

Verification Procedures for Modified Numeral Quantifiers

Jorie Koster-Moeller¹, Jason Varvoutis², and Martin Hackl¹
¹Pomona College and ²MIT

1. Overview

This paper presents two experimental studies of verification procedures for the modified numeral quantifiers *more than k*, *at least k*, *at most k*, and *fewer than k* in order to investigate to what extent the form of a particular quantifier determines its associated verification strategies. We show that the numeral *n* affects the counting component of the verification process, that the modifier (*more than*, *at least*, ...) affects the decision stage, and that these two factors don't interact. These findings are difficult to reconcile within Generalized Quantifier Theory (GQT), which is too coarse to exploit morpho-syntactic differences between denotationally equivalent quantifiers.

2. Coarseness of Generalized Quantifier Theory

Quantificational structures are standardly analyzed within the framework of Generalized Quantifier Theory (GQT) (Mostowsky 1957, Barwise and Cooper 1981, Keenan and Stavi 1986, etc.). According to GQT, quantification in natural language is a form of second order predication. The centerpiece of a quantificational structure is an expression that denotes a relation between sets of individuals. Thus, a quantifier like *every* is said to denote the subset relation, (1)a, rather than a variable binding first order operators, (1)b.

- (1) a. $\llbracket \text{every} \rrbracket (A)(B) = 1$ iff $A \subseteq B$
b. $\llbracket \text{every} \rrbracket (A)(B) = 1$ iff $\forall x [A(x) \rightarrow B(x)]$

Importantly, relations between sets are assumed to be semantic primitives in GQT, i.e. semantic objects that can be denoted without the help of compositional processes. This implies that it is not necessary to provide a compositional analysis for quantifiers even when they are morpho-syntactically complex (such as the modified numeral quantifiers listed above) to give a full characterization of their semantics. They are simply treated as quasi-idiomatic expressions whose particular morpho-syntax is assumed to have no effect on their semantic properties.¹

A consequence of this view is that GQT is too coarse to distinguish between denotationally equivalent determiners even when they are morpho-syntactically quite different. Thus the modified numeral quantifiers in (2) and (3) are claimed to be respectively interdefinable and the choice of how their truth-conditional import is given is seen as entirely arbitrary.

- (2) a. $\llbracket \text{more than } k \rrbracket (A)(B) = \llbracket \text{at least } k+1 \rrbracket (A)(B) = 1$ iff $|A \cap B| > k$
b. $\llbracket \text{at least } k+1 \rrbracket (A)(B) = \llbracket \text{more than } k \rrbracket (A)(B) = 1$ iff $|A \cap B| \geq k+1$
- (3) a. $\llbracket \text{fewer than } k+1 \rrbracket (A)(B) = \llbracket \text{at most } k \rrbracket (A)(B) = 1$ iff $|A \cap B| < k+1$
b. $\llbracket \text{at most } k \rrbracket (A)(B) = \llbracket \text{fewer than } k+1 \rrbracket (A)(B) = 1$ iff $|A \cap B| \leq k$

From a morpho-syntactic point of view, however, this insensitivity to form seems counterintuitive, since the format in which the truth-conditions are given in (2)a seems to yield a better correspondence

¹ Keenan and Westerstahl (1997:17) use the terms "lexical and near lexical determiners" for morpho-syntactically complex quantifiers like *more than half* while van Benthem (1986) calls them "tightly knit compounds" to indicate that their semantics is not a function of their parts.

for *more than k* than for *at least k+1* while the opposite is true for (2)b. Similarly, the format in which the truth-conditions are stated in (3)a seems to be a better match for *fewer than k+1* than for *at most k* as it employs operators that have direct morpho-syntactic counterparts in *fewer than k+1* but not in *at most k*.

These considerations amount to a conceptual argument against the coarseness inherent to GQT. Language internal empirical challenges against GQT's treatment of modified numeral quantifiers have been presented in Krifka (1998), Hackl (2000), Geurts and Nouwen (2007), and Buring (2008) among others. Common to all of these is the claim that *more than k* and *at least k+1* and *fewer than k+1* and *at most k* are not actually interdefinable and that one needs to take the particular form of the modified quantifier into account to fully understand their semantic properties. In particular, all of these authors agree that one needs to recognize two components: the modifier, which is either comparative or superlative in nature, and the numeral.

