Witness Sets, Polarity Reversal and the Processing of Quantified Sentences

Oliver Bott[1], Udo Klein[2] and Fabian Schlotterbeck[1]

[1] Tübingen University, Tübingen, Germany
{oliver.bott,fabian.schlotterbeck}@uni-tuebingen.de
[2] Bielefeld University, Bielefeld, Germany
udo.klein@uni-bielefeld.de

Experimental results show that the monotonicity of a quantifier (Q) affects how it is processed [4, 7, 6]. $Q's$ are upward entailing (UE) if they permit inferences to supersets, i.e. from $Q\ A\ B$ follows $Q\ A\ B'$ for any $B \subseteq B'$ (e.g. *every*, *at least five*). If from $Q\ A\ B$ it follows that $Q\ A\ B'$ for any $B' \subseteq B$, the quantifier Q is downward entailing (DE) (e.g. *no*, *at most five*). As compared to UE Q's, DE Q's are more difficult to verify [4, 7] and draw inferences from [6]. Recent attempts at deriving predictions about the processing of Q's are built on the notion of semantic automata [12], and have been presented e.g. in [11] and [10]. E.g., [11] improve on results by [8] by showing not only that the computational distinction between Q's recognized by finite-automata and push-down automata is psychologically relevant, but also that there are differences in the time required for verifying statements involving Q's even among the class of quantifiers recognized by finite state automata. However, since these approaches employ essentially the same kind of semantic automata for both DE and UE Q's, they cannot explain why DE Q's are more difficult to process than UE Q's. To explain this, we formulate a quantification theory which predicts that the (expansion) operation employed in processing DE Q's is more difficult than that for processing UE Q's, because it involves (i) inferences from negative information and (ii) polarity reversal. The predictions of our account were tested in two experiments investigating the online comprehension and verification of simply (Exp. 1) and doubly quantified sentences (Exp. 2) with UE vs DE Q's.

1 An algorithmic theory of quantification

In the following, we aim to provide an algorithmic theory of meaning in the sense of [9]: The meaning of an expression is the algorithm which computes its denotation. Here is the basic idea of what will be worked out below. In the case of a quantificational sentence Q_1 *boys tickled* Q_2 *girls*, with Q_1 *boys* scoping over Q_2 *girls*, the interpretation consists in specifying an algorithm which outputs whether S is true given <u>any</u> model, and the verification consists in applying this algorithm to <u>a specific</u> model.

Given a binary relation R (e.g. *tickle*), we first construct a polarity relation R^*, such that $\langle a, b, + \rangle \in R^*$ if $\langle a, b \rangle \in R$; and $\langle a, b, - \rangle \in R^*$ otherwise. We now specify an algorithm which derives $\langle Q_1\ boys, Q_2\ girls, + \rangle$ if and only if our input sentence is true (relative to a certain scoping). The algorithm for the verification of the doubly quantified sentence above (relative to the subject wide scope reading) consists of the following steps:

(1) a. For every *boy* x who *tickles* some girl, determine which *girls* he *tickles*. If the set of girls tickled by x is a witness set of Q_2 *girls*, add the positive information $\langle x, Q_2\ girls, + \rangle$. Otherwise, add the negative information that $\langle x, Q_2\ girls, - \rangle$.

b. For every *boy* x who tickles no girl, add $\langle x, Q_2\ girls, + \rangle$ if Q_2 *girls* has the empty set as a witness set. Otherwise, add the negative information $\langle x, Q_2\ girls, - \rangle$.

(2) a. If for some boy x we have $\langle x, Q_2\ \textit{girls}, +\rangle$, and the set of boys y with $\langle y, Q_2\ \textit{girls}, +\rangle$ is a witness set of $Q_1 boys$, then add $\langle Q_1 boys, Q_2 girls, +\rangle$, else $\langle Q_1 boys, Q_2 girls, -\rangle$.
b. If for every boy x we have $\langle x, Q_2\ \textit{girls}, -\rangle$, and the empty set is a witness set of $Q_1 boys$ then add $\langle Q_1 boys, Q_2 girls, +\rangle$. Otherwise, add $\langle Q_1 boys, Q_2 girls, -\rangle$

Q_1 *boys tickled* Q_2 *girls* is true iff $\langle Q_1 boys, Q_2 girls, +\rangle$ can be added. Importantly, if neither Q_1 *boys* nor Q_2 *girls* have the empty set as a witness set, then the general algorithm can be simplified as follows:

(3) For every *boy* x who *tickles* some girl, determine which *girls* he *tickles.* If the set of girls tickled by x is a witness set of Q_2 *girls*, add the positive information $\langle x, Q_2\ \textit{girls}, +\rangle$.

