Superlative quantifiers and epistemic interpretation of disjunction

Maria Spychalska

Ruhr-University Bochum, Department of Philosophy

Abstract. We discuss semantics of superlative quantifiers *at most n* and *at least n*. We argue that the meaning of a quantifier is a pair specifying a verification and a falsification condition for sentences with this quantifier. We further propose that the verification condition of superlative quantifiers should be interpreted in an epistemic way, that is as a conjunctive list of possibilities. We also present results of a reasoning experiment in which we analyze the acceptance rate of inferences with superlative and comparative quantifiers in German. We discuss the results in the light of our proposal.

1 Introduction

There is an ongoing debate (Look inter alia: (Geurts & Nouwen, 2007), (Koster-Moeller et al, 2008), (Geurts et al., 2010), (Cummins & Katsos, 2010), (Nouwen, 2010), (Cohen & Krifka, 2011)) concerning the right semantical interpretation of so-called superlative quantifiers, such as *at most n* and *at least n*, where *n* represents a bare numeral, e.g *two*. Generalized Quantifier Theory (referred here as a "standard account") defines superlative quantifiers as equivalent to respective comparative quantifiers: *fewer than n* and *more than n*, that is:

$$at\, most\, n\, (A, B) \iff fewer\, than\, n + 1(A, B)^{1} \tag{1}$$

$$at\, least\, n\, (A, B) \iff more\, than\, n - 1(A, B) \tag{2}$$

It has been observed that in natural languages those equivalences (1) and (2) might not hold, or at least they might not be accepted by language users based on pragmatical grounds. There are numerous differences between comparative and superlative quantifiers involving their linguistic use, acquisition, processing and the inference patterns in which they occur. First of all, it seems that superlative and comparative quantifiers are not freely exchangeable in same linguistic contexts. Geurts & Nouwen (2007) provide, among many others, such examples:

(a) I will invite at most two people, namely Jack and Jill.
(b) I will invite fewer than three people, namely Jack and Jill.

[1] Q(A,B) means Q A's are B, where Q is a $\langle 1, 1\rangle$ generalized quantifier

where (a) is considered a good sentence, while (b) is less felicitous. The contrast between (a) and (b) suggest that while embedding an indefinite expression (*two*) in a superlative quantifier licenses a specific construal (*namely Jack and Jill*), the same is not licensed in the case of a comparative modifier.

Secondly, it has been demonstrated that superlative quantifiers are mastered later than the comparative ones during language development (Musolino, 2004), (Geurts et al., 2010). Furthermore, there is ample data concerning processing of those quantifiers. It has been for instance shown that verification of sentences with superlative quantifiers requires a longer time than verification of sentences with respective comparative quantifiers (Koster-Moeller et al, 2008),(Geurts et al., 2010). Moreover, the processing of quantifiers is influenced by their monotonicity. A quantifier $Q(A, B)$ is upward monotone in its first argument A if it licences inferences from subsets to supersets, that is if $Q(A, B)$ and $A \subseteq A'$, then $Q(A', B)$. A quantifier $Q(A, B)$ is downward monotone in its first argument A if it licences inferences from supersets to subsets, that is if $Q(A, B)$ and $A' \subseteq A$, then $Q(A', B)$. Understood as $\langle 1, 1\rangle$ generalized quantifiers, *at least n* and *more than n* are upward monotone in both their arguments, while *at most n* and *fewer than n* are downward monotone. It has been shown that although the downward monotone *at most n* and *fewer than n* take a longer time to be verified than the upward monotone *at least n* and *more than n*, they are actually falsified faster (Koster-Moeller et al, 2008).

Finally, important arguments against the semantical equivalence between the comparative and superlative quantifiers come from the analysis of people's acceptance of inferences with those quantifiers. Empirical data show that a majority of responders usually reject inferences from *at most n* to *at most n+*, (where $n+$ denotes any natural number greater than n), although they accept presumably equivalent inferences with comparative quantifiers (Geurts et al., 2010), (Cummins & Katsos, 2010). To illustrate it with an example: while people are unlikely to accept that *if at most 5 kids are playing in this room, then at most 6 kids are playing in this room* (2-14% in Cummins' and Geurts' experiments for this inference scheme) they are more likely to accept that *if fewer than 5 kids are playing in this room, then fewer than 6 kids are playing in this room* (between 60-70% in Cummins'and Geurts'). Such data seem to directly contradict the standard account.

2 Modal semantics or clausal implicature?

Several theories have been developed to explain why seemingly logically equivalent quantifiers show such big differences in the way people use them in the language. Geurts (2007), (2010) proposes modal semantics for superlative quantifiers and rejects the assumption that equivalences (1) and (2) hold in natural languages. According to this proposal (referred here as a "modal account"), while *more than n* and *less than n* have a conventional meaning defined in terms of inequality relation, *at least n* and *at most n* have a modal component, namely:

> *at least n A's are B* means that: a speaker is certain that there are *n* elements which are both A and B, and considers it possible that there are more than *n*.
>
> *at most n A's are B* means that: a speaker is certain that there is no more than *n* elements that are both A and B, and considers it possible that there are *n* elements.

According to this proposal, as semantically richer, superlative quantifiers are expected to be harder to process than the respective comparative quantifiers. Finally, defined as above, *at most n A are B* does not imply *at most n+ A are B*: The latter implies that it is possible that there are *(exactly) n+ A that are B*, which is contradicted by the semantics of *at most n* in the premise.

There are strong arguments against the modal account. For instance this account seems unsatisfactory with regard to superlative quantifiers embedded in conditional and various other contexts. Authors (Geurts & Nouwen, 2007),(Geurts et al., 2010) realize themselves this problem and illustrate it with the following example:

> *If Berta had at most three drinks, she is fit to drive. Berta had at most two drinks. Conclusion: Berta is fit to drive.*

Such inferences, which are indeed licensed by the inference from *at most n (2)* to *at most n+ (3)*, are commonly accepted by speakers (over 96% in Geurts's experiment).