This paper presents language external evidence against GQT's treatment of modified numeral quantifiers by showing that the particular form in which truth-conditions are stated affects language external cognitive systems. Specifically, we show that verification procedures triggered by denotationally equivalent quantifiers differ systematically in ways that parallel the difference in the internal, morpho-syntactic makeup of the quantifier.

3. Verifying Quantified Statements

It has been acknowledged already in Barwise and Cooper (1981) that the denotations GQT provides for quantifiers are unrealistic for the purpose of verification. Consider the semantic objects that a comprehender would need to have in mind to verify statements like those in (4).

- (4) a. Some students are happy.
b. No students are happy.

Taking GQT's analysis literally, verifying (4)a requires (among other things) determining the set of students, the set of happy entities, the set of sets of entities that do (or do not in the case of (4)b) intersect with the set of students (the generalized quantifier) and finally checking whether the set of happy entities is in that set.

To address this concern, Barwise and Cooper (1981) provide a processing amendment based on the notion of a witness set as defined in (5). Given the notion of a witness set, truth-conditions for monotone increasing and monotone decreasing generalized quantifiers can be restated as in (6).²

- (5) A witness set for a generalized quantifier $D(A)$ living on A ³ is any subset W of A st.
 $D(A)(W) = 1$.
- (6) a. If $D(A)$ is monotone increasing then for any X , $D(A)(X) = 1$ iff $\exists W [W \subseteq X]$.
b. If $D(A)$ is monotone decreasing then for any X , $D(A)(X) = 1$ iff $\exists W [X \cap A \subseteq W]$.

If the statements in (6) are combined with a natural economy principle governing information gathering for the purpose of verification – Look for the smallest witness set! – we can derive the following general characterizations for verifying modified numeral quantifiers, where W_n stands for a witness set with cardinality n .

- (7) a. $\llbracket \text{more than } k \rrbracket (A)(B) = 1$ iff $\exists W_{k+1} [W_{k+1} \subseteq B]$
b. $\llbracket \text{at least } k+1 \rrbracket (A)(B) = 1$ iff $\exists W_{k+1} [W_{k+1} \subseteq B]$
- (8) a. $\llbracket \text{fewer than } k+1 \rrbracket (A)(B) = 1$ iff $\exists W_k [A \cap B \subseteq W_k]$
b. $\llbracket \text{at most } k \rrbracket (A)(B) = 1$ iff $\exists W_k [A \cap B \subseteq W_k]$

² It is transparent that the notion of a witness set is an attempt to recover aspects of first order analyses of quantifiers.

³ A generalized quantifier $D(A)$ "lives on" A iff D is conservative.wrt. to A .

Applied to a concrete example, in order to verify a statement like *More than 6 of the dots are blue* comprehenders are expected to look for a 7-membered set of dots such that all of them are blue, while verifying *At most 6 of the dots are blue* would mean to look for a 6-membered set of dots that contains all the blue dots.

Since the strategy for monotone decreasing quantifiers is more complex than the one for monotone increasing quantifiers, this theory expects verifying statements with decreasing quantifiers to be more difficult than verifying statements with increasing quantifiers. Interestingly, this approach also expects that falsifying decreasing quantifiers should be easier than falsifying increasing quantifiers. For example, as soon as it is determined that the number of blue dots exceeds 6, *At most 6 of the dots are blue* is false and the search can stop, while in order to answer “false” for *More than 6 of the dots are blue* one will have to consider the entire set of dots.