(4) If for some boy x we have $\langle x, Q_2\ \textit{girls}, +\rangle$, and the set of boys y with $\langle y, Q_2\ \textit{girls}, +\rangle$ is a witness set of $Q_1 boys$ then add $\langle Q_1 boys, Q_2 girls, +\rangle$.

In other words, if neither quantifier is an *empty-set quantifier*, the negative information can safely be ignored in every model, and the algorithm can be restricted to positive information. We will refer to this simplified procedure as *simple expansion* (**s-exp**). If one (or both) quantifiers are empty-set quantifiers, and the antecedent of (2b) holds, then negative information becomes relevant, since step (2b) allows the addition of positive information based on negative information, and therefore the more complex algorithm (**c-exp**) is required. The differentiation between **s-exp** and **c-exp** sets our theory apart from existing algorithmic proposals of quantification, as for instance [10]'s automata theory building on [12].

2 Quantification theory

First, we assume that simple NL determiners denote unary functions from restrictor sets to pairs consisting of the restrictor set itself and a set of subsets of the restrictor set, called the set of witness sets, i.e. as functions $q : \mathcal{P}(E) \mapsto \mathcal{P}(E) \times \mathcal{P}(\mathcal{P}(E))$.

Definition 1 (w-function, w-quantifier, witness sets): *Let q be a function from the set $\mathcal{P}(E)$ of subsets of the domain of entities E into $\mathcal{P}(E) \times \mathcal{P}(\mathcal{P}(E))$. Then q is called a w-function*[1] *if for any $A \subseteq E$ there is a $W \subseteq \mathcal{P}(A)$, such that $q(A) = \langle A, W\rangle$. If q is a w-function and $A \subseteq E$, then $q(A) = \langle A, W\rangle$ is called a w-quantifier, and W the set of witness sets of q at A. If a w-quantifier $q(A) = \langle A, W\rangle$ is such that $\emptyset \in W$, it is called an empty-set quantifier.*

To illustrate, for any subset A of the domain of individuals E, we have:

(5)
$$\begin{array}{lcl} [\![\textit{some}]\!](A) & = & \langle A, \{X : X \subseteq A \wedge |X| \geq 1\}\rangle \\ [\![\textit{most}]\!](A) & = & \langle A, \{X : X \subseteq A \wedge |A \cap X| > |A - X|\}\rangle \end{array}$$

Note that these w-quantifiers naturally correspond to the standard generalized quantifier denotations:

(6)
$$\begin{array}{lcl} [\![\textit{some}]\!]_E(A) & = & \{B : B \subseteq E \wedge |A \cap B| \geq 1\} \\ [\![\textit{most}]\!]_E(A) & = & \{B : B \subseteq E \wedge |A \cap (A \cap B)| > |A - (A \cap B)|\} \end{array}$$

Importantly, note that for most type $\langle 1, 1\rangle$ quantifiers there is no corresponding w-quantifier:

(7)
$$\begin{array}{lcl} EQUI_E(A) & = & \{B : B \subseteq E \wedge |A| = |B|\}\rangle \\ TOTAL_E(A) & = & \{B : B \subseteq E \wedge E = A\}\rangle \end{array}$$

[1] Mnemonic for witness-set function.

In fact, a type $\langle 1,1\rangle$ quantifier Q has a corresponding w-quantifier if and only if Q satisfies both conservativity (CONS) and extension (EXT).[2]

In order to explicitly encode positive and negative information we define the polarity relation P^* of an n-ary predicate: if $\langle a,b\rangle \in P$ then $\langle a,b,+\rangle \in P^*$, and if $\langle a,b\rangle \notin P$ then $\langle a,b,-\rangle \in P^*$.