Furthermore, while inferences from *at most n* to *at most n+(1)* are rejected by majority of people (ca 84% in Geurts' experiment) there are subjects who do accept them (14% in Geurts' and even more in our experiment — ca. 23%). If *at most n* logically implies *possible that n and not possible that more than n*, then the inference from *at most n* to *at most n+* should be inaccessible (except for cases of random mistakes) for any language users, due to the apparent contradiction between the premise and the conclusion. Last but not least, to say that *possible that n* is a part of the semantics of *at most n*, implies that *at most n* cannot be paraphrased by *not more than n*. However such a paraphrase seems totally eligible.

A slightly different account was proposed by Cummins & Katsos (2010), who observe that the considered linguistical phenomena could be better explained on pragmatical grounds. The authors show that people do not evaluate *at most n* and *exactly n-1* as equally semantically incoherent as cases of obvious logical incoherence, e.g. *at most n* and *exactly n+1* or *more than n* and *exactly n-1*. While sentence pairs, such as:

> *Jean has at most n houses. Specifically she has exactly n+1 houses.*

get average coherence judgments very low, i.e. -4, in the scale from -5 (incoherent) to $+5$ (coherent), sentence pairs:

> *Jean has at most n houses. Specifically she has n-1 houses*

get already +1.9. This result speaks against the "modal account", whose direct consequence is semantical incompatibility of *at most n* and *exactly n-*.

Consequently, Cummins et al. agree with Geurts that *at most n* and *at least n* both imply *possible that n*, but they claim that this is a pragmatical rather than a logical inference, namely a co-called *clausal implicature* (Levinson, 1983). Clausal implicature is a quantity implicature inferred due to use of epistemically weak statement. Since the expressed statement with a superlative quantifier e.g. *at most n A are B*, as equivalent to a disjunctive statement *there are exactly n or fewer than n elements that are both A and B*, does not imply the truth of its subordinate proposition $p =$ *there are exactly n elements that are A and B*, the possibility that p might or might not be true is inferred.

Although we agree with the intuitions concerning a modal component in the reasoning with superlative quantifiers, we reject the assumption by Geurts et al. that this component is a part of their meaning. Furthermore, although we agree with Cummins et al. that the mechanism that results with the observed inference patterns is more of a pragmatical nature, we are not satisfied with the "causal implicature" account. What we lack is a deeper insight into the source of this kind of a pragmatical inference and how it interacts with the logical meaning of those quantifiers in different reasoning contexts.

Our motivation to further experimentally investigate reasoning with superlative quantifiers is based inter alia on: (i) the lack of data concerning whether people accept inferences: *at least n* $\rightarrow$ *at least n -*, (ii) the lack of satisfactory data about how people accept inference with logically equivalent forms of superlative quantifiers: such as *not more than n/not fewer than n* or *n or fewer than n/n or more than n*, as well as how they accept mutual equivalences between these forms, (iii) finally, the lack of data concerning people's acceptance of logically incorrect inferences with the quantifiers considered here.

3 Two semantic conditions for *at most n*

Krifka (1999) points out that semantic interpretation of a sentence is usually a pair that specifies when the sentence is true and when it is false. Following Krifka, we propose to define meaning of a quantifier as a pair $\langle C_F, C_V \rangle$, where C_V is a verification condition (specifies how to verify sentences with this quantifier) and C_F is a falsification condition (specifies how to falsify sentences with this quantifier). Furthermore, we propose that the interpretation of a quantifier depends on a semantic context in which this quantifier is used, namely whether the context requires the use of the verification condition or the falsification condition. Verification and falsification conditions are to be understood algorithmically, with the "else" part of the conditional instruction being empty - thus, they verify (or falsify) the formulas only if their conditional test is satisfied. From a perspective of classical logic, these conditions should be dual, namely if C is a C_V condition for sentence ϕ, then C is a C_F condition for sentence $\neg\phi$, and vice versa. We further, however, observe that in the case of superlative quantifiers, there is a split between these two conditions. We suggest, that this split is a result of a pragmatic focus on the expressed borderline n.

One can think of the meaning of logical operators, thus also quantifiers, in terms of algorithms, that have to be performed in order to verify (or falsify) sentences with those operators. (See also (Szymanik, 2009), (Szymanik & Zajenowski, 2010), (Szymanik & Zajenowski, 2009)) Krifka (1999) observes, that a sentence *at most n x:* $\phi(x)$[2] says only that *more than n x:* $\phi(x)$ is false, and leaves a truth condition underspecified. In other words, the meaning of *at most n* provides an algorithm for falsifying sentences with this quantifier, but not (immediately) for verifying them. This corresponds with the experimental data showing that it is easier to falsify sentences with *at most* than to verify them (Koster-Moeller et al, 2008). Consequently, the primal semantical condition of *at most n x:* $\phi(x)$ could be understood as an algorithm: "falsify when the number of x that are ϕ exceeds n", and would constitute what we understand by the falsification condition.

Definition 1 *(falsification condition for at most)*

$$C_F(at\,most\,x : \phi(x)) := If\ \exists^{>n}x(\phi(x)),\ then\ falsify$$

But how can we know when a sentence with *at most n* is true? From the point of view of an algorithm it is a so-called "otherwise" condition that defines in this case the truth-condition. However a negation of a falsification condition is in sense *informationally empty*: it does not describe any concrete situation in which the given sentence can be verified. As a result, in those contexts that require to directly verify a sentence, we refer to a verification condition, which is specified independently. As expressing a positive condition, *at most n* may be understood as a disjunction *n or fewer than n* ("disjunctive *at most*").