Importantly, GQT-based verification does not expect an effect of the number mentioned in the quantifier (n). Rather, because it is based entirely on denotation (ignoring form) it predicts an effect of the critical number (N), the number of dots that allows a comprehender to determine the truth-value. For statements with increasing quantifiers *More than 6/At least 7 of the dots are blue*, $N = 7$. As long as there are less than 7 blue dots the statement is false and as soon as the 7th blue dot is encountered the statement is true. For statements with decreasing quantifiers like *At most 6/Fewer than 7 of the dots are blue*, $N = 7$ as well, although in these cases the switch in truth-value is from true to false as the 7th blue dot is encountered. Notice, however, the number mentioned, n , varies such that for *More than 6/At most 6*, $n = 6$ while for *At least 7/Fewer than 7*, $n = 7$. Thus, unlike GQT, a decompositional approach that recognizes n as an independent semantic unit inside the quantifier might expect n to have an effect on verification.

We can summarize these considerations as follows: a GQT model of verification for modified numeral quantifiers that is based on witness sets predicts an effect of monotonicity and an effect of N but crucially, not an have an effect of n . To see whether these predictions are indeed borne out we conducted two experiments discussed in the next section.

4. Investigating Verification Procedures using Self-Paced Counting

4.1. Self-Paced Counting (SPC)

Solving a verification problem that involves counting the number of objects in a scene is a complex task with many degrees of freedom, especially when the scene is displayed all at once. This makes it rather difficult to relate an observed difference in verification to a difference in linguistic form. To sidestep this difficulty, Self-Paced Counting reveals a given scene in a step-by-step and self-paced fashion quite similar to the widely used Self-Paced Reading methodology.⁴ In a typical trial, subjects hear a sentence such as *More than six of the dots are blue* or *At least seven of the dots are blue* played over speakers attached to a computer. They will then see an array of covered hexagonal plates displayed on a computer screen (Frame1, Figure 1). As they press the space bar, the plates are opened in increments of 2 or 3, revealing colored dots, while previously seen dots are covered again with masked plates, allowing participants to see only a small portion of the array at a time. To vary the total number of dots across arrays, some plates may not cover a dot, revealing only empty space. Subjects answer true or false by pressing the appropriate response key on a keyboard at any time during the trial and are encouraged to answer as fast and as reliably as possible.

⁴ See Hackl (forthcoming) for a more detailed description of SPC and supporting evidence documenting the soundness of the paradigm.

