SUPERLATIVE QUANTIFIERS AS SPEECH ACT MODIFIERS: EVIDENCE FROM HEBREW[1]

AVIYA HACOHEN
DANA KOZLOWSKI
ARIEL COHEN
Ben-Gurion University of the Negev

1 Background: The Meaning of Superlative Quantifiers

The superlative quantifiers (henceforth SQs) *at least* and *at most* are commonly assumed to have the same truth-conditions as comparative quantifiers (Keenan & Stavi 1986). Thus, (1a) and (2a) are taken to be equivalent to (1b) and (2b), respectively.

(1) a. Mary petted at least three rabbits
    b. Mary petted more than two rabbits
(2) a. Mary petted at most three rabbits
    b. Mary petted fewer than four rabbits

However, Geurts & Nouwen (2007) have convincingly shown that this assumption is wrong. Instead, they propose that SQs are complex epistemic operators. Specifically, (1a) is claimed to mean that it is epistemically necessary that Mary petted three rabbits, and it is epistemically possible that she petted more. Formally:

(3) $\exists x(\textbf{RABBITS}(x) \wedge |x|=3 \wedge \textbf{pet}(\textbf{m},x)) \wedge \Diamond\exists x(\textbf{RABBITS}(x) \wedge |x|>3 \wedge \textbf{pet}(\textbf{m},x))$

Alternative theories are proposed by Büring (2007) and Cummins & Katsos (2010), according to whom SQs are disjunctions.[2] Specifically, (1a) means that Mary petted exactly three rabbits, or

[1]This study is supported by the Israeli Science Foundation, grant # 376/09. We wish to thank Yifat Faran, Yarden Kedar,Chris Kennedy Desiree Meloul, David Mesika, and an anonymous referee for their helpful suggestions.
[2]These are two distinct theories, but for our purposes we can treat them as one.

she petted more than three rabbits. Formally:

(4) $|\mathbf{RABBITS} \cap \lambda x.\mathbf{pet}(\mathbf{m},x)|=3 \vee |\mathbf{RABBITS} \cap \lambda x.\mathbf{pet}(\mathbf{m},x)|>3$

Cohen & Krifka (2010) take a different approach, according to which SQs are illocutionary operators. Their starting point is the notion of *Speech Act Denegation* (Searle 1969; Hare 1970). Note the difference between (4a) and (4b):

(5) a. I promise not to come.
b. I don't promise to come.

Sentence (4a) is a straightforward promise: the speaker expresses an obligation to not come, i.e. to make true the negation of the proposition that the speaker comes. Assuming an operator PROMISE, this can be formalized as:

(6) PROMISE(¬'I come')

In contrast, (4b) is not really a promise: the speaker is explicitly refraining from performing the speech act. This is the denegation of the speech act of PROMISE, formalized by the symbol '~', so that (4b) is represented as:

(7) ~PROMISE('I come')

1.1 The Speech Act of GRANTing

Sometimes, rather than assert a proposition, we grant it to the interlocutor (cf. Merin 1994). GRANTs are denegations of asserting the contrary: to GRANT Φ is to refrain from ASSERTing $\neg\Phi$. In fact, the following equivalences (similar to modal logic) can be proved:

(8) a. $\mathrm{GRANT}(\Phi) = {\sim}\mathrm{ASSERT}(\neg\Phi)$
b. $\mathrm{ASSERT}(\Phi) = {\sim}\mathrm{GRANT}(\neg\Phi)$

Cohen & Krifka propose that SQs express Quantification over GRANTs. Specifically, (1a) means that the minimal number n such that the speaker GRANTs that Mary petted n rabbits is n = 3:

(9) $\min\{n|[\mathrm{GRANT}(|\mathbf{rabbit} \cap \lambda x[\mathbf{pet}(\mathbf{m},x)]|=n)\} = 3$

This statement can receive the formulation in (9a), which is equivalent to (9b).