Definition 2 *(verification condition for at most)*

$$C_V(at\,most\,n : x\phi(x)) := If\ (\exists^{=n}x\phi(x) \vee \exists^{<n}x\phi(x)),\ then\ verify$$

The disjunction in 2 could be further broken down to: $\bigvee_{i=1}^{n} \exists^{=i}x\phi(x) \vee \neg\exists x\phi(x)$, in short: $\bigvee_{i=0}^{n} \exists^{=i}x\phi(x)$, where the disjunct $\exists^{=o}x\phi(x)$ means that $\neg\exists x\phi(x)$.[3]

And $\exists^{=n}x\phi(x)$ means *precisely n x are* ϕ, that is:

$$\exists^{=n}x\phi(x) \iff \exists x_1 ... \exists x_n[\bigwedge_{i=1}^{n} \phi(x_i) \wedge \bigwedge_{1\leq i<j\leq n} (x_i \neq x_j) \wedge \forall y(\bigwedge_{i=1}^{n} y \neq x_i \to \neg\phi(y))] \quad (3)$$

[2] at most n x are ϕ

[3] Let us observe that $\exists^{<n}x\phi(x)$ is a short notation that can be misleading, since it is not an existential sentence. As existential, *fewer than n* would imply that there has to be at least one (though less than n) such x that is ϕ. However we would like a sentence *less than n x:* $\phi(x)$ to be also true if no x's are ϕ. Therefore, in fact, such a downward entailing sentence is a disguised universal sentence: $\forall x_1..x_n(\bigwedge_{i=1}^{n} \phi(x_i) \to \bigvee_{1\leq i<j\leq n}(x_i = x_j))$

3.1 Epistemic interpretation of disjunction

Following Zimmermann (2000) we adopt the view that disjunctive sentences in natural language are likely to get so-called epistemic reading, that is they are interpreted as *conjunctive lists of epistemic possibilities.* According to the proposed solution a disjunction P_1 *or...or* P_n is interpreted as an answer to a question: *Q: What might be the case?* and, thus, is paraphrased as a (closed) list L:

> L: P_1 (might be the case) [and]... P_n (might be the case) [and (*closure*) nothing else might be the case].

This results in the following reading of a disjunctive sentence:

Definition 3 *(Zimmermann, 2000)*

$$P_1 \vee ... \vee P_n \iff \diamond P_1 \wedge ... \wedge \diamond P_n$$

and (closure):

$$\forall P[\diamond P \rightarrow [P \cap P_1 = \emptyset \vee ... \vee P \cap P_n = \emptyset]]$$

The character of the closure requires a bit of our attention. Zimmermann (2000) observes, that disjunctive sentences in natural languages could be understood as *closed* (exhaustive lists of possibilities) or open (when other possibilities are not excluded), which in the spoken language is usually marked by intonation. Closure in Definition 3 indicates that the list is exhaustive. There are good reasons to treat NL disjunctions as generally closed, and it would make sense obviously also in the case of superlative quantifiers. In classical logic the truth of $\exists^{=n} x\phi(x) \vee \exists^{<n} x\phi(x)$ semantically excludes the option that $\exists^{>n} x\phi(x)$, as contradictory. In our analysis, however, we want to treat closure as a merely optional condition. This results from regarding the verification and falsification conditions as independent from each other.

3.2 Verification condition for *at most*

If we assume that disjunctions in natural language are likely to be interpreted as conjunctions of epistemic possibilities, then we get the following verification condition for *at most*:

Definition 4 *(epistemic interpretation of the verification condition for at most)*[4]

[4] A detailed description of the modal predicate logic needed for providing semantics of this kind of sentences is beyond the scope of this paper. For our present purposes it is just enough to assume that, for each possible world, we have a different domain of objects over which we quantify. We assume also standard semantics for modal operators, and we restrict to reflexive and transitive Kripke models.

$$C_W(at\,most\,n\,x : \phi(x)) := \; If\ (\diamond\exists^{=n}x\phi(x) \wedge \diamond\exists^{<n}x\phi(x)),\ then\ verify$$

and (closure)

$$If\ \diamond\exists^{>n}x\phi(x),\ then\ falsify$$

Where:

$$(\diamond\exists^{=n}x\phi(x) \wedge \diamond\exists^{<n}x\phi(x)) \iff \bigwedge_{i=o}^{n} \diamond\exists^{=i}x\phi(x) \tag{4}$$

Now we can see the asymmetry between the falsification and verification condition of *at most n*. While the falsification condition specifies precisely when the sentence has to be rejected as false, the verification condition (in the epistemic reading) provides a conjunctive list of epistemic possibilities that should all be the case in order to verify it.

The epistemic reading of the verification condition is also what differentiates superlative quantifiers from the comparative ones. We propose that the disjunctive form of the verification condition in the case of superlative quantifiers is a result of the focus on the borderline n. The n mentioned in the superlative quantifier constitutes the borderline of the truth-conditions, however the numeral n that occurs in a comparative quantifier is not a part of the truth-conditions. The borderline of truth-conditions (that is *n-1* for *fewer than n*, and *n+1* for *more than n*) remains silent. Consequently, the disjunctive form of comparative quantifiers (and hence also the epistemic reading), though logically possible, is not pragmatically justified. In principle, a comparative quantifier can be as well interpreted in the epistemic way (as a conjunction of epistemic possibilities). Such a reading is, however, not equally likely to occur as in the case of a superlative quantifier, since the borderline is not explicitly expressed.

Let us observe that closure of the verification condition is stronger than the falsification condition. Intuitively, if $\neg\diamond\psi$ (equivalently $\Box\neg\psi$), then as well $\neg\psi$ (here: $\psi = \exists^{>n}x\phi(x)$), i.e. $\Box\xi \Vdash \xi$, however the theorem holds only in reflexive Kripke frames. Furthermore, the optional character of the closure bases on our assumption that the falsification and verification conditions are in a sense independent and only as a pair constitute the full semantic interpretation. Since the falsification condition, as defined in 2, is sufficient to account for the right semantical criterion of when the sentence with *at most n* is false, the closure of the verification condition is redundant and might or might not be considered in the reasoning process. The optional character of closure turns out crucial in evaluating validity of inferences with *at most n*.