Definition 2 (**n-ary predicate**, **polarity relation**)**:** *A subset P of $(E \cup \mathcal{P}(E) \times \mathcal{P}(\mathcal{P}(E)))^n$, with $n \geq 1$, is called an n-ary predicate. For any n-ary predicate P and $\pi_i(\sigma)$ the i-th element of an n-ary tuple σ (for any $1 \leq i \leq n$), let $P^* = \{\langle \pi_1(\sigma),\ldots,\pi_n(\sigma),+\rangle : \sigma \in P\} \cup \{\langle \pi_1(\sigma),\ldots,\pi_n(\sigma),-\rangle : \sigma \notin P\}$ be the polarity relation of P.*

To illustrate, if $E = \{a,b,c\}$, $A = \{a,b\}$, and $P = \{\langle a,a\rangle,\langle a,b\rangle,\langle c,c\rangle\}$ is a binary predicate, then $P^* = \{\langle a,a,+\rangle,\langle a,b,+\rangle,\langle c,c,+\rangle\} \cup \{\langle a,c,-\rangle,\langle b,a,-\rangle,\langle b,b,-\rangle,\langle b,c,-\rangle,\langle c,a,-\rangle,\langle c,b,-\rangle\}$ is polarity relation of P.

Next, we define the set of i-fillers $[\sigma]_i^{P^*}$ of a tuple $\sigma \in P^*$ as the set of elements in the i-th position of an i-variant $\tau \in P^*$ of σ (τ is an i-variant of σ iff τ and σ differ at most at the i-th position).

Definition 3 (**i-fillers**)**:** *Let P^* be the polarity relation of some n-ary predicate P and let $\sigma \sim_i \tau$ hold iff σ and τ differ at most at the i-th element. For every $\sigma \in P^*$ and any integer $1 \leq i \leq n$, let $[\sigma]_i^{P^*} = \{\pi_i(\tau) : \tau \in P^* \wedge \tau \sim_i \sigma\}$ be the i-fillers of σ in P^*. Let $[\sigma]_i^{P^*,A} = [\sigma]_i^{P^*} \cap A$ be the set of those i-fillers of σ in P^* which are also in the set A (called restrictor set).*

For example, if $P^* = \{\langle a,a,+\rangle,\langle a,b,+\rangle,\langle b,a,-\rangle,\langle b,b,+\rangle\}$, then $[\langle a,a,+\rangle]_1^{P^*} = \{a\}$, $[\langle a,a,+\rangle]_2^{P^*} = \{a,b\}$ and $[\langle b,a,-\rangle]_1^{P^*} = \{b\}$, $[\langle b,a,-\rangle]_2^{P^*} = \{a\}$. Finally, we formulate two expansion operations which, when applied to a quantifier $q(A)$ and a polarity relation P^*, add those tuples representing negative or positive information involving $q(A)$ at position i in P^*.

Definition 4 (**simple and complex i-expansion**)**:** *Let $q : \mathcal{P}(E) \mapsto \mathcal{P}(E) \times \mathcal{P}(\mathcal{P}(E))$ be a function such that $q(A) = \langle A,W\rangle$, where $W \subseteq \mathcal{P}(A)$, and P be an n-ary predicate over E ($n \geq 1$). For any $\sigma \in P^*$ and any $1 \leq i \leq n$, let $\sigma[i/x]$ be the result of replacing the i-th element of σ by x. Then the simple expansion of P^* by $q(A)$ at position i, written as* $\mathbf{s\text{-}exp}_i(q(A),P^*)$ *is the smallest set Q such that $P^* \subseteq Q$ and clause 1 holds. The complex expansion of P^* by $q(A)$ at position i, written as* $\mathbf{c\text{-}exp}_i(q(A),P^*)$ *is the smallest set Q such that $P^* \subseteq Q$ and clauses 1-4 below hold:*

1. $\sigma^+ \in P^* \wedge [\sigma^+]_i^{P^*,A} \in W \rightarrow \sigma^+[i/q(A)] \in Q$ *(positive $\triangleright$ positive)*
2. $\sigma^+ \in P^* \wedge [\sigma^+]_i^{P^*,A} \notin W \rightarrow \sigma^-[i/q(A)] \in Q$ *(positive $\triangleright$ negative)*
3. $\sigma^- \in P^* \wedge [\sigma^-]_i^{P^*,A} = A \wedge \emptyset \in W \rightarrow \sigma^+[i/q(A)] \in Q$ *(negative $\triangleright$ positive)*
4. $\sigma^- \in P^* \wedge [\sigma^-]_i^{P^*,A} = A \wedge \emptyset \notin W \rightarrow \sigma^-[i/q(A)] \in Q$ *(negative $\triangleright$ negative)*