(10) a. $\forall n[n<3 \rightarrow {\sim}\mathrm{GRANT}(|\mathbf{rabbit} \cap \lambda x[\mathbf{pet}(\mathbf{m},x)]| = n]$
b. $\forall n[n<3 \rightarrow \mathrm{ASSERT}(|\mathrm{rabbit} \cap \lambda x[\mathrm{pet}(m,x)]| \neq n]$

The formulations in (9) say that, for all n<3, the speaker asserts that Mary did *not* pet exactly n rabbits. Note that they do not say anything about $n \geq 3$: for example, it is quite compatible with

(9) that the speaker also denies that Mary petted exactly 3, 4, 5, and 6 rabbits. However, this interpretation would violate the *implicature* generated by (9): since the speaker bothered to indicate that she denies that Mary petted 0, 1, or 2 rabbits, if the speaker intended to deny this for higher values of n too, she should have indicated this. Hence, by implicature, but only by implicature, we can conclude that the speaker GRANTs that Mary petted n rabbits for $n \geq 3$:

(11) a. $\forall n[n \geq 3 \rightarrow \text{GRANT}(|\textbf{rabbit} \cap \lambda x[\textbf{pet}(\textbf{m},x)]| = n]$
b. $\forall n[n \geq 3 \rightarrow \sim\text{ASSERT}(|\textbf{rabbit} \cap \lambda x[\textbf{pet}(\textbf{m},x)]| \neq n]$

It therefore follows that the speaker who utters (1a) performs the following speech acts:

(12) $\text{ASSERT}(..\neq 0..) \wedge \text{ASSERT}(..\neq 1..) \wedge \text{ASSERT}(..\neq 2..) \wedge$
$\sim\text{ASSERT}(..\neq 3..) \wedge \sim\text{ASSERT}(..\neq 4..) \wedge \ldots$

How are truth conditions generated from this formulation? Suppose first that Mary petted exactly two rabbits. Then (1a) ought to be false. Indeed, uttering (1a) includes, among other speech acts, $\text{ASSERT}(|\textbf{rabbit} \cap \lambda x[\textbf{pet}(\textbf{m},x)]| \neq 2)$. But this is an assertion of a false proposition, and, consequently, (1a) is false, as desired.

A more interesting case is one where Mary petted exactly four rabbits; now (1a) ought to be true. Uttering (1a) includes the following speech acts $\text{ASSERT}(..\neq 0..)$, $\text{ASSERT}(..\neq 1..)$, $\text{ASSERT}(..\neq 2..)$; all these are assertions of true propositions. However, this, in itself, does not yet account for the truth of (1a); we need to know that the speaker does not assert any false proposition, and, in particular, does not $\text{ASSERT}(..\neq 4..)$. We *can*, in fact, conclude this, by implicature—and only by implicature. Hence, all the assertions are true, and no false assertions are made, which accounts for the truth of (1a) in this case.

Note that a rather interesting situation arises: the falsity of SQ sentences is determined semantically, whereas their truth is determined pragmatically, via implicature. In fact, this captures the intuition that when one says (1a), one doesn't know how many rabbits Mary petted, but one does know how many rabbits she did *not* pet. Crucially, this implicature does not strengthen already determined truth conditions; rather, the implicature is necessary in order to have truth conditions in the first place. Without the contribution of implicature, we know that the speaker denies that Mary petted 0, 1, or 2 rabbits, but we do *not* know whether the speaker also denies that Mary petted 3, 4, and 5… rabbits.

There is some linguistic evidence for the involvement of implicature in the interpretation of SQs, which we will present briefly (see Cohen & Krifka 2010 for the full details). One piece of evidence involves cancelability: since it is proposed that (1a) implicates that the speaker GRANTs that Mary petted 3, 4, 5… rabbits, we expect this implicature to be cancelable, which indeed it is:

(13) Mary petted at least three rabbits, in fact five.

It might be argued that this is not really cancelation, but merely further specification of the range of numbers of rabbits petted by Mary. However, this is not the case, since cancelation becomes harder the more GRANTs are canceled:

(14) #Mary petted at least three rabbits, in fact 1000.

Note that no such contrast is exhibited by comparative quantifiers, which straightforwardly allow range specification:

(15) Mary petted more than two rabbits, in fact five/1000.