When *at most n* is interpreted as in Definition 4, then the inference from *at most n* to *at most n+1* is not valid. Namely, from $\diamond\exists^{=n}\phi(x) \wedge \diamond\exists^{<n}x\phi(x)$ one cannot infer $\diamond\exists^{=n+1}x\phi(x) \wedge \diamond\exists^{<n+1}x\phi(x)$. It is easy to observe that the conjunct $\diamond\exists^{=n+1}x\phi(x)$ cannot be proven based on the premise, though it can be excluded only if the closure of the premise is applied.

On the other hand, the inference: *at most n* $\rightarrow$ *at most n-1*, which is invalid in the standard account, in the epistemic interpretation is blocked only due to closure of the conclusion. That is the $\diamond\exists^{=n}x\phi(x)$ implied by the premise is contradicted by the closure of the conclusion, i.e. $\neg\diamond\exists^{>n-1}x\phi(x)$. It is important to notice that without the closure the implication holds (if the epistemic reading of the verification condition is applied).

In some aspects our proposal might seem similar to Geurts' "modal" approach, namely we define the verification condition of *at most n* and *at least n* (see below) in modal terms. The main difference is that, in our account, it is *only* the verification condition that is defined modally, while the falsification condition remains standard. This results in a specific split or ambiguity in the meaning of superlative quantifiers.

4 *At least* and bare numerals

The quantifier *at least n* might seem perhaps less interesting than *at most n*, as it seems, as is also evident from our results (see Section (5)), less problematic from the point of view of reasoning. As an upward monotone quantifier, *at least n* appears to provide a clear verification algorithm: "verify when *n x* (that are ϕ) are found". Such a semantical interpretation would not, however, account for the linguistical differences between *at least n* and *more than n-1*.

Let us start with defining a falsification condition for *at least n* as follows:

Definition 5 *(falsification condition for at least)*

$$C_F(at\ least\ n : x\phi(x)) := \ If\ \exists^{<n}x\phi(x),\ then\ falsify$$

What we expect from the verification condition is that it express the whole range of epistemic possibilities in which the sentence with *at least n* can be true. Understood as in formula (2), thus as equivalent to *more than n-1*, a sentence with *at least n* can be expressed as an existential sentence that merely says that *there are n* (x that are ϕ), but does not exclude the possibility that there are more:

$$\exists^{\geq n}x\phi(x) \iff \exists^{>n-1}x\phi(x) \iff \exists x_1...\exists x_n[\bigwedge_{i=1}^{n}\phi(x_i) \wedge \bigwedge_{1\leq i<j\leq n}(x_i \neq x_j)] \quad (5)$$

This formula however, which would be a right verification condition for the comparative *more than n-1*, does not outline the borderline n, which is emphasized in the superlative *at least n*. Therefore we further break down formula (5) into a disjunctive formula: *(exactly) n or more than n x are* ϕ, which constitutes our verification condition for *at least n*.

Definition 6 *(verification condition for at least n)*

$$C_V(at\ least\ n : x\phi(x)) := If\ (\exists^{=n}x\phi(x) \vee \exists^{>n}x\phi(x)),\ then\ verify$$

The latter can be handled as a conjunctive list of possibilities.

Definition 7 *(epistemic interpretation of the verification condition for at least n)*

$$C_F(at\ least\ n : x\phi(x)) := \ If\ (\diamond\exists^{=n}x\phi(x) \wedge \diamond\exists^{>n}x\phi(x)),\ then\ verify$$

and (closure)

$$If\ (\bigvee_{i=0}^{n-1} \diamond\exists^{=i}x\phi(x)),\ then\ falsify$$

Having introduced the semantical conditions for *at least n*, we further analyze how they affect inferences with this quantifier, in particular we mean here the inference (*at least n*→ *at least n-1*), as well as its presumably equivalent disjunctive form: (*n or more than n* → *n-1 or more than n-1*). We propose that the way people handle these inferences depends on how they interpret bare numerals, such as n.

From a logical perspective, a bare numeral n (e.g. "two") can be interpreted as denoting any set of *at least n* elements, or a set of *exactly n* elements.[5] Thus $\psi = nx : \phi(x)$ can simply mean that *there are n x that are* ϕ, without any further constraints on whether there are more. Then, ψ gets the reading as in formula (5). On the other hand, ψ could be understood with a kind of closure, that is that there are *exactly n x's are* ϕ, thus as in formula (3).

It has been a matter of a wide debate in formal semantics and pragmatics what is the right approach to interpreting bare numerals in natural languages. It has been proposed that: (i) the literal meaning of n is *at least n*, while the condition *exactly* comes as a scalar implicture (Horn, 1972), (ii) the basic meaning of n is *exactly n*, while both *at least* and *at most* readings would be context-based (Breheny, 2008) (iii) n is ambiguous between *at least n* and *exactly n* (Geurts, 2006), (iv) n is underspecified and can receive *at least*, *at most* or *exactly* readings depending on the context (Carston, 1998).

Let us now show how the interpretation of n interacts with the validity of inferences (*n or more than n*)→ (*n-1 or more than n-1*), given the epistemic interpretation of disjunction. Suppose now that n is interpreted with a closure: *exactly n*. It is easy to observe that, in such a case, *possible that n and possible that more than n* does not imply *possible that n-1 or possible that more than n-1*. The premise which is interpreted as in Definition 7 does not imply $\diamond\exists^{=n-1} \wedge \diamond\exists^{>n-1}$ (with closure $\bigwedge_{i=0}^{n-2} \neg\diamond\exists^{=i}x\phi(x)$) While $\diamond\exists^{>n-1}$ follows from both $\diamond\exists^{>n}$ and $\diamond\exists^{=n}$, the problematic element is $\diamond\exists^{=n-1}$, which is directly contradicted by the closure of the premise. But suppose that n does not get the "exact" reading, but it is interpreted barely as *there are n*, so as (5). Then from *possible that n* we can infer *possible that n-1*, since the latter does not exclude the possibility that there is a bigger set of elements.