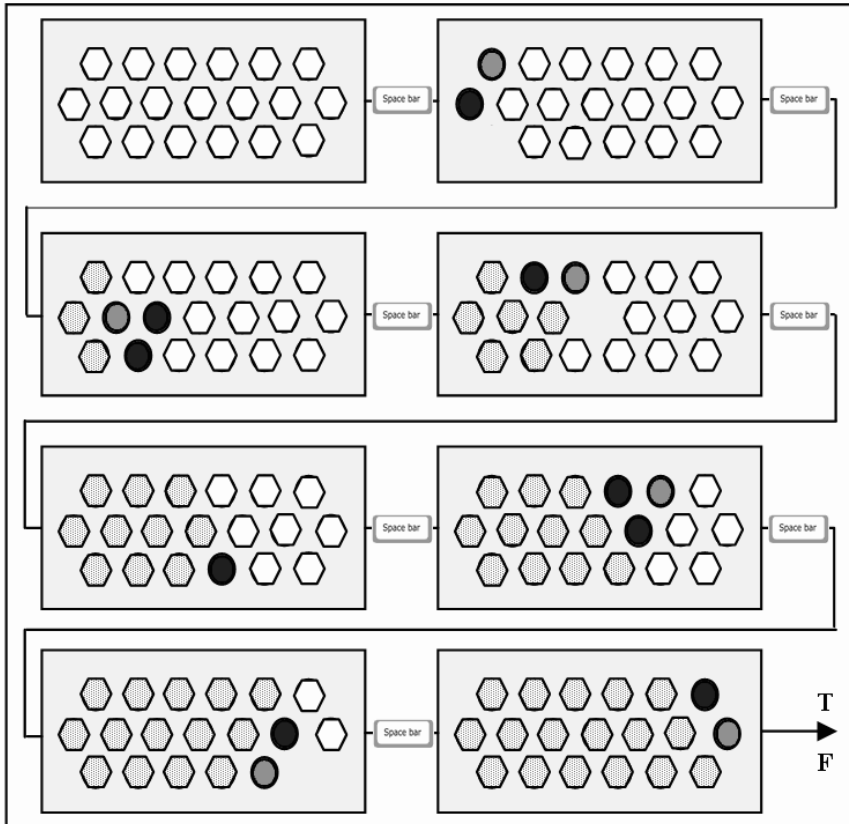


Figure 1: Sequence of Events in Self-Paced Counting Trials

4.2. Experiment 1

Methods and Materials: The first study reported here used 120 target sentences, 30 *more than k*, 30 *at least k+1*, 30 *fewer than k+1*, and 30 *at most k*. These sentences were combined with arrays so that half of them were true and half of them false. The arrays varied evenly by the number of target dots such that 40 arrays had 6 dots in the target color, 40 had 7, and another 40 had 8. The total number of dots ranged from 16 to 19. No target items could be answered before the final (seventh) frame was revealed. All of the target items were designed such that one less than the critical number of dots needed to verify was seen on Frame 4. For *more than k* and *at most k* this was also n , the number mentioned in the quantifier itself, while for *at least k+1* and *fewer than k+1* it was not. Furthermore, on frames 5 and 6 no target dots were presented. (This can be clearly seen in the drop of in RTs on these frames in Figure 2 below.) In Frame 7, the truth value of a given item was determined by the presence or absence of a final dot, which determined whether N , the critical number was reached. For monotone increasing quantifiers (*more than k* and *at least k+1*) the presence of an additional dot in Frame 7 supported an answer of “True” while for decreasing quantifiers the correct answer was “False.”

To camouflage our target items, we used 288 filler items which employed quantifiers like *only k*, *exactly k*, *k*, *many*, *few*, *some*, *all but k*, etc. Half of the filler items were true and half false and most could be answered before the final frame was revealed. In addition, subject got 10 practice items to familiarize themselves with the task, with additional practice items after every break.

Stimuli were presented using the Presentation software package from Neurobehavioral Systems on a PC, which recorded the time between successive space bar presses, the time it took to press an answer key, the screen at which the answer key was pressed, and whether the given answer was correct.

25 undergraduate students at the Claremont Colleges, all native speakers of English, participated in this experiment. Subjects received either course credit or 10 dollars in cash as compensation. Only subjects whose overall success rate across all items of the experiment was at least 75% were considered. No subject had to be excluded by this criterion and only reaction times (RTs) from correctly answered items were included in the analysis.

Results: Figure 2 provides a frame by frame representation of the average RTs for our 4 target categories, averaging across true and false items. Two areas (shaded in Figure 2) are of particular interest. In Frame 4 where $N-1$ dots are reached, we see that RTs for *more than k* and *at most k* are higher than for *at least $k+1$* and *fewer than $k+1$* . In Frame 5, the opposite is true. A repeated measures ANOVA with Frame and n as factors revealed a significant interaction ($F(1,24) = 29.078$; $p < .001$) indicating that n , the number mentioned, and not N , the critical number needed to verify, affects the verification process. If it were N , which is the same across all four quantifiers, we would see a main effect of Frame and no interaction.

