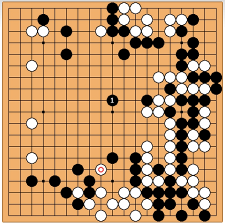


Pragmatic effects in processing superlative and comparative quantifiers: epistemic-algorithmic approach

Maria Spychalska, Institute of Philosophy II,

Ruhr-University Bochum

September 27, 2013

Introduction

Superlative quantifiers

- **At most three** people came to the party.
- Anna ate **at least three** apples.

GQT

- **at most n** A are B $\iff$ **fewer than n+1** A are B $\iff$ **not more than n** A are B $\iff$ **(exactly) n** of **fewer than n** A are B
- **at least n** A are B $\iff$ **more than n-1** A are B $\iff$ **not fewer than n** A are B $\iff$ **(exactly) n** of **more than n** A are B

Difference between comparative and superlative quantifiers in NL

- Linguistic use
- Acquisition
- Processing
- Reasoning

Differences in inference patterns

- Majority of responders reject inferences from **at most n** to **at most n+1**, [2-14% in (Geurts et al., 2010), (Cummins & Katsos, 2010)] e.g.

$$\frac{\text{At most 3 students came to the class.}}{\text{At most 4 students came to the class.}}$$

- They however do accept presumably equivalent inferences with comparative quantifiers (60-70%)

$$\frac{\text{Fewer than 4 students came to the class.}}{\text{Fewer than 5 students came to the class.}}$$

Different processing time

- Superlative quantifiers require longer verification time that respective comparative quantifiers (Geurts et al., 2010),(Koster-Moeller et al, 2008).
- Downward monotone quantifiers are verified slower, while upward monotone quantifiers are falsified slower (Koster-Moeller et al, 2008).
- Time: at most $>$ at least $\approx>$ fewer than $>$ more than $>$ exactly

Geurts' modal account

Comparative and superlative quantifiers are not semantically equivalent (Geurts & Nouwen, 2007), (Geurts et al., 2010):

- **more than n** and **less than n** have their conventional meaning defined in terms of inequality relation,
- the semantics of **at least n** and **at most n** have a modal component, namely:
    - $\textit{at most } n(A, B) \Rightarrow \textit{possibly exactly } n\,(A, B)$
    - $\textit{at least } n(A, B) \Rightarrow \textit{possibly more than } n\,(A, B)$

Problems of modal account

Superlative quantifiers embedded in conditional and various other contexts (deontic): (Geurts & Nouwen, 2007),(Geurts et al., 2010)

> *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 2** to **at most 3**, are commonly accepted by people (over 96% in Geurts's experiment).

Coherence judgements Cummins & Katsos (2010)

John has at most 5/at least 5 houses. Specifically he has exactly 4/6 houses.

are judged more coherent than semantically self-contradictory cases, e.g.

John has at most 5/at least 5 houses. Specifically he has exactly 6/4 houses.

Introduction

Clausal implicature

Cummins & Katsos (2010): **at most n** and **at least n** both imply **possibly exactly n**, but this is a pragmatical rather than a logical inference, namely a co-called **clausal implicature**. These implicatures are based on the fact that the semantics of both **at most n** and **at least n** can be represented in a disjunctive form.

$$\textit{at most } n(A, B) \iff \textit{n or fewer than n}\,(A, B)$$
$$\textit{at least } n(A, B) \iff \textit{n or more than n}\,(A, B)$$

Clausal Implicature

- Quantity Implicature concerning the truth of a proposition expressed in a particular subclause. The addressee infers that the proposition may or may not be true.
    - *weaker form:* **p or q**
    - $(p \vee q) \nrightarrow p$, $(p \vee q) \nrightarrow q$ *but* $p \rightarrow (p \vee q)$, $q \rightarrow (p \vee q)$
    - *stronger form:* **p and q**
    - $(p \wedge q) \rightarrow p$, $(p \wedge q) \rightarrow q$
    - *Clausal implicature of disjunction:* $\{P_S p, P_S \neg p, P_S q, P_S \neg q\}$

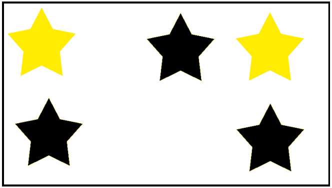
Clausal implicatures of superlative quantifiers

at most n As are B $\Longleftrightarrow$ exactly n or fewer than n As are B

Implicatures:

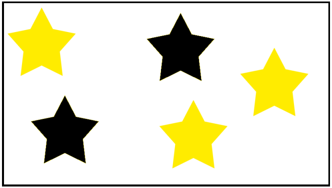
- It is possible that there are exactly n As that are B.
- It is possible that fewer than n As are B.