[5] or a set of *at most n* elements, however this interpretation seems to be counterintuitive and allowed only in special contexts.

5 Experiment

We conducted a pilot reasoning experiment (in German) to check how people reason with superlative quantifiers: *at least n* (*mindestens n*)[6] and *at most n* (*höchstens n*), as well as with logically equivalent but linguistically different forms of those quantifiers: a comparative negative form and a disjunctive form. These were quantifiers: *not more than n* (*nicht mehr als n*) and *n or fewer than n* (*n oder weniger als n*) (as logically equivalent to *at most n*), and *not fewer than n* (*nicht weniger als n*) and *n or more than n* (*n oder mehr als n*) (as equivalent to *at least n*).

We were particularly interested in comparing subjects' acceptance of inferences from *at most n* to *at most n+1* (type B: see Table 1) with their acceptance of logically equivalent forms of this inference: type D and type F. We were also interested in people's willingness to infer *at most n* from *not more than n*, and vice versa (type G). Furthermore, we checked the respective inferences with *at least n*, that is: *at least n* → *at least n-1* (type A), and its equivalent forms: type C and type E. Finally, we checked the inferences between *not fewer than n* and *at least n* (type H). Last but not least we tested the respective *incorrect* inferences with the considered quantifiers, in all forms (see Table 2).

In the premises of our inferences we used four different quantifiers: *at least three*, *at most three*, *at least four*, *at most four*, and their equivalent forms. All the numbers were throughout spelled out in words according to the requirements of German grammar. There was only one example for each distinct inference relation (i.e. two per inference type). Every sentence content was different. Additionally, to introduce more variation, sentences which had *at most 4/at least 4* (or their equivalent forms) in the premise had a quantifier in the subject position (e.g. *At least four computers are broken in the lab*), sentences which had *at least 3 /at most 3* in the premise had a quantifier in the object position (e.g. *Arthur has at least three cars.*)

As fillers we used inferences with so-called bare numerals (e.g. *four*): those whose correctness depends on the "at least" reading of bare numerals, i.e. $n \rightarrow n-$ (e.g. $4 \rightarrow 3$); and those that are logically incorrect independently of the presumed reading, such as $n \rightarrow n+$ (e.g. $4 \rightarrow 5$).

5.1 Procedure

The experiment was conducted on German native speakers, mainly students of philosophy, psychology, neuroscience and computer sciences. There were 17 subjects (7 male). Subjects were asked to respond "yes" or "no" to the question whether the second sentence (below the line) has to be true, assumed that they know that the first sentence (above the line) is true. For a better understanding of the task two examples were given: one of a valid inference, that should be given a "yes" response:

[6] We give in brackets the German translation used in the experiment

$$\frac{\text{Inga has done three exercises.}}{\text{Inga has done more than two exercises.}}$$

and one of an invalid inference, that requires a “no” response:

$$\frac{\text{Eva has done three exercises.}}{\text{Eva has done fewer than two exercises.}}$$

Note that the examples were selected in such a way that their validity did not depend on the understanding of any of the tested inference relations, that is the examples served as an instruction of what is an inference in general, but not how to evaluate the inferences, that were tested in the experiment.

After reading the instruction subjects saw 40 randomly ordered reasoning tasks: one task at a time, displayed on a computer screen. There was no time limit in the test.

At the end of the experiment two additional control questions were asked, in which the inference from *at most n* to *at most n+1* was embedded in a deontic context. Note that the logically correct response to first question is “yes”, while to the second is “no”.

(1) Erika promised to drink at most six beers. She drank at most four. Did she keep her promise?

(2) Thomas is allowed to eat at most three cookies. He ate at most two. Did he break the rule?

Tables: 1 and 2 summarize the inferences as well as the results.

5.2 Results

Our first observation is that people accepted the logically correct inferences much more frequently than the logically incorrect ones. The incorrect inferences (apart from disjunctive inferences: *E'* and *F'* which turned out specially problematic) were mostly rejected and their acceptance rate was low enough ($1-9\%$) to be considered as a result of random mistakes (Table 2). On the other hand all correct inferences were accepted on the level of at least 20%[7], with high variance depending on the form of an inference, in this: inferences of type B and F seemed the most problematic.

The important result is that nearly 100% of responders did accept inference of type G and H, that is (*at most n* → *not more than n*) as well as (*not more than n* → *at most n*), and respective inferences between *at least n* and *not fewer than n*, which suggests that they do see those expressions as equivalent. Furthermore, while inferences from type B, namely the problematic (*at most n* → *at most n+1*) were accepted only by 23% of responders, the inferences of type D (*not more than n* → *not more than n+1*), were already accepted by 44%. The difference was statistically significant: $z = -2,333, p = .02, r = -.4$ [8] Thus, it seems that paraphrasing *at most n* to the negative form *not more than n* facilitates the inference.