To illustrate, consider the evaluation of *Every boy (Q_1) tickled exactly three girls (Q_2)* by means of **s-exp** in a model with $B = \{b_1,b_2\}, G = \{g_1,g_2,g_3,g_4\}$ and $T = \{\langle b_1,g_1\rangle,\langle b_1,g_2\rangle,\langle b_1,g_3\rangle, \langle b_2,g_1\rangle,\langle b_2,g_2\rangle,\langle b_2,g_3\rangle,\langle g_1,g_2\rangle,\langle g_1,g_3\rangle,\langle g_1,g_4\rangle,\}$. By the first clause of $\mathbf{s\text{-}exp_2}$ we can add $\langle b_1,[\![Q_2]\!],+\rangle$, $\langle b_2,[\![Q_2]\!],+\rangle$ and $\langle g_1,[\![Q_2]\!],+\rangle$, and by clause 1 of $\mathbf{s\text{-}exp_1}$ we add $\langle [\![Q_1]\!],[\![Q_2]\!],+\rangle$, since $\{b_1,b_2,g_1\} \cap B$ is a witness set of $[\![$*every boy*$]\!]$. The subject-wide-scope reading of $Q_1\ V\ Q_2$ is true in $\mathcal{M}$ iff $\langle [\![Q_1]\!],[\![Q_2]\!],+\rangle \in \mathbf{c\text{-}exp}_1([\![Q_1]\!],\mathbf{c\text{-}exp}_2([\![Q_2]\!],[\![V]\!]^*))$. Clearly, if $\langle [\![Q_1]\!],[\![Q_2]\!],+\rangle \in \mathbf{s\text{-}exp}_1([\![Q_1]\!],\mathbf{s\text{-}exp}_2([\![Q_2]\!],[\![V]\!]^*))$, then $\langle [\![Q_1]\!],[\![Q_2]\!],+\rangle \in \mathbf{c\text{-}exp}_1([\![Q_1]\!],\mathbf{c\text{-}exp}_2([\![Q_2]\!],[\![V]\!]^*))$. Importantly, the truth evaluation of this sentence in a model requires neither inferences based

[2] A quantifier Q satisfies CONS iff for all domains E and for all $A,B \subseteq E$, it holds that $Q_E(A,B)$ iff $Q_E(A,A\cap B)$, and EXT iff iff for all domains E and all $A,B \subseteq E$, it holds that $Q_E(A,B)$ iff $Q_A(A,B)$.

on negative information nor polarity reversal. On the other hand, whether *Less than two boys (Q_1) tickled exactly three girls (Q_2)* can be shown to be true by **s-exp** depends on the situation. If $B = \{b_1, b_2\}, G = \{g_1, g_2, g_3, g_4\}$ and $T = \{\langle b_1, g_1\rangle, \langle b_1, g_2\rangle, \langle b_1, g_3\rangle$, then $\langle b_1, [\![Q_2]\!], +\rangle$ is added by $\mathbf{s\text{-}exp}_2$, and by $\mathbf{s\text{-}exp}_1$ we add $\langle [\![Q_1]\!], [\![Q_2]\!], +\rangle$, since $\{b_1\}$ is a witness set of *less than two boys*. But in the situation where $B = \{b_1, b_2\}, G = \{g_1, g_2, g_3, g_4\}$ and $T = \{\langle g_1, g_2\rangle\}$ **c-exp** is required: By clause 4 of $\mathbf{c\text{-}exp}_2$ we add the negative information that $\langle b_1, [\![Q_2]\!], -\rangle$ and $\langle b_2, [\![Q_2]\!], -\rangle$, and by clause 3 of $\mathbf{c\text{-}exp}_1$ we add the positive information $\langle [\![Q_1]\!], [\![Q_2]\!], +\rangle$ based on the negative information added before (polarity reversal). An important consequence of this theory is the following proposition:

Proposition 1: *Let $q_1(A_1) = \langle A_1, W_1\rangle$ and $q_2(A_2) = \langle A_2, W_2\rangle$ be w-quantifiers with $\emptyset \notin W_1$ and $\emptyset \notin W_2$, and let P^* be the polarity relation of some predicate $P \subseteq E^n$.*[3] *Then $\langle q_1(A_1), q_2(A_2), +\rangle \in \mathbf{s\text{-}exp}_1(q_1(A_1), \mathbf{s\text{-}exp}_2(q_2(A_2), P^*))$ if and only if $\langle q_1(A_1), q_2(A_2), +\rangle \in \mathbf{c\text{-}exp}_1(q_1(A_1), \mathbf{c\text{-}exp}_2(q_2(A_2), P^*))$.*