At most three stars are yellow



True Infelicitous?

At most three stars are yellow



True

Asymmetry between the two implicatures

- **Fewer than n** is still possible when there are n, in a sense of a **subset**.
- There is some hierarchy of these implicatures or, as I propose, an order in checking the conditions.

First assumption

- Epistemic/non-descriptive character of superlative quantifiers
- Analogous to disjunction

There is a cat or a ball.

Epistemic interpretation of disjunction

According to Zimmermann (2000) 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].*

Thus, disjunctive sentences in natural language are interpreted as **conjunctive lists of epistemic possibilities**.

Epistemic interpretation of superlative quantifiers

- Pragmatic enrichment of the reading of **at most n**
- **at most n** $\rightsquigarrow$ **possibly exactly n** AND **possibly fewer than n**
- “Both are agreeable with my knowledge”

Second assumption: Meaning as an algorithm

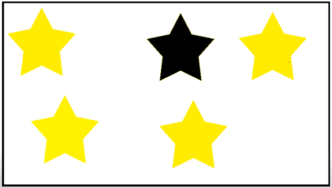
- Meaning of a quantifier as a pair $\langle C_F, C_V \rangle$, where C_V is a **verification condition** (specifies how to verify sentences with this quantifier) and C_F is a **falsification condition** (specifies how to falsify sentences with this quantifier).
- Verification and falsification conditions are to be understood **algorithmically** (as partial algorithms), 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.

Falsification of at most n

- Krifka (1999) observes, that a sentence **at most n x:** $\phi(x)$[1] says only that **more than n x:** $\phi(x)$ is false, and leaves a truth condition underspecified.
- If find **more than n (A,B)**, then falsify.

[1] at most n x are ϕ

Falsification: At most three stars are yellow



Verification condition: at most n (A,B)

- **At most n (A,B)** $\Longleftrightarrow$ **exactly n (A,B)** or **fewer than n (A,B)**
- **at most n (A,B)** $\rightsquigarrow$ B_S **exactly n** AND B_S **fewer than n**
- Then: **at most n(A,B)** $\not\rightarrow$ **at most n+1 (A,B)**

Introduction

Verification

1. **at most n (A,B)** $\rightsquigarrow$ B_S **exactly n** AND B_S **fewer than n**
2. *If exactly n, then accept;*
   Else;
   If fewer than n, then accept.
3. Model: there are n (A,B)
4. Update of belief set: K_S **exactly n** (AND **not** B_S **fewer than n**)

Verification

1. **at most n (A,B)** $\rightsquigarrow$ B_S **exactly n** AND B_S **fewer than n**
2. Model: there are n-1 (A,B)
3. Update of belief set: K_S exactly n-1
4. cancel the first option: **not** B_S **exactly n**
5. exactly n-1 is fewer than n, hence K_S **fewer than n**

Comparative: fewer than n

1. Model: there are n-1 (A,B)
2. Update of belief set: K_S exactly n-1
3. exactly n-1 is fewer than n, hence K_S **fewer than n**

Downward monotone context

1. 1. (a) If at most six people come, then we will be able to fit them all.
   1. (b) At most five people come.
   1. (c) We will be able to fit them all.
2. The only case when you reject the consequent of (a) is when the antecedent of (a) is false.
3. But (b) does not allow us to falsify the antecedent, thus we accept (c).

Introduction

At least also have some problems in UP contexts but not in DM

1. 1. (a) If Anna comes before midnight, then she will get at least 3 drinks.
   1. (b) Anna came before midnight.
   1. (c) She will get at least 2 drinks.
2. 2. (a) If at least six people come, then we will have a show.
   2. (b) At least seven people come.
   2. (c) We will have a show.