Table 1. Logically correct inferences

	Premise	*Conclusion*	*Correct Responses*	*Percentage*
A *At least*	*at least 4*	*at least 3*	13	79%
	at least 3	*at least 2*	14	
B *At most*	*at most 4*	*at most 5*	3	23%
	at most 3	*at most 4*	5	
C *Not fewer than*	*not fewer than 4*	*not fewer than 3*	8	59%
	not fewer than 3	*not fewer than 2*	12	
D *Not more than*	*not more than 4*	*not more than 5*	7	44%
	not more than 3	*not more than 4*	8	
E *N or more than n*	*4 or more than 4*	*3 or more than 3*	10	67%
	3 or more than 3	*2 or more than 2*	13	
F *N or fewer than n*	*4 or fewer than 4*	*5 or fewer than 5*	6	23%
	3 or fewer than 3	*4 or fewer than 4*	2	
G *"At most" Equivalence*	*not more than 3*	*at most 3*	16	98%
	at most 3	*not more than 3*	17	
	not more than 4	*at most 4*	17	
	at most 4	*not more than 4*	17	
H *"At least" Equivalence*	*not fewer than 3*	*at least 3*	17	93%
	at least 3	*not fewer than 3*	17	
	not fewer than 4	*at least 4*	16	
	at least 4	*not fewer than 4*	13	
K *Numerical*	*5*	*4*	9	65%
	6	*2*	11	
	7	*6*	13	
	8	*5*	11	
embedded	*at most 4*	*at most 6*	17	100%
L	*at most 2*	*at most 3*	16	94%

Table 2. Logically incorrect inferences

	Premise	*Conclusion*	*Correct Responses*	*Percentage*
A' *At least*	*at least 4*	*at least 5*	1	6%
	at least 3	*at least 4*	1	
B' *At most*	*at most 4*	*at most 3*	0	6%
	at most 3	*at most 2*	2	
C' *Not fewer than*	*not fewer than 4*	*not fewer than 5*	1	9%
	not fewer than 3	*not fewer than 4*	2	
D' *Not more than*	*not more than 4*	*not more than 3*	1	9%
	not more than 3	*not more than 2*	2	
E' *N or more than n*	*4 or more than 4*	*5 or more than 5*	3	20%
	3 or more than 3	*4 or more than 4*	4	
F' *N or fewer than n*	*4 or fewer than 4*	*3 or fewer than 3*	8	50%
	3 or fewer than 3	*2 or fewer than 2*	9	
K' *Numerical*	*4*	*5*	0	1,5%
	2	*6*	1	
	6	*7*	0	
	3	*8*	0	

The inferences of type A, that is *at least n*→ *at least n-1*, turned out relatively unproblematic for subjects, who accepted them in ca. 80% of cases. Interestingly a paraphrase to the negative comparative form *not fewer than n*, made the task more difficult (59% accepted; means comparison: $z = -2.07, p = .038, r = -.355$). However, it is worth to note that inferences of type A were still rejected

[7] We give an overall result for a given type of an inference

[8] In all the cases we used Wilcoxon Signed Ranks test to compare means.

by ca. 20%, which suggests that there is some, at least pragmatic, mechanism suppressing this inference.

The results for the disjunctive inferences (E and F) are especially interesting. First of all the response pattern for disjunctive counterpart of *at least* corresponds with the predictions of classical logic: While logically valid inferences (E) were accepted on a relatively high level of 67% (which is lower, though not significantly lower, compared to the acceptance of the basic form (type A)), the invalid inferences (E') were mostly rejected (only 20% accept). The opposite effect, however, we got for the disjunctive form of *at most*. The logically valid inferences (F) were mostly rejected (only 23% accept), whilst invalid inferences (F') were accepted in exactly 50% of cases. Interestingly the acceptance rate of (F) inferences was similar to the acceptance rate of the basic form of inferences with *at most* (B). In both cases the differences between acceptance rate of correct and incorrect forms were statistically significant, and were as follows: The differences between correct and incorrect inferences with "disjunctive *at most*" (F and F'): $z = -2.491, p = .013, r = -.43$ and correct and incorrect "disjunctive *at least*" (E and E'): $z = -2.165, p = .030, r = -.37$. The differences between disjunctive *at most* and *at least*: incorrect (E' and F') $z = -2.057, p = .040, r = -.36$: and correct (E and F) :$z = -2.697, p = .007, r = -.46$.

The "correct" inferences with bare numerals were accepted in ca. 65% of cases. There was only one mistake in the incorrect inferences with bare numerals. Finally, both embedded *at most* inferences got nearly 100% correctness rate (one mistake only for question 2).

6 Discussion

Although our results cannot be treated as a final evidence of our theory, our experiment certainly provides various important observations that support the plausibility of our proposal.

First of all, all the implications between the negative comparative and superlative forms of considered quantifiers were almost without exceptions accepted by our subjects. This result supports the assumption that those are semantically equivalent forms in natural language.

Secondly, the inferences *(at least n → at least n-)*, although accepted by a majority of responders, were not as obvious as the standard theory would predict, and 20% of subjects rejected them (A, Table 1). This suggests existence of some, at least pragmatic, mechanism interfering in subject's reasoning with *at least*. What is also worth reminding, valid inferences with "disjunctive *at least*" were rejected even more often (F, Table 1). We consider that this effect can be explained in terms of the epistemic interpretation of *at least n* and its interaction with the reading of bare numerals. Let us notice that our results provide a weak evidence for the interplay between the reading of bare numerals and the treatment of "disjunctive *at least*" inferences. In our experiment inferences of type K: $n \rightarrow n-$, which base on the "at least" reading of numerals were accepted in ca. 65% of cases, thus our responders in 35% cases integrated the "exact" reading of bare numerals, which resulted in their rejection of considered inferences. However, "disjunctive *at least*" inferences (type E) ware rejected also in ca. 33%. We suggest that rejection of type E inferences was as well a result of an exact reading of a bare numeral n, as we have explained above. Although a correlation between subject's acceptance of type E and type K inferences failed to reach significance, it was close to significant (Spearman's rho= .426, $p = .08$) and we expect that with a bigger sample it could reach the significance level.

A similar effect is presumably the reason why 20% of subjects rejected type A inferences: (*at least n*$\rightarrow$ *at least n-1*). Namely, the application of the epistemic verification condition of *at least n* together with an exact reading of bare numerals results in rejection of such inferences. This effect might be however weaker and less likely to occur than in the case when the disjunctive form is given explicitly.