Another piece of linguistic evidence for the involvement of implicature regards embedding. It is well known that scalar implicatures generally do not survive downward entailing contexts (Chierchia 2004). For example, (15a) implicates that you will drink or smoke, but not both. But this implicature disappears in the antecedent of a conditional, and (15b) means that if you drink or smoke *or both,* you will become ill.

(16) a. You will drink or smoke
b. If you drink or smoke, you will become ill

Since the interpretability of SQs requires scalar implicature, Cohen & Krifka predict them to be bad in downward entailing contexts. This prediction is borne out, as can be seen by the following example, from Geurts &Nouwen (2007):

(17) a. *None of the guests danced with at least/most three of the waitresses.
b. ?Betty didn't have at least/most three martinis.[3]

The following examples, from Nilsen (2007), make the point even more clearly:

(18) a. ??John hardly ate at least/most three apples.
b. ??Policemen rarely carry at least/most two guns.
c. ??This won't take at least /most 50 minutes.

In this paper, however, we will concentrate not on the linguistic evidence, but on experimental evidence for the crucial role that implicature plays in the interpretation of SQs.

2 Testing the Theories: Predictions for Processing

Above we briefly described three theories of SQs:

(i) Geurts & Nouwen (2007)
(ii) Büring (2007) and Cummins & Katsos (2010)
(iii) Cohen & Krifka (2010)

According to all three theories, SQs are more complex than comparative quantifiers; they are epistemic operators, disjunctions, or quantifiers over speech acts. Hence, all three theories make a common prediction regarding the processing of SQs: that it will take longer than the processing

[3]Geurts and Nouwen propose a different explanation for these examples. See Cohen and Krifka (2010) for problems with this account.

of comparative quantifiers. This prediction has been borne out, as reported by Geurts, Katsos, Cummins, Moons & Noordman (2010) and Cummins & Katsos (2010).

However, Cohen & Krifka (2010) make an additional prediction. Recall that according to this theory, judgments of true SQ sentences require computation of scalar implicature, whereas judgments of false SQ sentences do not. Since computing scalar implicatures takes time (e.g., Bott & Noveck 2004), the theory predicts that true SQ sentences will take longer to process than false ones. In contrast, the competing theories do not draw any distinction between true and false SQ sentences, and therefore make no such prediction. In the following, we describe the experiments we conducted to test these predictions.

3 Methods

3.1 Experiment 1

Before presenting the experimental materials, we must first address the issue of frequency effects. We know that in English *at least* is much more frequent than *at most*[4]; in order to control for these potential frequency effects, we chose Hebrew as the language of the stimuli. Hebrew has two forms (*lexol hapaxot* 'at least' and *lexol hayoter* 'at most') with roughly the same (low) frequency; for completeness, we also added the much more frequent form *lefaxot* 'at least'. This way, we controlled for any potential confounds that may result from frequency differences.

3.1.1 Materials and Design

The experiment tested two classes of quantifier: Superlative Quantifiers and Comparative Quantifiers. The former class consisted of three levels and the latter included two levels; together they rendered the following five experimental conditions:

Table 1: Experimental conditions

Superlative Quantifier	Comparative Quantifier
lexol hapaxot 'at least'	*yoter me-* 'more than'
lexol hayoter 'at most'	*paxot me-* 'less than'
lefaxot 'at least'	

Participants were presented with Hebrew sentences of the structure *I see Q N Xs*, where Q is a superlative or comparative quantifier; N is a number between 3-5; and X is an everyday object. Each sentence was accompanied by picture and the participant had to judge whether the sentence truthfully described the accompanying picture. Example items from the three SQ conditions are presented below:

[4] A simple Google search revealed a 26:1 ratio between *at least* and *at most* (approximately 1,740,000,000for *at least* and circa 67,700,000 for *at most*)

Figure 1: sample experimental items

The trials consisted of five types of sentence, as seen in Table 1, each of which appearing 36 times, comprising a total of 180 sentences per participant.

