Experimental detection of embedded implicatures[*,**]

Emmanuel Chemla[1,2] and Benjamin Spector[1]

[1] Institut Jean-Nicod (CNRS - EHESS - ENS, Paris, France)
[2] LSCP (CNRS - EHESS - ENS, Paris, France)

1 Theories of scalar implicatures: globalism vs. localism

According to the Gricean approach to scalar implicatures (SIs for short), SIs are pragmatic inferences that result from a reasoning about the speaker's communicative intentions. In recent years, an alternative view of SIs (let us call it the 'grammatical view' of SIs) has been put forward, according to which they result from the optional presence of a covert so-called exhaustivity operator in the logical form of the relevant sentences and are thus reducible to standard semantic entailment (cf. Chierchia 2006, Fox 2007, Chierchia et al. in press, a.o).

While these two radically different approaches do not make distinct predictions in simple cases, they do for more complex ones. In particular, if the 'grammatical approach' is correct, then the exhaustivity operator should be able to occur in an embedded position (just like *only*), so that the strengthening, say, of 'some' into 'some but not all' could occur 'locally', under the scope of linguistic operators. This approach is often called 'localist', as opposed to pragmatic, so-called 'globalist' approaches (See also Landman 1998, Chierchia 2004).

Consider for concreteness the following example:

(1) Every student solved some of the problems.

The standard neo-Gricean mechanism predicts that (1) should be interpreted as implying the negation of its scalar alternative, i.e. the negation of 'Every student solved all of the problems'. Hence, (1) should give rise to the following reading (henceforth, we'll refer to this reading as the 'global reading'):

(2) Every student solved some of the problems and at least one student didn't solve them all.

If, however, the strengthening of 'some' into 'some but not all' can occur at an embedded level, as predicted by localist approaches, one expects that another possible reading for (1) is the one expressed by (3) below (which we will henceforth call the 'local reading'):

(3)   Every student solved some but not all the problems.

It thus seems that determining the possible readings of sentences like (1) should provide decisive evidence in the debate between localism and globalism. This is unfortunately not so. For several formalized *globalist* theories of SIs (e.g., Spector 2003, 2006, van Rooij and Schulz 2004, Chemla 2008, 2009b) also predict that (3) is a possible reading of (1).[3]

The first goal of this paper is to provide new experimental data which show, contrary to claims put forward in a recent paper by Geurts and Pouscoulous (Geurts and Pouscoulous 2009), that (3) *is* a possible reading for (1). A second goal of this paper is to examine a case where localism and globalism are bound to make different predictions, and to test it with a similar experimental paradigm.

2   Geurts and Pouscoulous' results

G&P collected truth-value judgments for sentence-picture pairs, asking subjects to evaluate the relevant sentence as true, false, or ambiguous between a true and a false reading. One of their crucial conditions consisted of the sentence 'All the squares are connected with some of the circles', paired with the picture in Fig. 1.
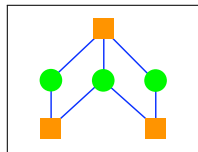

Fig. 1: Item from Geurts and Pouscoulous's (2009) experiment 3 (their Fig. 2)

Here are the three relevant potential readings for the sentence they used:

(4)   a.   <u>Literal Reading</u>. Every square is connected with at least one circle.

b.   <u>Global Reading</u>. Every square is connected with at least one circle, and it's not the case that every square is linked with all the circles.

c.   <u>Local Reading</u>. Every square is connected with at least one circle, and no square is connected with all the circles.

G&P found that virtually all the subjects considered the sentence to be true in Fig. 1, even though it is false under the local reading (the top square is linked to *all* the circles), and concluded that the local reading does not exist. We challenge this interpretation, by pointing out that there are several reasons why the strong reading, even if it existed, might have been very hard to detect:

---

[3] These theories do not derive this reading by localist means, of course. They argue instead that the proposition: 'Some students solved all the problems' should be added to the list of negated scalar alternatives of (1).

– **(i)** G&P's pictures are hard to decipher; in particular, the unique falsifier of the local reading (i.e. the top square) is hard to identify as such.