Thirdly, the surprising result that subjects accepted the invalid inferences with "disjunctive *at most*" more frequently than the valid ones can be explained by our proposal. As we have proposed above, closure in the verification condition is optional, since the falsification condition is sufficient to account for the right semantics. However, if context enforces applying one of the semantical conditions (verification or falsification), then the other one tends to be ignored. While, from the perspective of classical logic it should be enough to use only one of the conditions (since the other can be defined via the first one), in the case of superlative quantifiers the epistemic reading of the verification condition creates the bifurcation in the meaning. This results in different inferential patterns in which those quantifiers occur, depending on what the context primarily enforced: the verification or falsification condition.

In what follows, and as we have explained above, when the verification condition is used, then *n or fewer then n* does not imply *n+ or fewer then n+* due to the epistemic interpretation. Though, it also does not exclude it if no closure is applied. However, *n or fewer then n* does imply *n- or fewer then n-* if the verification condition is used but no closure is applied. Now we can explain why inferences (F', Table 2): (*n or fewer than n*) $\rightarrow$ (*n-1 to fewer than n-1*) got a 50% rate of acceptance, although they are invalid both in the standard account and in the epistemic account. Based on Definition 4, $\exists^{=n}x\phi(x) \vee \exists^{<n}x\phi(x)$ is interpreted as $\diamond\exists^{=n}x\phi(x) \wedge \diamond\exists^{<n}x\phi(x)$ (with closure: $\neg\diamond\exists^{>n}x\phi(x)$). But $\diamond\exists^{<n}x\phi(x)$ can be broken down to: $\gamma = \diamond(\diamond\exists^{n-1}x\phi(x) \wedge \diamond\exists^{<n-1}x\phi(x))$ Now γ implies $\diamond\exists^{=n-1}x\phi(x) \wedge \diamond\exists^{<n-1}x\phi(x)$ (here we use the assumption that the world accessibility relation is transitive). In such a case it might happen that the closure of the conclusion, that is: $\neg\diamond\exists^{>n-1}x\phi(x)$ which contradicts the assumption that $\diamond\exists^{=n}x\phi(x)$ is ignored by subjects, which results in the high logical mistake ratio.

7 Conclusions

We have argued that the meaning of a quantifier can be defined as a pair $\langle C_F, C_V\rangle$, in which the verification (C_V) and the falsification (C_F) condition for sentences with this quantifier are specified separately. Though from the logical point of view those conditions should be dual, in the case of superlative quantifiers they are not. Namely, pragmatic focus on the borderline n in both *at least* n and *at most* n enforces a disjunctive verification condition, which is further interpreted as a conjunctive list of epistemic possibilities.

Finally, we would like to say few words about why we want to understand the verification and falsification conditions in terms of algorithms. Let us make an observation that semantical equivalence and procedural identity of algorithms are different things. Let us consider algorithms A_1 and A_2:

A_1 : Count all x that are ϕ. If the number m of x that are ϕ is smaller than $n-1$, then $verify$.

A_2 : Count all x that are ϕ. If the number m of x that are ϕ equals n or is smaller than n, then $verify$.

A_2 and A_3 are semantically equivalent, namely they verify logically equivalent formulas, e.g. ψ_1, ψ_2 and ψ_3, however in a sense of procedures that are executed they are not identical.

$$\begin{aligned}
&\psi_1 \iff \neg\exists^{>n} x\phi(x)\\
&\psi_2 \iff \exists^{<n+1}\phi(x)\\
&\psi_3 \iff \exists^{=n} x\phi(x) \vee \exists^{<n} x\phi(x)\\
&\psi_1 \iff \psi_2 \iff \psi_3
\end{aligned}$$

Furthermore, A_3

A_3 : When the number m of x that are ϕ is bigger then n, then $falsify$.

is dual to both A_1 and A_2, namely adding an *otherwise verify* condition to A_3 and *otherwise falsify* condition to A_1 and A_2 would make A_3 semantically equivalent to both A_1 and A_2. However, without the "otherwise" condition, A_3 does not allow to verify any of the given sentences, while A_2 or A_1 do not allow to falsify them. Then, $\langle A_1, A_3\rangle$, or $\langle A_2, A_3\rangle$ could be considered pairs of partial algorithms. Each pair could constitute a full semantical interpretation of each of the sentences ψ_1, ψ_2, ψ_3.

We consider that logically equivalent, but linguistically different, natural language sentences may trigger different kinds of such partial algorithms, or pairs of partial algorithms. First of all two equivalent sentences that differ in the linguistical form can trigger as primary only one of the algorithmic conditions: verification or falsification, while the complement condition is ignored. Secondly, they might trigger non-identical verification/falsification procedures. Consequently, it might happen that the executed procedures differ in complexity. Additionally, if we take into account that some extra mechanisms, e.g. the above-discussed epistemic interpretation of disjunctive conditions, might play a role, we not only obtain different algorithms in the procedural sense but also *semantically non-equivalent* algorithms for logically equivalent or even same sentences. For instance, A_2 would be replaced by:

A'_2: if [it is possible that there are exactly n x that are ϕ AND it is possible that there are fewer than n x that are ϕ], then $verify$.

While A_2 is dual to A_3, A'_2 is not anymore. However a pair $\langle A_3, A'_2 \rangle$ could constitute a semantical interpretation of a natural language sentence *at most* $n : x\phi(x)$, which would explain the non-semantically coherent (from the point of view of classical semantics) inference patterns in which this quantifier occurs.

Appendix: The complete list of pairs of sentences used in the experiment (premise, conclusion), translated from German.