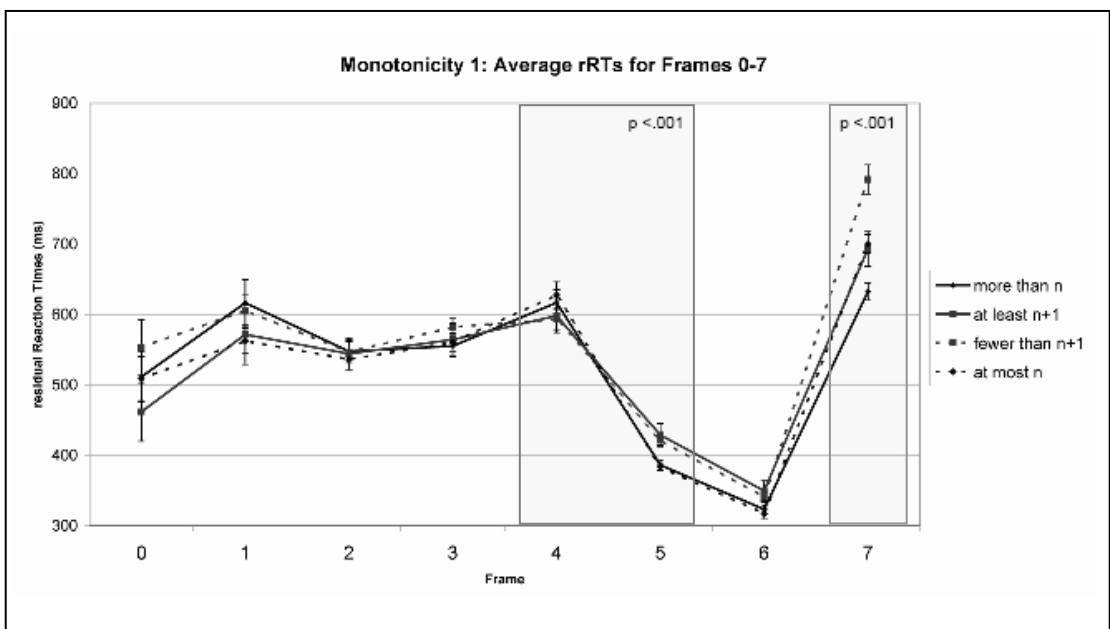


Figure 2: Results of Experiment 1

The second area of interest is the answer frame, Frame 7, where we observe two main effects. First, we observe a main effect of monotonicity ($F(1,24) = 17.09$; $p < .001$) such that monotone decreasing quantifier take more time to answer than monotone increasing ones. This is in line with Barwise and Cooper (1981). The second main effect is an effect of n again ($F(1,24) = 13.228$; $p < .001$) such that *at least $k+1$* and *fewer than $k+1$* take significantly longer than their counterparts *more than k* and *at most k* . Since in half of the items Frame 7 contains the $k+1^{\text{th}}$ dot, this main effect is reminiscent of what we have seen in Frame 4, such that reaching the number heard in the quantifier causes a slow down. Again this is unexpected under GQT.

4.3. Experiment 2

In Experiment 1, we saw a main effect of Monotonicity such that statements with decreasing quantifiers are harder to answer than ones with increasing quantifiers. Recall that we collapsed across true and false answers and thus did not report any effects of truth. In fact, the design of Experiment 1 did not allow us to investigate whether it is in fact easier to verify increasing quantifiers than decreasing ones and whether it is easier to falsify decreasing quantifiers than increasing ones. This is so because when an additional blue dot was seen statements with increasing quantifiers were true

while statements with decreasing quantifiers were false and so any possible effects of truth could in fact have been due the presence or absence of an additional blue dot. Experiment 2 avoids this confound.

Methods and Materials: Target and filler items in Experiment 2 were identical to those in Experiment 1 except that all 8 target conditions (4 determiners by true/false) were matched to an identical array (e.g. total number of blue dots was 7). Thus the number heard, n , varied across true and false items e.g. *More than 6* (true)/*More than 7* (false), *At least 7* (true)/*At least 8* (false), *Fewer than 8* (true)/*Fewer than 7* (false), *At most 7* (true)/*At most 6* (false). We used 15 different arrays, for a total 120 target items. In all arrays only $N-2$ dots (in the cases above, 5 blue dots) were seen by Frame 6, which ensured that the number mentioned, n , was not reached until the last frame. On the last frame, either one or two blue dots were revealed, determining truth or falsity. This sidestepped the previous confound by ensuring that at least one dot in the target color was seen on the answer frame for both true and false items⁵.

Results: As can be seen in Figure 3, which gives average reaction times for 12 subjects on the answer frame (Frame 7), Truth and Monotonicity interacted as expected ($p < .05$). Specifically, monotone increasing quantifiers (*more than/at least*) are quicker to verify than falsify, monotone decreasing quantifiers (*fewer than/at most*) are quicker to falsify than verify, monotone increasing quantifiers are quicker to verify than monotone decreasing, and monotone decreasing quantifiers are quicker to falsify than monotone increasing quantifiers.