Two issues regarding the experimental material ought to be noted. First, our pilot study revealed that when the same number of objects appears in both verbal and visual stimulus (e.g., the sentence is *I see at least three glasses*, and the picture contains exactly three glasses), participants are often confused; therefore, in order to avoid this potential confusion, the visual stimulus (the picture) always depicted either one more or one fewer objects than the number of objects indicated in the verbal stimulus.

Second, it is important to note that the objects in each picture were all of the same kind, i.e., all plates or all glasses but never gasses and plates together. This was done in order to control for a different reading of SQs, discussed by Kadmon (1987). According to this interpretation, a sentence such as *I see at least three plates* can mean *I see exactly three glasses, and possibly some other kind of dishware.*

3.1.2 Procedure

We used a sentence verification task, in which participants were asked to judge whether a written sentence truthfully described an accompanying picture. The experiment was conducted in Hebrew and began with twenty practice items, each of which included feedback, in order to ensure that the participant understood the task. Trial and experimental procedures were identical except for the absence of feedback in the experimental phase and the fact that participants were encouraged to ask questions during the practice but not during the actual experiment. Each experimental session was divided into three blocks in order to allow participants some time to pause.

Each trial began with a presentation of a fixation point, followed by the appearance of an experimental item, such as the ones presented in (18). The item remained on the screen until the participant responded. Trials were counterbalanced across participants and the following categories were counterbalanced across the experiment: Condition, response-type (T/F), number of items in visual stimulus, number of items in verbal stimulus, object-type in visual stimulus.

Two lists were created such in order to counterbalance the response buttons: in one of the lists, the *p* button was assigned the True response and in the second list the True response was indicated by the *q*. This was very important for the current study, as it controlled for the possibility that any potential processing differences between true versus false SQs could be merely the consequence of the "true button" being on the right or on the left.

The experiment was programmed, the stimuli were presented and the reaction time data were recorded using the E-Prime software (Psychology Software Tools Inc., Pittsburgh, Pennsylvania, USA).

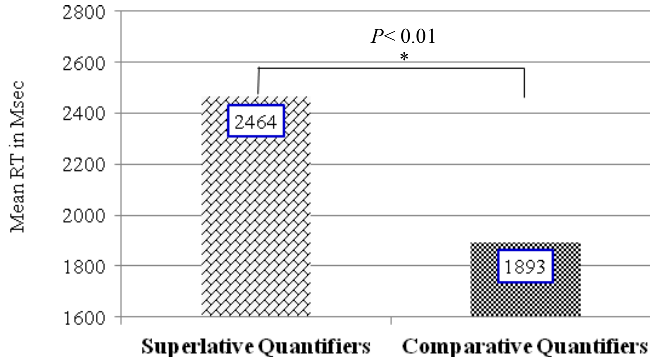
3.1.3 Participants

Twenty eight native monolingual speakers of Hebrew, most of them students at Ben-Gurion University, volunteered to participate in this experiment. There were 17 female and 11 male participants.

3.1.4 Results and Discussion

Mean reaction times were calculated for every trial and each participant. In our report of the data we excluded incorrect responses as well as extremely long or extremely short reaction times, which indicate that no relevant processing actually took place.

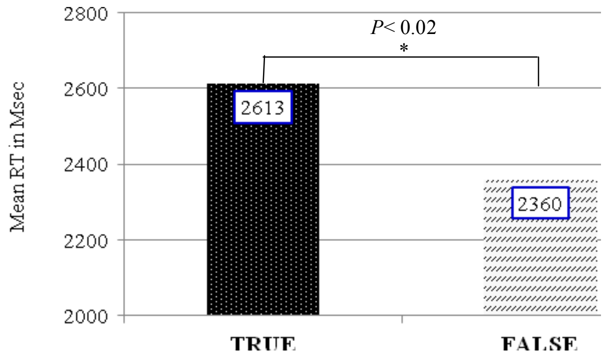
Figure 2: RTs for superlative vs. comparative quantifiers



Not surprisingly, and as can be seen from the graph above, prediction 1 was borne out: SQs took on average nearly 2500 msec to process, while the mean reaction time for comparative quantifiers was less than 1900 msec. An unpaired t-test of performance revealed that this difference between superlative and comparative quantifiers was significant ($P < 0.01$). We have thus been able to replicate the results of Geurts *et al.* (2010) and Katsos & Cummins (2010), regarding the difference in processing time between comparative and superlative quantifiers.