That is, if a quantified statement contains no empty-set quantifiers, then $\langle q_1(A_1), q_2(A_2), +\rangle$ can be added by **s-exp** if and only if it can be added by **c-exp** – in other words, the evaluation of such a statement's truth requires only inferences of positive information from positive information. If, on the other hand, a quantified statement does contain an empty-set quantifier, then it cannot be evaluated for truth by **s-exp** in every model, as shown above.

Based on our theory we derive the following predictions about the processing of quantified sentences. **Prediction 1:** Given that negation and polarity reversal increase processing difficulty [2], we predict first that quantified statements which have to be evaluated by **c-exp** are more difficult to process than statements which can be evaluated by **s-exp**. Non-empty-set quantifiers (e.g. UE quantifiers like *at least/more than n* or *every*[4]) can be evaluated by **s-exp** in all models/situations. On the other hand, the evaluation of empty-set quantifiers (all DE quantifiers, among others) depends on the model: in some models **s-exp** suffices, in others **c-exp** is necessary. We predict increased processing difficulty whenever **s-exp** does not suffice. In particular, processing difficulty should be increased if a rule has to be applied that involves polarity reversal (clauses 2 and 3) or is based on negative information (clauses 3 and 4). The most difficult cases should be applications of clause 3 because this clause involves polarity reversal and inference from negative information. **Prediction 2:** If the comprehension of a quantified sentence involves the specification of the simplest algorithm for checking this sentence *in any model*, we predict not only differences in ease of verification but also differences in ease of comprehension: sentences involving no empty-set Q should be read faster than sentences involving empty set Q's, because the former requires the specification of an algorithm involving **s-exp**, whereas the latter involves the complex expansion **c-exp**.

3 Experiments

Experiment 1: 72 German participants read simply quantified intransitive sentences of the type *Q dots are blue* manipulating Q: (a) *mindestens ein...* (*at least one*, non-empty-set), (b) *höchstens ein...* (*at most one*, empty-set) and (c) *weniger als zwei* (*less than two*, empty-set). After reading the sentence, a picture was presented that showed either zero objects (0-model), one object (1-model), or three objects (3-model) of the relevant color among other objects of a different color.

[3]Note that this proposition also holds for empty predicates.

[4]Assuming existential import *every* can also be considered a non-empty-set quantifier since it presupposes a non-empty restriction [3]. Note that without this presupposition the empty set is a witness set of *every*(A) in case A is empty.

All objects (total number between 4 to 7 varying across items) in the display were elements of the restrictor set. 27 items were constructed in a 3×3 within design and distributed to nine lists using a latin square. The descriptive statistics are presented in Table 1.

	error rates			judgment times		
	0-model	1-model	3-model	0-model	1-model	3-model
at least one	0%	0%	4.9%	1511ms	1300ms	1453ms
fewer than two	10.3%	2.0%	0%	2076ms	1744ms	1755ms
at most one	22.5%	1.0%	1.0%	1939ms	1385ms	1478ms

Table 1: Mean judgments and judgment times in Exp. 1.

The analysis of error rates revealed that, except for the two empty-set quantifiers with 0-models, all of the conditions had error rates well below 5%. The 0-models led to 10.3% errors for *less than* and to even 22.6% errors for *at most*[5]. In a logit mixed effects model analysis, the observed differences led to a significant *quantifier* $\times$ *model* interaction ($z = 2.42$). The increased error rates for empty-set Q's when the predicate is empty are fully expected since this case requires the application of the most difficult clause of **c-exp**, namely clause 3, which involves polarity reversal and inference from negative information. If the predicate is not empty, then there is an $a \in E$ with $\langle a, + \rangle \in P^*$, so clause 1 or 2 can be applied. The verification times[6] also showed a significant interaction between *quantifier* and *model* (ANOVAs: $p_1 < .01$; $p_2 < .05$) which lends further support to our theory (prediction 1). Thus, the evaluation of 0-models proves to be a source of quantificational complexity in empty-set quantifiers even in simply quantified intransitive sentences.