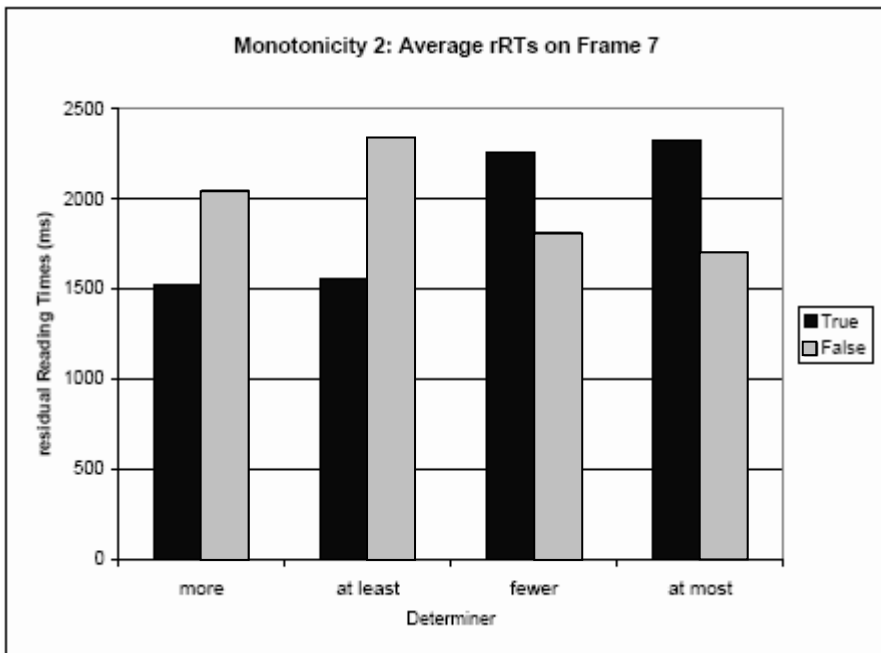


Figure 3: Results of Experiment 2.

Discussion of Experiments 1 and 2: Verifying quantified statements of the sort used in these experiments clearly involved at least two separate stages – the counting stage, which involves counting the number of target dots, and the decision stage, where based on the number of targets seen, a decision about truth or falsity is made. Experiment 1 revealed that the counting stage is affected by the number mentioned in the quantifier, rather than the critical number of dots needed to verify the statement. Specifically, we saw that participants' response times for denotationally equivalent quantifiers varied such that there was a slow down on frames where the total dot count equaled the

⁵ In pilot studies, we found that there was no significant difference between conditions where 1 versus 2 target dots were seen in a given frame, while there was a significant difference between 0 and 1.

number mentioned in the quantifier. A purely denotational, Generalized Quantifier Theory based perspective would instead predict that the counting stage would be affected only by the critical number needed to verify, thus expecting no difference between denotationally equivalent quantifiers. We also saw that the decision stage is affected by the monotonicity properties of a given quantifier. Specifically we saw that decreasing quantifiers are harder to verify than increasing ones when collapsing across true and false answers (Experiment 1), and that monotonicity interacts with truth (Experiment 2), such that decreasing quantifiers are easier to falsify while increasing quantifiers are easier to verify.

It is clear, then, that GQT alone does not provide an adequate framework for understanding verification procedures for quantified statements and that even the processing amendment given in Barwise and Cooper (1981) together with an economy principle can characterize only the decision stage adequately. It does not characterize how the linguistic form of a particular quantifier affects the counting procedure that comprehenders use to gather the necessary information. Thus GQT seems to be too coarse to provide an adequate base for a processing account of modified numeral quantifiers.

5. Conclusion

Formal semantic analyses focus their attention on capturing the correct set of entailment patterns associated with a particular expression. They usually do not concern themselves with comprehension or verification in real time. However, semantic theories do have the responsibility to furnish the pieces that a processing theory requires to draw systematic distinctions that occur during real time comprehension. Whether a particular theory manages to live up to this obligation depends to a large extent on what type of semantic primitives it assumes, since the “size” of the primitives determines the resolution with which the theory can analyze its phenomena.