Table: Logically correct inferences

		Premise	Conclusion	Percentage	total	total
A	At least	at least 4	at least 3	80,6%	79,4%	**75,7%**
		at least 3	at least 2	77,7%		
		at least 8	at least 6	69,4%	72,2%	
		at least 7	at least 5	75 %		
B	At most	at most 4	at most 5	16,7%	12,5%	**13,9%**
		at most 3	at most 4	8,3%		
		at most 8	at most 10	16,7%	15,3%	
		at most 7	at most 9	13,9%		
C	Not fewer than	not fewer than 4	not fewer than 3	52,8	55,6%	**63,2%**
		not fewer than 3	not fewer than 2	58,3		
		not fewer than 8	not fewer than 6	61,1	70,8%	
		not fewer than 7	not fewer than 5	80,6		
D	Not more than	not more than 4	not more than 5	30,6%	27,8%	**32%**
		not more than 3	not more than 4	25%		
		not more than 8	not more than 10	38,9%	36,1%	
		not more than 7	not more than 9	33,3%		
E	N or more than n	4 or more than 4	3 or more than 3	72,2%	70,8%	**65,3%**
		3 or more than 3	2 or more than 2	69,4%		
		8 or more than 8	6 or more than 6	58,3%	59,7%	
		7 or more than 7	5 or more than 5	61,1%		
F	N or fewer than n	4 or fewer than 4	5 or fewer than 5	13,9	18%	**16,7%**
		3 or fewer than 3	4 or fewer than 4	22,2		
		8 or fewer than 8	10 or fewer than 10	13,9	15,3%	
		7 or fewer than 7	9 or fewer than 9	16,7		

Table: Logically correct inferences

M	More than n-1	more than 3	more than 2	83,3%	77,8%	**80,6%**
		more than 2	more than 1	72,2%		
		more than 7	more than 5	83,3%	83,3%	
		more than 6	more than 4	83,3%		
N	fewer than n+1	fewer than 5	fewer than 6	52,8	58,7%	**57,6%**
		fewer than 4	fewer than 5	66,7		
		fewer than 9	fewer than 11	47,2	55,6%	
		fewer than 8	fewer than 10	63,9		
K	Numerical	4	3	58,3	58,3%	**59%**
		3	2	58,3		
		8	6	69,4	59,7	
		7	5	50		

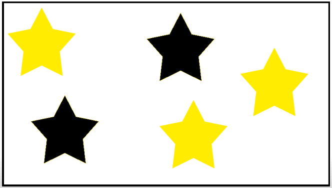
Introduction

At least n (and more)

- Falsification: if can't find n, then falsify
- Verification: **at least n** $\rightsquigarrow$ **possibly more than n** AND **possibly exactly n**
- Reversed order: monotonicity effect **fewer than n** $\subseteq$ **n**

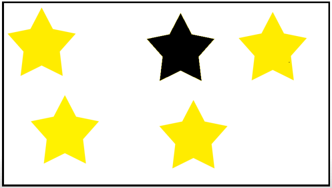
  n $\subseteq$ **more than n**

Verification: at least three stars are yellow



Introduction

Verification: at least three stars are yellow



Experiment

- Sentence-picture verification experiment with reaction time measure
- **Goal:** compare subjects' accuracy and response time in evaluating sentences with superlative quantifiers **at most three/ at least three**, and their equivalent comparative and disjunctive forms in models, in which they might involve different algorithmic procedures.

- Participants: 56 (29 women) right-handed German native speakers (mean age: 24.25, SD: 4.26)
- Quantifier forms

Quantifiers	superlative	comparative	disjunctive	negative comparative
Upward monotone	at least three	more than two	three or more than three	not fewer than three
Downward monotone	at most three	fewer than four	three or fewer than three	not more than three
Bare numeral	three			

- Models

Number of target objects	1	2	3	4	5
Upward monotone	False	False	True	True	True
Downward monotone	True	True	True	False	False

Hypotheses

[**Modal account**:] If superlative quantifiers have modal semantics, which means that the sentence ϕ: **At most n As are B** logically implies ψ: **It is possible that there are exactly n As that are B**, then ϕ should be rejected in models in which there are fewer than *n* **As** that are **B**. Since ψ, which is a logical consequence of ϕ, is false in such models, then ϕ cannot be true in those models either.