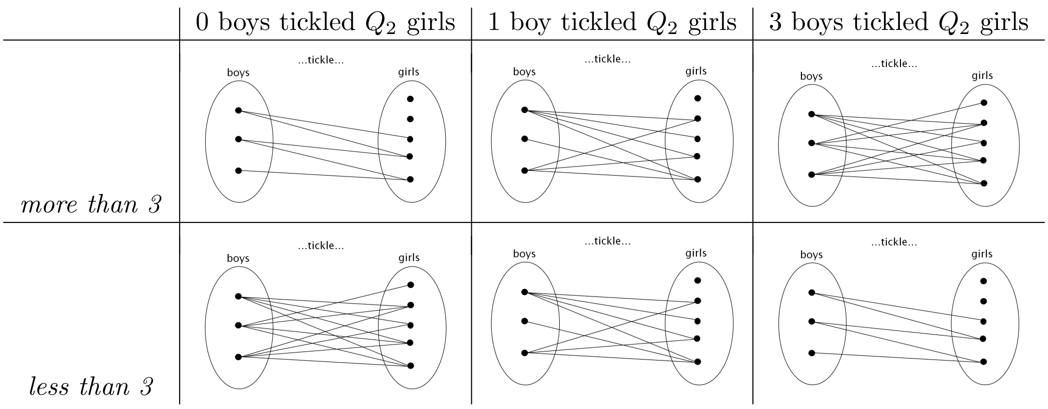
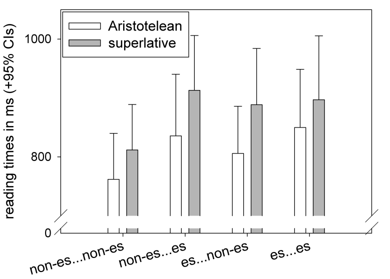


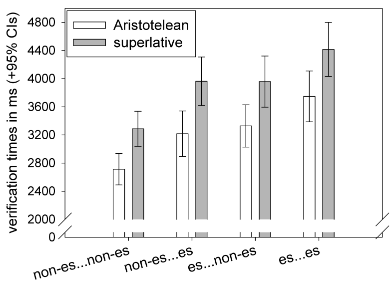
Figure 1: Sample models in Exp. 2 for Q_1 *boys tickled more than/less than three girls*. Note: diagrams corresponding to 0- and 3-models had to be swapped for Q_2=*more than* and Q_2=*fewer than*, but the set of diagrams was kept constant across sentence conditions.

Experiment 2: 72 participants read German sentences of the form Q_1 *boys tickled* Q_2 *girls* which have independently been shown to exhibit surface scope only [cf. 1]. Q_1 was either one of the Aristotelian quantifiers *jeder* (*each*, non-empty-set) and *kein* (*no*, empty-set) or one of the superlative quantifiers *mindestens ein* (*at least one*, non-empty-set) and *höchstens*

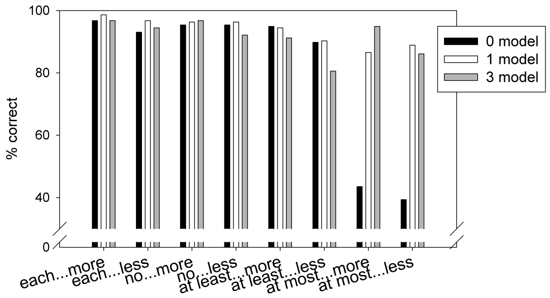
[5] Being well aware of the fact that these "errors" may be (at least partly) caused by pragmatic factors.
[6] All judgment times and reading times in Exp. 1/2 were corrected for outliers by excluding RTs from the analysis that were below 100ms or above a participant's mean RT plus 2.5sd's.



(a) mean reading times of the critical region



(b) mean judgment times



(c) mean judgments

Figure 2: Results of Experiment 2. *es* = empty-set Q; *non-es* = non-empty-set Q.

ein (*at most one*, empty-set). Q_2 was either *mehr als zwei/drei* (*more than two/three*, non-empty-set) or *weniger als zwei/drei* (*less than two/three*, empty-set). (8) is a sample sentence, asterisks indicate segmentation for self-paced reading. The underlined noun phrase of the second quantifier marks the critical region since this is the earliest point at which a verification algorithm can be fully specified.

