Exp1

Fabian Schlotterbeck

March 7, 2011

1 Introduction

The present experiment was a picture verification experiment. It served two purposes. Firstly, we wanted to see whether the findings of Chemla & Spector (to appear) could be replicated using a different method. In particular, we collected truth value judgments about set diagrams instead of acceptability ratings and also included distractor items in order to keep participants from developing strategies. Secondly, we tested another type of linguistic construction which has not yet been investigated experimentally. These were constructions like (1) containing *every* embeded under the non-monotone operator *exactly one*. The scalar item *every* is the utmost item in its scale. There has been a debate about whether constructions like these should give rise to global scalar implicatures like in (2) (see).

(1) Exactly one student was parsised by all the teachers.

(2) $\rightsquigarrow$ It is not true that exactly one student was praised by some teachers.

1.1 Procedure

The present experiment was part of a dual task experiment employing picture verification and a sternberg task (Sternberg 1966). In addition to the picture verification task participants had to solve a memory task in some trials. In the *high load* condition participants had to remember six letters while performing the picture verification. In the *low load* condition they had to remember one letter and in the *no load* condition they didn't have to remember anything while they were performing the picture verification. In the *memory only* conditions they only had to perform a memory task (ie. remember one or six letters) but no picture verification. All the critical items of the present experiment were presented in the *no load* condition. Only distractor items were presented in one of the dual task conditions.

In each trial participants were first asked to indicate readiness by pressing a button. In the *no load* condition, after they had pressed the button an asterisk was flashed on the screen six times within $7.2s$. Then a sentence and a picture were presented simultaneously. Within $12s$ participants had to judge whether the sentence was true with respect to the picture. They indicated their judgment by pressing one of two buttons. After they had provided their judgment the next trial started automatically. If they did not respond within twelve seconds, the next trial was started. Judgments and reaction times were logged.

In the *low load* and *high load* condition a trial looked similar to a trial in the *no load* condition except for two things. Firstly, in the dual task conditions instead of asterisks letters were flashed on the screen. Secondly, after finishing the picture verification task a question appeared on the screen probing for a letter. Participants had to indicate by pressing a button whether they thought the letter had been present in the array from the beginning of the trial.

Except for one thing, the *memory only* trials were identical to the dual task trials. In *memory only* trials the picture verification part was replaced with a blank screen. That is, in the middle of such trials the screen stayed blank for $5s$.

The experiment started with a practise session consisting of 25 trials devided into three blocks (1st block: memory task only, ten trials; 2nd block: picture verification only, five trials; 3rd block: dual task, ten trials). Then one list consisting of 249 trials devided into three blocks was presented. An experimental session took about 75 minutes. During the whole experiment no feedback was provided to the participants.

1.2 Materials

There were three sentence conditions. An example of each sentence condition is given in (3)−(5). We will refer to the sentence type in (3) as the $\forall\exists$ type of sentence. Such sentences contained the determiner *einige* (some) embedded under *jeder* (every). Sentences like the one in (4) contained *einige* embedded under *genau ein* (exactly one), they will be referred to as $\exists!\exists$ sentences. Finally, sentences like in (5) containig *jeder* embedded under *genau ein* will be referred to as $\exists!\forall$. Within every one sentence the two quantificational determiners were separated by a clause boundary as to enforce linear scope readings. We constructed 48 experimental items, so 48 sentences of each type were constructed yielding a total of 144 sentences.

Because of the logical relations between the different readings of the sentences in (3)−(5), it was not possible to construct all types of disambiguating diagrams one might wish to test. However, testing some of the possible sentence-picture combinations can already provide crucial information about the status of global and local implicatures of embedded scalar items. Below we will introduce the picture materials for each sentence condition in turn. Every sentence was paired with diagrams of four types. In total, this yielded twelve conditions and 576 pictures. The conditions were distributed over twelve lists accordding to a latin square design ensuring that in each list every item was only present in one condition. In addition to the target sentences, 201 (?) distractors were included in each list. Of these, n were *memory only* trials. One third (?) of them was presented in the *no load* condition, another third (?) in the *low load* and n in the *high load* condition. The picture verification of n of the distractor items was intended to yield a "yes true" response, n were false.