For the modified numeral quantifiers considered here the standard treatment is provided by Generalized Quantifier Theory, which takes the denotations of these expressions to be semantic primitives. This assumption entails that the properties of the component parts of modified numeral quantifiers do not affect the quantifiers’ external behavior. The experimental evidence presented here reveals that a GQT-based approach is too coarse to fully account for the behavior of modified numeral quantifiers during verification. The evidence instead supports a decompositional approach that recognizes degrees (numerals) and degree operators (*more than*, ...) among the basic building blocks of quantification in natural language.

Contact: Martin.Hackl@pomona.edu
Jorie.Koster-Moeller@pomona.edu

Acknowledgements

We would like to thank the audiences at CUNY 21, WCCFL 27 for helpful comments and suggestions, and especially Andrea Gottstein for her help with the poster and slides. This research was funded through NSF grant 0642748.

References

- Barwise, Jon and Robin Cooper (1981): ‘Generalized Quantifiers and Natural Language’, *Linguistics and Philosophy* 4: 159-219.
- Benthem, Johan van. (1986): *Essays in Logical Semantics*, Dordrecht: Reidel Publishing Co.
- Fox, Danny and Martin Hackl. (2006): ‘On the universal density of measurement’, *Linguistics and Philosophy* 29:537-586.
- Geurts, Bart and Rick Nouwen (2007): ‘At least” et al.: the semantics of scalar modifiers’, *Language* 83: 533-559.
- Geurts, Bart (2007): Experimental support for a modal analysis of "at least" and "at most". In: Proceedings of the ESSLI workshop on quantifier modification
- Hackl, Martin (2000): *Comparative Quantifiers*, PhD dissertation MIT, MITWPL, Cambridge, MA.
- Hackl, Martin (forthcoming): ‘On the Grammar and Processing of Proportional Quantifiers: Most versus More Than Half’, *Natural Language Semantics*

- Keenan, Edward and Jonathan Stavi (1986): 'A Semantic Characterization of Natural Language Determiners', *Linguistics and Philosophy* 9: 253-326.
- Krifka, Manfred (1999): 'At least some Determiners aren't Determiners', in Ken 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
- Mostowski, Andrzej (1957): 'On a Generalization of Quantifiers', *Fundamenta Mathematicae* 44: 12-36.

Proceedings of the 27th West Coast Conference on Formal Linguistics

edited by Natasha Abner
and Jason Bishop

Cascadilla Proceedings Project Somerville, MA 2008

Copyright information

Proceedings of the 27th West Coast Conference on Formal Linguistics
© 2008 Cascadilla Proceedings Project, Somerville, MA. All rights reserved

ISBN 978-1-57473-428-7 library binding

A copyright notice for each paper is located at the bottom of the first page of the paper.
Reprints for course packs can be authorized by Cascadilla Proceedings Project.

Ordering information

Orders for the library binding edition are handled by Cascadilla Press.
To place an order, go to www.lingref.com or contact:

Cascadilla Press, P.O. Box 440355, Somerville, MA 02144, USA
phone: 1-617-776-2370, fax: 1-617-776-2271, e-mail: sales@cascadilla.com

Web access and citation information

This entire proceedings can also be viewed on the web at www.lingref.com. Each paper has a unique document # which can be added to citations to facilitate access. The document # should not replace the full citation.

This paper can be cited as:

Koster-Moeller, Jorie, Jason Varvoutis, and Martin Hackl. 2008. Verification Procedures for Modified Numeral Quantifiers. In *Proceedings of the 27th West Coast Conference on Formal Linguistics*, ed. Natasha Abner and Jason Bishop, 310-317. Somerville, MA: Cascadilla Proceedings Project.

or:

Koster-Moeller, Jorie, Jason Varvoutis, and Martin Hackl. 2008. Verification Procedures for Modified Numeral Quantifiers. In *Proceedings of the 27th West Coast Conference on Formal Linguistics*, ed. Natasha Abner and Jason Bishop, 310-317. Somerville, MA: Cascadilla Proceedings Project. www.lingref.com, document #1845.