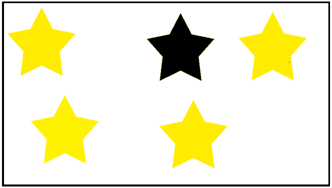
Hypotheses

[**Pragmatic account:**] If ψ is merely a pragmatic inference from ϕ, then it should be defeasible and ϕ should be evaluated as true in models that have fewer than n **As** that are **B**. But the evaluation of superlative quantifier will take longer time /result in higher mistakes ration than the evaluation of comparative quantifiers due to the need for cancelling the clausal implicature.

Hypotheses

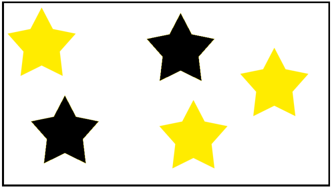
[**Epistemic-algorithmic account:**] The evaluation of sentences with superlative quantifiers will result in a longer response-time and/or a higher mistakes ratio compared to the evaluation of sentences with comparative quantifiers only in those models in which the use of superlative quantifiers requires cancelation of the “first” of epistemic possibilities, i.e. the models that have fewer than N **As** that are **B**, where N is the borderline of the sentence’s truth-conditions ($N = n$ for the quantifiers **at most n**).

At most three stars are yellow.



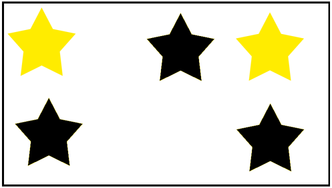
False

At most three stars are yellow.



True and easy

At most three stars are yellow.



True and hard

Accuracy results

- The analysis of subjects' accuracy revealed that subjects generally accepted (**at most 3**/**at least n**) (mean accuracy in each model over 2.6, value range: $0 - 3$) in all models in which they are semantically true.
- For **at most 3**: they made significantly more mistakes in algorithmically "harder" models, i.e. models with 1 target objects ($z = -3.392$, $p = .001$, $r = -.320$), or with 2 target objects ($z = -2.324$, $p = .02$, $r = -.219$) than in the "easy" ones, i.e. with 3 target objects.

Accuracy

- A similar effect was obtained for the DM disjunctive form: subjects were significantly more often correct when accepting this form in the models with 3 target objects compared to models with 1 target object ($z = -2.840$, $p = .005$, $r = -.268$), though not compared to models with 2 target objects ($p = .058$).
- For the upward monotone group, the differences between subjects' correctness in felicitous and infelicitous models, when corrected for multiple comparisons, were not significant for the superlative or the disjunctive form.

Reaction time analysis

- Repeated Measure (**Mon** (2) × **Qform** (4) × **Mform** (5)):
- To calculate the mean reaction time only those trials were taken into account in which a subject gave a correct response.
- Greenhouse-Geisser correction was applied in those cases in which the assumption of sphericity was not met.
- Bonferroni correction was applied familywise for pairwise comparisons, and all the reported *p*-values are already corrected.

Reaction time analysis

- Monotonicity (**Mon**) ($F(1,49) = 34.681$, $p < .001$, $\eta^2 = .414$) and Quantifier form (**Qform**)($F(2.25, 110.27) = 45.093$, $p < .001$, $\eta^2 = .479$) had a significant effect on subjects' time taken to respond correctly.
- Pairwise comparisons for **Qform** showed, however, that only the negative comparative form was significantly slower evaluated than every other form ($p < .001$), but the comparisons between other forms were not significant.

Reaction time analysis

- All interactions between our three factors (monotonicity, quantifier, model) turned out significant ($p < .001$).

  Mon $\times$ **Qform**: $F(2.06, 100.9) = 23.443$, $p < .001$, $\eta^2 = .324$,

  Mform $\times$ **Mon** $F(3.2, 157.35) = 13.407$ $p < .001$, $\eta^2 = .215$,

  Mform $\times$ **Qform**: $F(6.86, 336.4) = 5.596$, $p < .001$, $\eta^2 = .103$,

  Mon $\times$ **Qform** $\times$ **Mform**: $F(6.96, 340.86) = 6.26$, $p < .001$, $\eta^2 = .116$.