(8) Mindestens ein * Junge * kitzelte * mehr * als drei * <u>Mädchen</u>.
At least one * boy * tickled * more * than three * <u>girls</u>.

Participants read these sentences in a self-paced reading experiment with moving window presentation. After each sentence, a set diagram appeared on the screen of the types shown in Figure 1 and they had to provide a truth value judgment. As in Exp. 1 there were three types of diagrams: 0-, 1- and 3-set diagrams which showed a) no, b) one or c) three boys tickling Q_2 girls, respectively. Accordingly, the experiment employed a factorial 2 (*empty-set* Q_1 vs. *non-empty-set* Q_1) $\times 2$ (*type of* Q_1: Aristotelian vs. superlative) $\times 2$ (*empty-set* Q_2 vs. *non-empty-set* Q_2)) $\times 3$ (*diagram*) within design. Diagrams were constructed in such a way that across sentence conditions always the same set of pictures was used for the 0- vs 1- vs 3-set diagram. Since 0- and 3-models had to be swapped for empty-set vs. non-empty-set-Q_2s, judgment times had to be collapsed over diagram types to allow for comparison. 72 experimental items plus 78 fillers were constructed and distributed to 24 lists in a latin square. Across the experiment, 50% of the sentences were true. Sentence-picture-pairs were presented in randomized order.

	each$\uparrow$	each$\downarrow$	**at least**$\uparrow$	at least$\downarrow$	no$\uparrow$	no$\downarrow$	at most$\uparrow$	at most$\downarrow$
0	1(1/1/1)	4(2/2/2)	1(1,1,1)	4(2/2/2)	3(2/2/2)	3(2/2/2)	3(2/2/2)	3(2/2/2)
1	1(1/1/1)	2(1/2/2)	1(1,1,1)	1(1/2/2)	2(1/2/2)	2(1/2/2)	1(1/2/2)	1(1/2/2)
3	1(1/1/1)	1(1/1/1)	1(1,1,1)	1(1/1/1)	2(1/1/1)	2(1/1/1)	2(1/1/1)	2(1/1/1)

Table 2: Clauses of **c-exp** that had to be applied in Exp. 2. The arrows indicate whether Q_2 was an empty-set quantifier ($\downarrow$) or not ($\uparrow$). The clauses for **c-exp**$_2$ are given in parentheses after the clause numbers for **c-exp**$_1$. Conditions in boldface indicate quantifier combinations that can be evaluated by **s-exp** in any model. *0* = 0-model, *1* = 1-model and *3* = 3-model.

We analyzed reading times as well as verification latencies by computing repeated measures ANOVAs. Error rates of the subsequent verification stage were analyzed in a logit mixed effects model. The full set of results is presented in Figure 2. Up to the critical noun phrase the only significant effect was the main effect of *empty set* Q_2 ($p_{1/2} < .01$) at regions *fewer/more* and *than n*. Unsurprisingly, *fewer than n* took longer to read than *more than n*. Effects due to the semantic properties of the first quantifier only emerged at the critical noun phrase of the second quantifier phrase. Conditions with two non-empty-set quantifiers were read faster than the conditions involving empty-set quantifiers (ANOVAs: *empty set* Q_1 $\times$ *empty set* Q_2, $p_1 < .05$; $p_2 = .09$) in line with our second prediction that an algorithm based on **s-exp** is inherently less complex than an algorithm based on **c-exp**.

Both empty-set subject and empty-set object quantifiers led to a slowdown of judgment times as reflected by significant main effects of *empty set* Q_1 ($p_{1/2} < .01$) and *empty set* Q_2 ($p_{1/2} < .01$) but no reliable interaction. Note that this is slightly different from what we would expect on the basis of prediction 1. According to our theory, one empty-set quantifier may already suffice to trigger **c-exp**. Table 2 summarizes which clauses had to be applied in each of the 24 conditions. As can be seen from the table, we would have expected an interaction between *witness sets* Q_1 and *empty set* Q_2 because for an empty-set Q_1 the same clauses were required for empty-set and non-empty-set Q_2. To test whether this unexpected difference was reliable we computed ANOVAs for the subset of conditions with an empty-set Q_1. This analysis revealed a main effect of *empty set* Q_2 ($p_{1/2} < 0.1$). Thus, the judgment times provide partial evidence against our theory. As it seems, the combination of two empty-set quantifiers adds processing difficulty beyond what would be expected on the basis of the complexity of the proposed verification algorithms.