– **(ii)** Note that the local reading a-symmetrically entails the global reading, which in turn asymmetrically entails the literal reading. Meyer and Sauerland (2009), among others, argue that subjects, due to some kind of a charity principle, tend to interpret ambiguous sentences under their weakest readings, unless a stronger available reading is particularly 'accessible' (see also, e.g., Crain and Thornton 2000, Abusch 1993, Reinhart 1997). If the global and the local readings are equally accessible, it follows that the local reading will be be hard to detect experimentally even if it exists.

3   Our experimental design

Like G&P, we used a sentence-picture matching task, but with some crucial modifications. We believe that our design improve on that of G&P's in the following respects:

– **(re i)** The falsifiers of the strong reading are easy to identify (see Fig. 2 below, and in particular the **weak** condition which is the counterpart of G&P's item represented in Fig. 1).

– **(re ii)** Instead of asking for absolute judgments of truth or falsity, we asked for graded judgments: subjects were asked to position a cursor on a continuous line going from 'No' (i.e. 'false') on the left, to 'Yes' (i.e. 'true') on the right.[4] By offering subjects more options, we hoped to get more fine-grained results, which could reveal differences that remained hidden when subjects were given only two or three options, and thus to overcome some of the consequences of the charity principle. More specifically, we hypothesized that given a sentence S and two distinct pictures $P1$ and $P2$, if the set of available readings for S that are true in $P1$ is a proper subset of those that are true in $P2$, then the degree to which S will be judged true will be lower in the case of $P1$ than with $P2$.

4   Experiment 1: scalar items in universal sentences

In this experiment, we showed that the local reading is available for sentences like (1) above: French scalar items like 'certains' (*some*)[5] and 'ou' (*or*), when embedded under universal quantifiers, can give rise to readings in which they seem to be equivalent to, respectively, 'some but not all' or an *exclusive* disjunction.

---

[4] See Chemla (2009a,c) for the use of a similar methodology to collect judgments in pragmatics, and the references cited therein.

[5] Note that French *certains*, unlike its singular counterpart *un certain* or English *certain*, does not force a specific reading.

4.1 Experimental items

The items explicitly discussed in the instructions were presented first to allow participants to get used to the display and to the task.[6] After that, participants ran a first block of items in which all target conditions were repeated several times (in pseudo-random order). Participants then could take a short break before moving to a second block of items instantiating the same experimental conditions (with superficially different pictures). In a last experimental block of items, some control conditions were administered.

Target conditions: universal sentences. Each item consisted of a sentence and a picture. We used the two distinct sentence-types, illustrated in (5) and (6). For each of them, we were interested in the availability of three distinct potential readings, namely the literal, the global and the local readings:

(5)   Chaque lettre est reliée à certains de ses cercles.
      Each letter is connected to some of its circles.
   a.   Literal Reading: Each letter is connected to at least one of its circles.
   b.   Global Reading: Each letter is connected to at least one of its circles, and it is not the case that each letter is connected to all its circles.
   c.   Local Reading: Each letter is connected to at least one of its circles, and no letter is connected to all its circles.

(6)   Chaque lettre est reliée à son cercle rouge ou à son cercle bleu.
      Each letter is connected to its red circle or to its blue circle.
   a.   Literal Reading: Each letter is connected to its red circle, its blue circle or both.
   b.   Global Reading: Each letter is connected to at least one of its circles, and it is not the case that each letter is connected to both the red and the blue circle.
   c.   Local Reading: Each letter is connected to its red circle or its blue circle but none is connected to both.

Each of these sentences was paired with various pictures, giving rise to the following four target conditions (see Fig. 2): **false:** no reading is true, **literal:** only the literal reading is true, **weak:** both the literal and the global readings are true but the local reading is false, **strong:** all readings are true.

Control conditions: downward entailing (DE) environments. When scalar items are embedded in the scope of 'No' as in (7a) or (8a), it is uncontroversial that the potential 'local' readings described in (7b) and (8b) are only marginally available at best.