At least four Anna's dress are red. At least three Anna's dresses are red.
Arthur has at least three cars. Arthur has at least two cars.
At most four books were stolen from the library. At most five books were stolen from the library.
Markus ate at most three pieces of cake. Markus ate at most four pieces of cake.
Not fewer than four cards are missing in the deck. Not fewer than three cards are missing in the deck.
A child has painted not fewer than three pictures. A chid has painted not fewer than two pictures.
Not more than four students came today to the philosophy seminar. Not more than five students came today to the philosophy seminar.
Sabine got not more than three presents. Sabine got not more than four presents.
Four or more than four students were sick this week. Three or more than three students were sick this week.
Christopher speaks three or more than three languages. Christopher speaks two or more than two languages.
Four or fewer than four students have passed the course. Five or fewer than five students have passed the course.
Beate has three or fewer than three children. Beate has four or fewer than four children.
Five people came to the party. Four person came to the party.
Monika invited six guests to her birthday party. Monika invited two guests to her birthday party.
Seven fruits in the basket have spoilt. Six fruits in the basket have spoilt.
Alicia bought eight bottles of beer. Alicia bought five bottles of beer.
At least four of Carolina's scarfs are blue. At least five of Carolina's scarfs are blue.
Thomas has read at least three books. Thomas has read at least four books.
At most four computers in the lab are broken. At most three computers in the lab are broken.
Andrea baked at most three pizzas. Andrea baked at most two pizzas.
Not fewer than four professors attended the meeting. Not fewer than five professors attended the meeting.
Hans had not fewer than three glasses of wine. Hans had not fewer than four glasses of wine.
Not more than four people have applied for this job. Not more than three people have applied for this job.
Natalie wrote not more than three exercises. Natalie wrote not more than two exercises.
Four or more than four girls in the class are good in arts. Five or more than five girls in the class are good in arts.
Christina's cat gave birth to three or more than three kittens. Christina's cat gave birth to four or more than four kittens.
Four or fewer than four students failed in the exam. Three or more than three students failed in the exam.
Tanja trains three or fewer than three times a week. Tanja trains two or fewer than two times a week.
Four new students joined the chess club this week. Five new students joined the chess club this week.
Stephanie baked two cakes for her birthday. Stephanie baked six cakes for her birthday.
Six members of the library club came to the meeting. Seven member of the library club came to the meeting.
Frank gave his mother three roses. Frank gave his mother eight roses.
Not more than three children have done their homework for today. At most three children have done their homework for today.
At most three girls took part in the maths competition. Not more than three girls took part in the maths competition.
Erika has not more than four necklaces. Erika has at most four necklaces.
Lena takes at most four courses at the university. Lena takes not more than four courses at the university.
Not fewer than three new animals were born in the city zoo. At least three new animals were born in the city zoo.
At least three exotic three have died in our botanic garden. Not fewer than three exotic threes have died in our botanic garden.
Daniel plays not fewer than four times a week football. Daniel plays at least four times a week football.
Albert has at least four exams this semester. Albert has not fewer than four exams this semester.

Bibliography

Breheny, Richard. 2008. A new look at the semantics and pragmatics of numerically quantified noun phrases. *Journal of Semantics* 25.93.

Carston, Robin. 1998. Informativeness, relevance and scalar implicature. In *Relevance theory: Applications and implications*, ed. by R. Carston & S. Uchida, 179-236. Amsterdam: John Benjamins.

Cohen, Ariel & Krifka, Manfred. 2011. Superlative quantifiers as modifiers of meta-speech acts. In. Partee, B.H., Glanzberg, M., & Skilters, J. (Eds.) (2011). *Formal semantics and pragmatics. Discourse, context and models. The Baltic International Yearbook of Cognition, Logic and Communication,* Vol. 6 (2010). Manhattan, KS: New Prairie Press.

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

Geurts, Bart. 2006. Take 'five': The meaning and use of a number word. In *Non-definiteness and Plurality*, ed. by S. Vogeleer & L. Tasmowski, 311-329. Amsterdam/Philadelphia: Benjamins.

Geurts, Bart and Nouwen, Rick. 2007. "At least" et al.: The semantics of scalar modifiers. *Language*, 83: 533-559.

Geurts, Bart, Katsos, Napoleon, Cummins, Chris, Moons, Jonas, and Noordman, Leo. 2010. Scalar quantifiers: Logic, acquisition, and processing. *Language and Cognitive Processes*, 25: 130-148.

Horn, Laurence R. 1972. *On the Semantics of Logical Operators in English.* Ph. D. thesis, Yale University dissertation.

Koster-Moeller, Jorie, Varvoutis, Jason & Hackl, Martin. 2008. Verification procedures for modified numeral quantifiers, *Proceedings of the West Coast Conference on Formal Linguistics*, 27.

Krifka, Manfred. 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 Science B.V., 257-291.

Levinson, Stephen C. 1983. *Pragmatics.* Cambridge, England: Cambridge University.

Musolino, Julien. 2004. The semantics and acquisition of number words: Integrating linguistic and developmental perspectives, *Cognition*, 93-1, 1-41.

Nouwen, Rick 2010. Two kinds of modified numerals. *Semantics and Pragmatics* 3 (3).

Szymanik, Jakub. 2009.*Quantifiers in TIME and SPACE. Computational Complexity of Generalized Quantifiers in Natural Language*, DS-2009-01, Amsterdam.

Szymanik, Jakub & Zajenkowski, Marcin. 2010. Comprehension of simple quantifiers. Empirical evaluation of a computational model. *Cognitive Science*, 34(3), pp. 521-532.

Szymanik, Jakub & Zajenkowski, Marcin. Improving methodology of quantifier comprehension experiments. 2009. *Neuropsychologia*, Vol. 47, No. 12, pp. 2682-2683.

Zimmermann, Thomas Ede. 2000. Free choice disjunction and epistemic possibility, *Natural Language Semantics*, 8 (2000), 255 - 290.