(3) Für jeden dieser Umweltschützer gilt: er boykottierte einige dieser
For each of these environmentalists holds: he boycotted some of these
Großkonzerne.
conglomerates.
Each of these environmentalists boycotted some of these conglomerates.

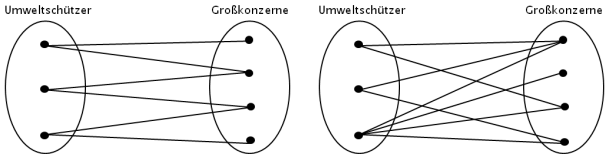
(4) Für genau einen dieser Umweltschützer gilt: er bykottierte einige
For exactly one of these environmentalists holds: he boycotted some

dieser Großkonzerne.
of these conglomerates.
Exactly one of these environmentalists boycotted some of these conglomerates.

(5) Für genau einen dieser Umweltschützer gilt: er bykottierte jeden
For exactly one of these environmentalists holds: he boycotted each
dieser Großkonzerne.
of these conglomerates.
Exactly one of these environmentalists boycotted eachl of these conglomerates.

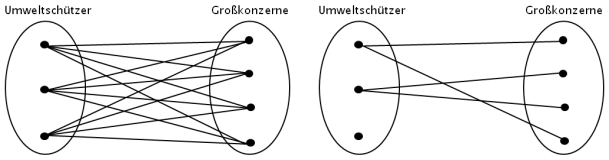
1.2.1 Diagrams for ∀∃ sentences

The ∀∃ sentences were paired with diagrams like those in Figure 1. Figure 1(a) is compatible with the global implicature, the local implicature and the literal meaning of (3). Using the terminology of Chemla & Spector we will refer to such diagrams as *strong* diagrams. In contrast to *strong* diagrams, Figure 1(b) is not compatible with the local implicature (*weak* diagrams) and Figure 1(c) is only compatible with the literal meaning of the sentence (*literal* diagrams). Finally, Figure 1(d) is not compatible with the sentence at all (*false* diagrams). These four conditions constitute a replication of Chemla & Spector (to appear, Experiment 1).



(a) *strong*: compatible with global, local and literal interpretation

(b) *weak*: compatible with global, literal but not with local interpretation



(c) *literal*: Only compatible with the literal meaning

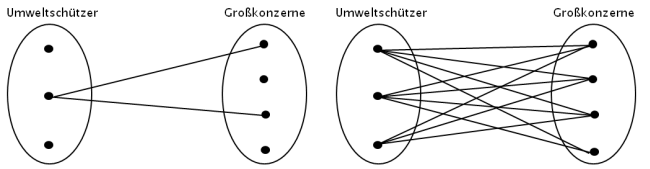
(d) *false*: Not compatible with the sentence at all

Figure 1: Diagrams for ∀∃ sentences

1.2.2 Diagrams for ∃!∃ sentences

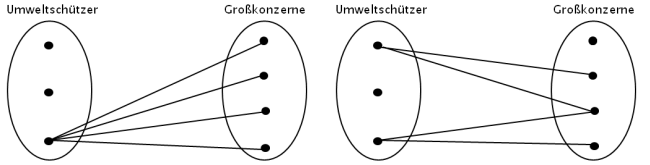
The ∃!∃ sentences were paired with digrams like those in Figure 2. The diagram in Figure 2(a) is compatible with all three interpretations. We use Chemla & Spector's (to appear) term *all* to refer to such diagrams. Figure2(b) is only compatible with the local interpretation (*local* diagrams) and Figure 2(c) is

only compatible with the literal interpretation (*literal* diagrams). Figure 2(d) is plainly false with respect to the sentence in (4) (*false* diagrams). These four conditions replicate Chemla & Spector (to appear, Experiment 2).