---

[6] The experiment involved 16 native speakers of French, with no knowledge of linguistics, ranging in age from 19 to 29 years (10 women)

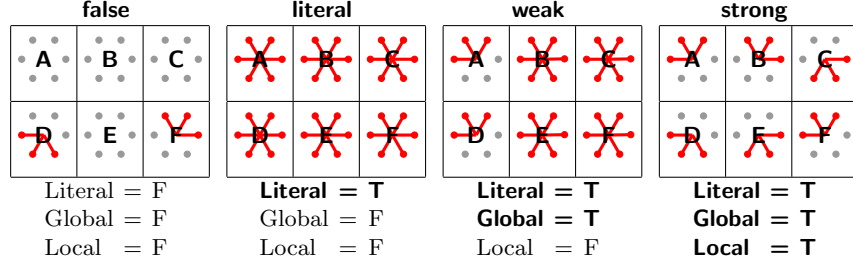| false	literal	weak	strong
Literal = F	**Literal = T**	**Literal = T**	**Literal = T**
Global = F	Global = F	**Global = T**	**Global = T**
Local = F	Local = F	Local = F	**Local = T**

Fig. 2: Illustrative examples of the images used in the different conditions **false**, **literal**, **weak** and **strong** for the test sentence (5). We also reported below each image whether the literal, global and local readings are true (T) or false (F).

(7) a. Aucune lettre n'est reliée à certains de ses cercles.
      No letter is connected to some of its circles.

     b. Potential local reading: No letter is connected to some but not all of its circles.

(8) a. Aucune lettre n'est reliée à son cercle rouge ou à son cercle bleu.
      No letter is connected to its red circle or its blue circle.

     b. Potential local reading: No letter is connected to exactly one of its two circles.

Sentences like (7a) and (8a) were thus used as controls, to check that participants do not access the 'local' reading for such sentences, or do so only marginally (given the marginal availability of the local reading).They were paired with pictures instantiating the following three conditions: **false:** no reading is true **?local:** only the local reading is true, **both:** both the local and the literal readings are true.

4.2 Results and interpretation

Main result: detection of the local reading. Fig. 3 reports the mean ratings in the target conditions. The relevant t-tests show that all differences between two consecutive bars are significant.[7]

The crucial result is that the ratings are higher in the **strong** condition than in the **weak** condition, even though the two conditions differ only according to the truth value of the local reading. This difference provides important support

---

[7] SOME: **false** vs. **literal**: $F(1, 15) = 14$, $p < .01$; **literal** vs. **weak**: $F(1, 15) = 27$, $p < .001$; **weak** vs. **strong**: $F(1, 15) = 25$, $p < .001$. OR: **false** vs. **literal**: $F(1, 15) = 6.2$, $p < .05$; **literal** vs. **weak**: $F(1, 15) = 22$, $p < .001$; **weak** vs. **strong**: $F(1, 15) = 17$, $p < .001$. Note that 4.6% of the responses were excluded as outliers or for technical reasons. Statistical analyses presented here are computed per subject, per item analyses yield similar resuts.
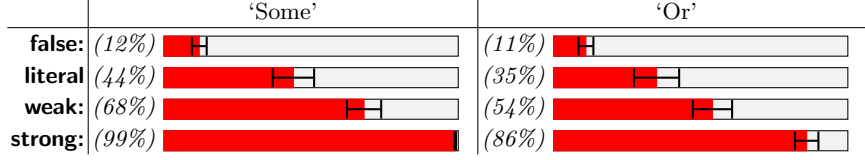
Fig. 3: Main results: Mean position of the cursor in the target conditions of Exp. 1. Error bars represent standard errors to the mean.

for the existence of the local reading. Indeed, these results are fully explained if we assume that a) the target sentence is ambiguous between the literal reading, the global reading and the local reading, and b) the more readings are true, the higher the sentence is rated. They are not expected if only the literal and the global readings exist.