But what about prediction 2, which provides a crucial test of Cohen & Krifka's theory, and constitutes the main prediction of the current study? The figure below presents the results of true versus false SQs:

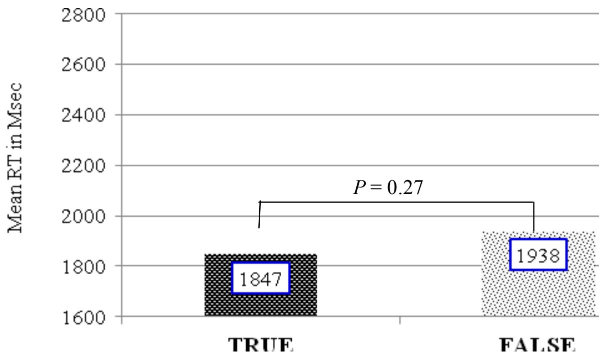
Figure 3: RTs for true vs. false superlative quantifiers



The data above clearly demonstrate that prediction 2 was also borne out: whereas true SQs took on average more than 2600 msec to process, false SQs were processed at 2360 msec, and this difference is statistically significant ($P<0.02$). Importantly, an ANOVA revealed that the interaction between quantifier-type (*lexol hapaxot*, *lexol hayoter* or *lefaxot*) and truth judgment (true or false) was not significant ($P=0.15$), indicating that all three SQs demonstrate a similar effect. Hence, our findings corroborate Cohen & Krifka's theory.

Yet, we wanted to ensure that the RT differences truly indicates the relevant processing differences predicted by Cohen & Krifka, rather than simply stemming from a potential artifact of the experimental design. We therefore also analyzed participants' responses in the comparative quantifier conditions. As can be seen below, result here revealed no RT difference between True and False responses:

Figure 4: RTs for true vs. false comparative quantifiers



True comparatives were processed on average at approximately 1850 msec and he mean RT for false comparatives was at around 1940 msec. The unpaired t-test revealed that this difference was not significant, as indicated by the relatively high *P* value ($P = 0.27$).Thus, this analysis verifies that the differences found in the SQ conditions were real and not simply the result of experimental design.

Taken together, our data clearly support Cohen & Krifka (2010), who predict that true SQs will require longer processing time than false ones. Recall that no other theory predicts this result. Nevertheless, could our findings somehow be made compatible with any competing theories? Before addressing this question, it is important to note here that we are not arguing that any of the competing theories presented above makes such a claim, either explicitly or implicitly; we are merely entertaining the possibility that an alternative theory could also account for our data.

The first alternative account was suggested to us by an anonymous reviewer. According to this reviewer, when *at least* ϕ is true, the participant must consider more possibilities than when *at least* ϕ is false. Consider, for example, the sentence *Mary petted at least three rabbits*. There are three possibilities for it to be false, namely if Mary petted exactly two rabbits, one, or no rabbit at all. In contrast, there are an infinite number of cases which will make the sentence true (i.e., Mary's petting three rabbits, four rabbits, five rabbits…). It could be—this reviewer argues—that this, rather than the difference between semantic versus pragmatic processing, is what is responsible for the response pattern attested.

However, if we follow this line of argumentation, then we must also assume that when *at most* ϕ is true, there are fewer scenarios to consider than when it is false. Yet, our experiment shows that *at most* ϕ takes longer to process when it is true than when it is false, just like *at least* ϕ. Hence, this hypothesis could not account for our data.

A second alternative was proposed by Chris Kennedy (p.c.). As Kennedy (rightfully) points out, *at least* ϕ is odd when all facts are known: if we know exactly how many rabbits Mary petted, *Mary petted at least three rabbits* is odd. In order to get to the correct judgment, one must suppress the fact that the sentence is infelicitous. In other words, there is a sort of mismatch between truth and felicity if an SQ is used in the context of known facts. Perhaps, Kennedy argues, the effort involved in inhibiting the 'infelicitous' response in favor of judging the sentence to be true is what results in longer RTs for true SQs.