(a) *all*: compatible with global, local and literal interpretation

(b) *local*: compatible with local interpretation only



(c) *literal*: compatible with literal interpretation only

(d) *all*: not compatible at all

Figure 2: Diagrams for ∃!∃ sentences

1.2.3 Diagrams for ∃!∀ sentences

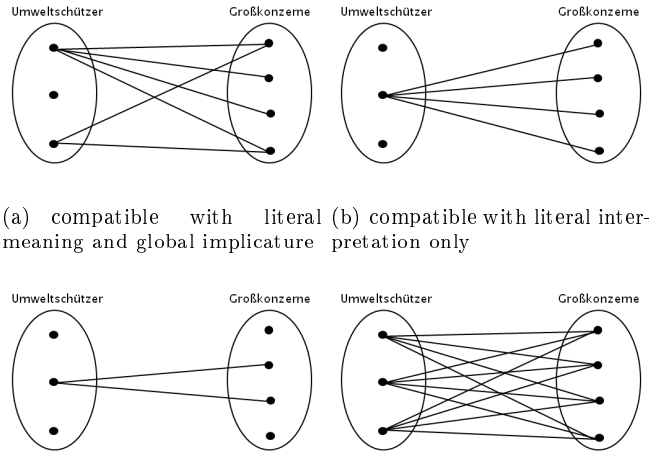
The ∃!∀ sentences were paired with digrams like those in Figure 3. Figure 3(a) is only true under the literal reading (*literal* diagrams) while Figure 3(b) is only true with respect to the global implicature (*global* diagrams), denying the weaker scalar alternative containig *einige* (some). Diagrams like those in Figures 3(c) and 3(d) were not compatible with the ∃!∀ sentences. Figure 3(c) is incompatible because there are too few connections. Figure 3(d) is incompatible because there are too many connections. We used two types of incompatible pictures only because of thechnical reasons.

1.3 Participants

60 native German speakers (40 female, mean age: 24, ranging between 20 and 40) from the University of Tübingen participated in the study. Participants were naïve to the purpose of the study. They were paid €10 compensation.

1.4 Hypotheses and Predictions

Since most of the present experiment replicates the study of Chemla & Spector (to appear), the major hypotheses and predictions can be found therein. We will shortly summarize the main points here. In monotone increasing environments local implicatures do generally imply global implicatures, but not reversely. In these cases the distinction between local and global accounts is rather subtle.



(a) compatible with literal meaning and global implicature

(b) compatible with literal interpretation only

(c) false because of insufficient connections

(d) false because of to many connections

Figure 3: Diagrams for ∃!∀ sentences

It is hard to test the existence of locally drawn implicatures in upward entailing contexts empirically (see, however, Chemla & Spector to appear, Geurts & Pouscoulous 2009). The same applies to the existence of global implicatures in upward entailing contexts. Generally, if one reading implies another one and both are rated equally, nothing can be concluded since the sentences may just be ambiguous.

If we do find evidence for the existence of global scalar implicatures, this does at least provide evidence that strict locality is not enough. This is because when applied to sentences containing embedded scalar items, local accounts are not able to generate global implicatures. If implicature computation was strictly local, global implicatures should not be drawn.

The heterogenity of global accounts is another complicating factor. Global accounts differ with respect to the readings they assume to be possible in upward entailing environments. Some global accounts (eg.) posit global strategies that produce readings equivalent to eg. the local implicature of ∀∃ sentences. Other more restricted global accounts do not manage to generate any of the local readings. Therefore, it is not possible to decide between globality and locality by testing scalar items in upward entailing contexts. Even clear evidence for so called local readings does only discard some global accounts, but not globality in general.

In non-monotone contexts implication from local to global readings does not hold. Furthermore, no global account can generate any of the local readings in such contexts. Therefore, as Chemla & Spector point out, non-monotone contexts are a good test case. If local implicatures are detected in such environments, strict globality is clearly insufficient.

Yet another distinction within the global accounts can be drawn on the basis of sentences like (5) which have not been investigated experimentally thus far. Replacing *jeder* (every) in (5) with any other item from its scale does not yield a

sentence that entails the original one. Therefore, most theories do not consider this substitution an alternative to be negated by implicature. Some global accounts do, however, allow for such implicatures. Luckily, for sentences like (5) global readings do not imply literal readings. Therefore acceptance of global readings of such sentences would have straightforward theorectical implications. We will now spell out the predictions for each sentence condition in some detail.

1.4.1 Predictions for ∃!∀ sentences

Local accounts let us predict that *strong* diagrams are accepted more often than *literal* ones. Furthermore, they do not predict any difference between acceptance of *literal* and *weak* diagrams. Global accounts, on the other hand, predict that *weak* diagrams are accepted more often than *literal* diagrams. At least more restricted global accounts do not predict any difference between *strong* and *weak* diagrams.

