridge Domestic Research Studentship and of funding from COST Action A33 for a Short Term Scientific Mission to the University of Nijmegen (COST-STSM-A33-4692). Napoleon Katsos acknowledges the support of the Arts and Humanities Research Council (Ref.: AH/E002358/1).

CHRIS CUMMINS
Research Centre for English and Applied Linguistics
University of Cambridge
English Faculty Building
9 West Road
Cambridge CB3 9DP
United Kingdom
e-mail: crc37@cam.ac.uk

NAPOLEON KATSOS
Research Centre for English and Applied Linguistics
University of Cambridge
English Faculty Building
9 West Road
Cambridge CB3 9DP
United Kingdom
e-mail: nk248@cam.ac.uk

REFERENCES

Barwise, J. & R. Cooper. (1981), 'Generalized Quantifiers and Natural Language'. *Linguistics and Philosophy* **4:**159–219.

The British National Corpus, version 3 (2007), Distributed by Oxford University Computing Services on behalf of the BNC Consortium (http://www.natcorp.ox.ac.uk/).

Büring, D. (2007), 'The least "at least" can do'. In C.B. Chang and H.J. Haynie (eds.), *Proceedings of WCCFL 26*. Cascadilla Press. Somerville, MA. 114–20.

Geurts, B. (2006), 'Take "five": The meaning and use of a number word'. In S. Vogeleer and L. Tasmowski (eds.), *Non-definiteness and Plurality*. John Benjamins. Amsterdam. 311–29.

Geurts, B, N. Katsos, C. Cummins, J. Moons., & L. Noordman. (2010), 'Scalar Quantifiers: Logic, Acquisition and Processing'. *Language and Cognitive Processes* **25:**130–48.

Geurts, B., & R. Nouwen. (2007), 'At least et al.: The semantics of scalar modifiers'. *Language* **83:**533–59.

Grice, H. P. (1975), 'Logic and conversation'. In P. Cole and J.L. Morgan (eds.), *Syntax and Semantics*, Academic Press. New York. 41–58.

Horn, L. (1972), *On the Semantic Properties of Logical Operators in English*. UCLA dissertation, distributed by Indiana University Linguistics Club.

Katsos, N. (2007), Experimental Investigations on the Effects of Structure and Context on the Ph.D. thesis, Generation of Scalar Implicatures. University of Cambridge.

Katsos, N. (2008), 'The Semantics/Pragmatics Interface from an Experimental Perspective: The Case of Scalar Implicature'. *Synthese* **165:**358–401.

Krifka, M. (1999), 'At least some determiners aren't determiners'. In K. Turner (ed.), *The Semantics/Pragmatics Interface from Different Points of View*, Current Research in the Semantics/Pragmatics Interface. Vol. 1. Elsevier. Oxford. 257–92.

Musolino, J. (2004), 'The Semantics and Acquisition of Number Words: Integrating Linguistic and Developmental Perspectives'. *Cognition* **93:**1–41.

Nouwen, R. (2008), 'Upperbounded No More: The Exhaustive Interpretation of Non-strict Comparison'. *Natural Language Semantics* **16:** 271–95.

Nouwen, R. (2010), 'Two Kinds of Modified Numerals'. *Semantics and Pragmatics* **3:**1–41.

R Development Core Team (2008), *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Austria. ISBN 3-900051-07-0 (http://www.R-project.org).

First version submitted: 28.05.2009
Second version submitted: 13.10.2009
Third version submitted: 12.02.2010
Accepted: 19.02.2010