As we consider this possibility, let us compare SQs with other constructions that are also odd when all the relevant facts are known:

(i) ϕ *or* ψ
(ii) *possibly* ϕ

Indeed, when speakers are asked to judge these cases in an experimental setting, correct responses amount to no more than 75% (Braine & Rumain 1981; Noveck 2001). Given the similarity in this respect between SQs, (i) and (ii), we would expect SQs to be as difficult as the latter two. Consequently, accuracy rates for SQs should be similar to those found for (i) and (ii); yet, our data reveal that correct responses to *lefaxot* 'at least', which has comparable frequency to that of (i) and (ii), reached 95%! Thus, based on these extremely high accuracy rates, it would be implausible to argue that the infelicity of true SQs can explain why they take longer to process than false SQs.

The third alternative is the result of the authors' attempt to play, so to speak, the devil's advocate; it is the following: arguably, there are logical forms that may take longer to evaluate for truth than for falsity. For example, a false conjunction takes fewer steps to evaluate than a true conjunction. Perhaps a competing theory could argue that the logical form it proposes is of this type; our findings could then be made compatible with such a theory. To control for this possibility, we carried out a second experiment.

3.2. Experiment 2

3.2.1 Materials, Design and Procedure

The materials, design and procedure of experiment 2 were the same as in experiment 1, with one exception: in each trial in experiment 2 the verbal stimulus preceded the visual stimulus by 2 seconds. In other words, participants first saw the written sentence for two seconds, and only then did they see the picture.

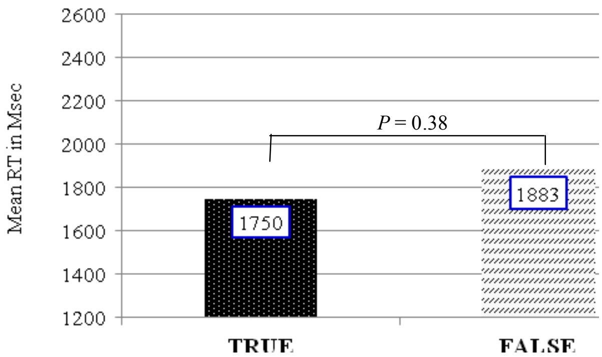
The reason this delay was introduced is that it should allow subjects enough time to compute the required implicature prior to making the actual evaluation (cf. Bott & Noveck 2004). If indeed such an implicature is necessary for the interpretation of true SQs, as Cohen & Krifka hypothesize, then the addition of extra time, allowing the participant to compute the implicature prior to the evaluation, should result in the disappearance of the discrepancies between true and false SQs found in experiment 1.

In contrast, a logical form can only be evaluated once the picture is seen. It follows, then, that if the alternative is correct, the delay we introduced in this experiment should not affect reaction times; the response pattern observed in experiment 1 will also be evinced in experiment 2.

3.2.2 Results and discussion

As the figure below clearly demonstrates, the addition of the delay yielded the response pattern predicted by Cohen & Krifka:

Figure 5: RTs for true vs. false superlative quantifiers (experiment 2)



Following the introduction of the delay, the mean reaction time for true SQs was 1750 msec and the processing of false SQs took on average approximately 1880 msec. This difference was not significant, as indicated by the high P value ($P = 0.38$). These data indicate that indeed the delay allows subjects to compute the implicature necessary for the interpretation of true SQs prior to the actual evaluation, resulting in the disappearance of the RT difference found in experiment 1, as predicted by Cohen & Krifka. If the results of experiment 1 were not due to implicature but to verification of logical form, the delay should not have made a difference. Hence, the results of experiment 2 confirm involvement of the implicature in the processing of true SQs, providing further support for Cohen & Krifka's hypothesis.

4 General Implications

As we saw above, the results of our study clearly support Cohen and Krifka's (2010) theory which argues that the meaning of SQs involves quantification over speech acts. If this theory is indeed correct, it has important implications concerning the nature of speech acts.