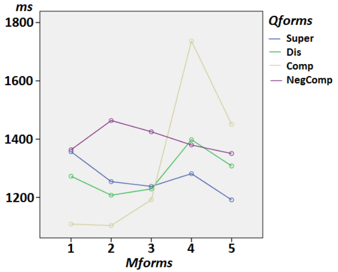
RM for each Mform and Mon

- We compare separately the RT in evaluating various forms in each model for each monotonicity.
- The time of evaluating disjunctive and superlative quantifiers (of both monotonicities) did not differ significantly in any of the models.
- The reaction time in evaluating superlative and comparative forms **differed only in some models:** note the differences with respect to the monotonicity (see below)

Results

Downward monotone quantifiers

Whereas in the “hard models” (1 or 2 target objects) the superlative quantifier (**at most three**) was evaluated significantly slower than the comparative one (**fewer than four**) ($p < .008$), in the “easy models” (3 target objects) there was no significant difference in the reaction time taken to verify these two forms.



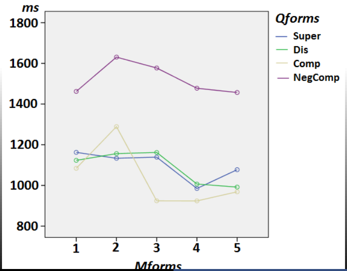
Results

Downward monotone

Model	1	2	3	4	5
Qform	F(3,159)= $= 7.462$ $\eta^2 = .123$ $p < .001$	F(2.195,120.713)= $= 11.545$ $\eta^2 = .173$ $p < .001$	F(1.54,83.134)= $= 13.781$ $\eta^2 = .203$ $p < .001$	F(2.445,134.49)= $=16.247$ $\eta^2 = .228$ $p < .001$	F(2.513,135.729)= $=3.88$ $\eta^2 = .067$ $p = .015$
sup vs. comp	$p < .001$	$p = .007$	–	$p < .001$	$p = .052$
sup vs. dis	–	–	–	–	–
sup vs. neg comp	–	—	$p = .002$	$p = .002$	–
comp vs. neg comp	$p = .001$	$p < .001$	$p = .001$	$p = .001$	–
comp vs. dis	–	–	—	$p < .001$	–
dis vs. neg comp	–	$p = .008$	$p = .001$	$p < .001$	–

Upward monotone quantifiers

In the models with 3 target objects there was a significant difference between the comparative and the other quantifiers: the comparative quantifier was evaluated significantly faster ($p < .004$). In the models with 4 or 5 target objects, there was, however, no difference between the processing time of the superlative and of the comparative quantifiers.



Results

Upward monotone quantifiers

Model	1	2	3	4	5
Qform	F(2.194,120.680) =17.227 $\eta^2 = .239$ $p < .001$	F(2.194,121.661) =18.201 $\eta^2 = .256$ $p < .001$	F(2.03,109.598) =29.875 $\eta^2 = .356$ $p < .001$	F(1.754,96.462) =23.374 $\eta^2 = .298$ $p < .001$	F(2.081,114.438) =26.222 $\eta^2 = .323$ $p < .001$
sup vs. comp	–	$p = .037$	$p < .004$	–	–
sup vs. dis	–	–	–	–	–
sup vs. neg comp	$p < .001$	$p < .001$	$p < .001$	$p < .001$	$p < .001$
comp vs. neg comp	$p < .001$	$p = .005$	$p < .001$	$p < .001$	$p < .001$
comp vs. dis	–	–	$p = .001$	–	–
dis vs. neg comp	$p < .001$	$p < .001$	$p < .001$	$p < .001$	$p < .001$

Conclusions

1. The modal semantics for superlatives cannot be supported
2. The pragmatic approach, where the possibility of **exactly n (A,B)** is a pragmatic inference from **at most n (A,B)**, is supported
3. The experiment shows that a more refined theory is needed, that takes into account verification procedures, in order the explain the difference between the felicity of “at most n” in models that have exactly n and those that have fewer than n target elements.

Acknowledgments

- Jarmo Kontinen (programming)
- Theresa Garwels (translation)

Bibliography I

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.

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.

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

Bibliography II

Bart Geurts, Napoleon Katsos, Chris Cummins, Jonas Moons, and Leo Noordman. 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.

Results

Bibliography III

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

Thomas Ede Zimmermann. 2000. Free Choice Disjunction and Epistemic Possibility, Natural Language Semantics, 8 (2000), 255 - 290.