1.4.2 Predictions for ∃!∀ sentences

Global accounts predict that *all* diagrams are rated better than *literal* diagrams. They cannot explain more acceptance of *local* diagrams than of *false* diagrams. Local accounts predict both *all* and *local* diagrams to be rated better than *litaral* and false *diagrams*.

1.4.3 Predictions for ∃!∀ sentences

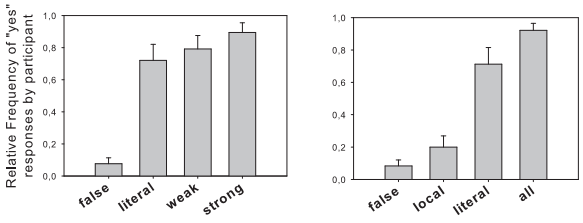
Some global accounts would predict *global* pictures to be rated better than *literal* pictures. Other global accounts do not predict any difference. Since the literal reading is true in both types of diagrams, they should be rated better than *false1* and *false2*.

1.5 Results

Overall participants were able to solve the task. In the distractor items the picture verification was done accurately in 88% (sd : 33%) of the cases. All participants were above 70% and significantly better than chance. On average it took participants $4989ms$ (sd : $2176ms$) to come up with their answer in the filler trials. We will now consider results for the three sentence conditions in turn.

1.5.1 Results for ∃!∀ sentences

The judgment pattern we obtained for ∃!∀ sentences is depicted in Figure 4(a). The *strong* diagrams were judged to be true in 90% of the cases, *weak* diagrams received "yes" judgments in 79% , *literal* diagrams in 72% and *false* in 0.9%. A logit mixed effects model with binomial link function containing a fixed effect of *diagram* as well as random effects of items and participants revealed a clear effect of *diagram*. In particular *false*, *literal* and *weak* diagrams all received fewer "yes" judgments than *strong* diargams (all $p < .001$). Another logit mixed effects model only modelling data for *literal* and *weak* diagrams also revealed an reliable effect of *diagram* ($p<.05$). Finally, a model only considering *false* and *literal* diagrams also revealed a significant effect of *diagram* ($p<.001$).



(a) ∀∃ sentences (b) ∃!∃ sentences

Figure 4: Mean judgments

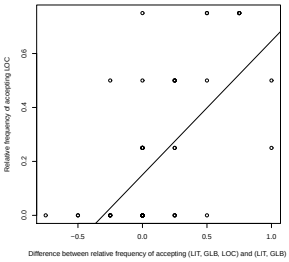


Figure 5: Correlation

1.5.2 Results for ∃!∃ sentences

The response pattern for ∃!∃ sentences is presented in Figure 4(b). Unsurprisingly, these sentences were accepted as truthfull description of *false* diagrams in only 0.8% of all cases. The *local* diagrams received "yes" responses in 20%, *literal* in 71% and *all* in 93% of the time. A logit mixed effects model with *diagram* as fixed effect as well as item and participant as random effects revealed a signficant effect of *diagram*. The *all*, *literal* and *local* diagrams were all accepted more often than the *false* diagram (all $p < .001$). A simpler model only considering *local* and *literal* diagrams reveald a reliable effect of *diagram* ($p<.001$) as did a model only considering *literal* and *all* diagrams ($p<.001$).

1.5.3 Tendency for acceptance of diagrams compatible with local readings

Participants who accepted *local* diagrams together with ∃!∃ sentences also on average made a greater difference between *strong* and *weak* diagrams when they were judging ∀∃ sentences. This is reflected in a significant positive correlation between these two variables. The first variable is the proportion of "yes" responses per participant to the *local* diagrams in the ∃!∃ sentences. The second one is the difference per participant between the proportion of "yes" responses to *strong* and *weak* diagrams in the ∃!∃ sentences. This correlation ($p < .001$) is visualized in Figure 1.5.3.

References

E. Chemla & B. Spector (to appear). 'Experimental evidence for embedded scalar implicatures'. *Journal of Semantics* .

B. Geurts & N. Pouscoulous (2009). 'Embedded implicatures?!?'. *Semantics and pragmatics* **2**(4):1–34.

S. Sternberg (1966). 'High Speed Scanning in Human Memory'. *Science* **5**:652–654.