Consistent with recent proposals from the semantic literature (e.g., [5]) superlative quantifiers were more difficult to process than Aristotelian quantifiers consistently across the reading and verification stages (main effects of *type of* Q_1, all $p_{1/2} < .01$) suggesting that they are inherently more complex. In fact, complexity due to the *type of* Q_1 and *empty set* Q_1/Q_2 added up. This was indicated by extraordinarily high error rates of sentences with a superlative empty-set Q_1. Accordingly, a logit mixed effects model analysis of the superlative conditions revealed a reliable two-way interaction between *empty set* Q_1 and *diagram type* ($z = 8.98$). As in Exp. 1, errors were limited to those models which required the evaluation of a empty-set quantifier with an empty predicate, further confirming prediction 1. This time, polarity reversal led to performance even below chance level, in particular in sentences with the two empty-set quantifiers *at most* and *less than* ($\chi^2(1) = 5.06$; $p < .05$).

The results of Exp. 2 are for the most part consistent with our quantification theory. The observed pattern of reading times supports the distinction between **s-exp** and **c-exp** and the verification data lend support to the predicted difficulty of empty-set quantifiers when having to evaluate a situation in which the scope of the subject quantifier phrase consists of the empty set. Only the judgment times are not fully compatible with our theory.

4 Conclusions

We have presented an algorithmic theory of quantifier interpretation which predicts that the operation employed in processing empty-set quantifiers is more difficult than that for processing non-empty-set quantifiers. This prediction was mainly confirmed in two experiments that investigated the online comprehension and verification of simply and doubly quantified sentences.

In order to further disentangle complexity effects that are due to the presence of empty-set quantifiers from complexity effects due to the monotonicity of the quantifiers, we plan to extend this line of research and investigate whether non-monotonic Q's which are not empty-set Q's (e.g. *exactly one boy or exactly three boys*) are easier to process than non-monotonic empty-set Q's (e.g. *no boy or exactly three boys*), a prediction which - to our knowledge - is unique to the present account.

References

[1] O. Bott and F. Schlotterbeck. Incremental truth value judgments. In B. Stolterfoht and S. Featherston, editors, *Empirical approaches to linguistic theory*, pages 3–28. Mouton de Gruyter, Berlin, 2012.

[2] H. Clark and W. Chase. On the process of comparing sentences against pictures. *Cognitive Psychology*, 3:472–517, 1972.

[3] B. Geurts. Existential import. In I. Comorovski and K. von Heusinger, editors, *Existence: Semantics and syntax*, pages 253–271, Dordrecht, 2007. Springer.

[4] B. Geurts, N. Katsos, J. Moons, and L. Noordman. Scalar quantifiers: logic, acquisition, and processing. *Language and Cognitive Processes*, 2010.

[5] B. Geurts and R. Nouwen. "At least" et al.: the semantics of scalar modifiers. *Language*, 83(3):533–559, 2007.

[6] B. Geurts and F. van der Slik. Monotonicity and processing load. *Journal of Semantics*, 22:97–117, 2005.

[7] J. Koster-Moeller, J. Varvoutis, and M. Hackl. Verification procedures for modified numeral quantifiers. In N. Abner and J. Bishop, editors, *Proceedings of the 27th West Coast Conference on Formal Linguistics*, Somerville, MA, 2008. Cascadilla Proceedings Project.

[8] C. T. Mcmillan, R. Clark, P. Moore, C. Devita, and M. Grossman. Neural basis for generalized quantifiers comprehension. *Neuropsychologia*, 43:1729–1737, 2005.

[9] Y. Moschovakis. Sense and denotation as algorithm and value. In J. Oikkonen and J. Väänänen, editors, *Lecture Notes in Logic, vol. 2*, pages 210–249. Association for Symbolic Logic, A.K. Peters, Ltd, Natick, MA, 1994.

[10] S. Steinert-Threlkeld and T. F. I. III. Iterating semantic automata. *Linguistics and Philosophy*, 36:151–173, 2013.

[11] J. Szymanik and M. Zajenkowski. Comprehension of simple quantifiers: Empirical evaluation of a computational model. *Cognitive Science*, 34(3):521–532, 2010.

[12] J. van Benthem. *Essays in Logical Semantics.* Reidel, 1986.