The classical view of speech acts (Stenius 1967) considers speech acts to be something quite different from semantic objects like entities or propositions; they are moves in a language game. In other words, speech acts are actions, not propositions.

This perspective on speech acts has clear consequences concerning the role of semantic representations in syntactic recursion. According to the common view, speech acts are distinct from regular semantic objects, and therefore cannot be arguments of semantic operators. Hence, speech acts cannot be embedded. In fact, often the fact that some expressions can be embedded is used as evidence that it is not interpreted as a speech act. In particular, quantification over speech acts, of the form proposed by Cohen and Krifka, ought not to be permitted.

However, the conclusion that embedding of speech acts is impossible has been challenged. Krifka (2001; to appear) formulates a theory of speech acts, according to which, although they differ from regular semantic objects, they can still be "folded back" into semantic meanings in some contexts. Krifka argues that in these contexts, speech acts can, in fact, be embedded.

Cohen and Krifka extend this theory of speech acts to SQs. As we have seen, SQs do not denote propositions, but can be "folded back" into propositions: truth conditions can be derived from the interpretation of SQs by way of implicature. And, indeed, as predicted, SQs can be embedded in some contexts, but not others: they can be embedded in contexts where implicatures survive, but not, as we have seen, in contexts where implicatures do not survive—the scope of downward entailing operators. Therefore, to the extent that the findings in this study support Cohen & Krifka's account of SQs, they provide evidence that speech acts, while they are not propositions, can nonetheless be embedded and quantified over—they are full-fledged participants in the semantic game.

References

Bott, Lewis and Ira A. Noveck. 2004. Some utterances are underinformative: The onset and time course of scalar inferences. *Journal of Memory and Language* 51(3): 437–457.

Braine, Martin D. S. and Barbara Roumain.1981. Children's comprehension of "or": Evidence for a sequence of competencies. *Journal of Experimental Child Psychology* 31:46–70

Büring, Daniel. 2007. The least "at least" can do. In *Proceedings of the West Coast Conference on Formal Linguistics*26114–20,Somerville, MA. Cascadilla Press.

Chierchia, Gennaro. 2004. Scalar implicatures, polarity phenomena and the syntax/pragmatics interface. In *Structures and Beyond*, ed. Adriana Belletti, 39–103. Oxford: Oxford University Press.

Cohen, Ariel and Manfred Krifka. 2010. Superlative Quantifiers as Modifiers of Meta-Speech Acts. In *The Baltic International Yearbook of Cognition, Logic, and Communication, vol 6: Formal Semantics and Pragmatics. Discourse, Context, and Models*, ed. Barbara H. Partee, Michael Glanzberg and Jurģis Šķilters. Manhattan, KS: New Prairie Press.

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

Geurts, Bart, Napoleon Katsos, Chris Cummins, Jonas Moons and Leo Noordman. 2010. Scalar quantifiers: Logic, acquisition, and processing. *Language and Cognitive Processes* 5:244–253.

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

Hare, Richard. 1970. Meaning and speech acts. *The Philosophical Review* 79:3–24.

Kadmon, Nirit. 1987. On unique and non-unique reference and asymmetric quantification. Doctoral dissertation, University of Massachusetts, Amherst.

Keenan, Edward, and Jonathan Stavi. 1986. A semantic characterization of natural language determiners. *Linguistics and Philosophy* 9:253–326.

Merin, Arthur. 1994. Algebra of elementary social acts. In *Foundations of Speech Act Theory: Philosophical and Linguistic Perspectives*, ed. Savas L. Tsohatzidis, 234–266. London: Routledge.

Nilsen, Aystein. 2007. At least: free choice and lowest utility. Talk given at the *ESSLLI Workshop On Quantifier Modification*, Dublin. http://www.let.uu.nl/Rick.Nouwen/qm/nilsen.pdf.

Noveck, Ira A. (2001). When children are more logical than adults: experimental investigations of scalar implicature. *Cognition* 78:165–188.

Searle, John R. 1969. *Speech Acts: An Essay in the Philosophy of Language*. Cambridge: Cambridge University Press.

Stenius, Erik 1967. Mood and language game. *Synthese* 17:254–274.