Control result : downward-entailing environments Fig. 4 reports the results for the control conditions. For the scalar item 'some', the relevant t-tests show a significant difference between all pairs of conditions, while for the scalar item 'or', there is no difference between the **false** condition and the **?local** condition.[8]
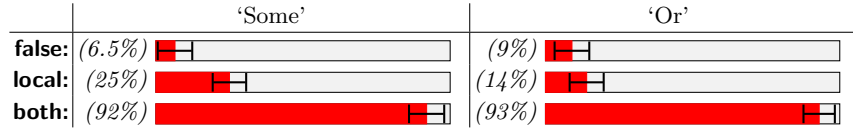

Fig. 4: Mean responses for the DE control conditions in exp. 1 (see §4.1).

In the case of 'some', we cannot exclude that participants perceived the 'local' reading, because the **?local** condition is judged a little higher than the **false** condition. But this result is not terribly disturbing for two reasons. First, it does not generalize to the scalar item 'or'. Second, the control sentences receive a much lower rating in the condition **?local** than in conditions where it is uncontroversial that the target sentence has a true reading. Note that even with the scalar item 'some', the condition **?local** is rated at a radically lower level than the condition **both** (25 % vs. 92 %); more importantly, in the case of 'some', the condition **?local** is rated much lower than conditions in which it is uncontroversial that the target sentence has a true reading (consider for instance the important difference between this **?local** condition and the **weak** condition – which involved the target sentences. This difference is statistically significant: $F(1, 15) = 22$, $p < .001$).

---

[8] SOME: **false** vs. **?local**: $F(1, 15) = 6.5$, $p < .05$; **?local** vs. **both**: $F(1, 15) = 43$, $p < .001$. OR: $F(1, 15) = .45$, $p = .51$ and $F(1, 15) = 60$, $p < .001$, respectively.

5   Experiment 2: non-monotonic environments

In this second experiment, we tested cases for which pragmatic and grammatical theories are bound to make different predictions. This happens with sentences where a scalar item like 'some' or 'or' occurs in a non-monotonic environment:

(9)   Exactly one letter is connected to some of its circles.

(10)   Exactly one letter is connected to its blue circle or its red circle.

The relevant potential readings (i.e. those that the sentence could in principle have according to various theories) can be paraphrased as follows:[9]

(11)   Potential readings of (9)

   a.   <u>Literal meaning</u>: one letter is connected to some or all of its circles, the other letters are connected to no circle.

   b.   <u>Global reading</u>: one letter is connected to some but not all of its circles, the other letters are connected to no circle.

   c.   <u>Local reading</u>: one letter is connected to some but not all of its circles, the other letters may be connected to either none or all of their circles.

(12)   Potential readings of (10)

   a.   <u>Literal meaning</u>: one letter is connected to its blue circle or its red circle or to both, the other letters are connected to no circle.

   b.   <u>Global reading</u>: one letter is connected to exactly one of its two circles, the other letters are connected to no circle.

   c.   <u>Local reading</u>: one letter is connected to exactly one of its two circles, the other letters may be connected to either none or both of their circles.

Because the scalar item now occurs in a non-monotonic environment, the local reading does not entail the global reading. In fact, it does not even entail the literal reading. This is of major importance for three reasons. First, globalist theories are bound to predict readings that entail the literal reading. Hence they cannot predict local readings like (11c) or (12c) in these non-monotonic cases. Second, the fact that the local reading does not entail any of the other two potential readings could automatically make it easier to detect (according to a charity principle). Finally, this very fact allowed us to construct cases where only the local reading is true and to assess its existence independently of the other readings.

5.1   Experimental items

The task and the instructions were essentially the same as in experiment 1.The items were presented just like in experiment 1: the examples from the instructions were presented first; then came two blocks of target conditions,and finally

---

[9] The global reading (11b) is obtained by adding to the literal reading the negation of the alternative sentence "Exactly one letter is connected to all its circles".

came a block with exactly the same control conditions as in experiment 1. The target conditions involved French translations of (9) and (10). Each of these sentences was paired with various pictures, giving rise to the following four target conditions, which represent all the possible combinations of true and false readings, and are illustrated in Fig. 5: **false:** no reading is true, **literal:** only the literal reading is true, **local:** only the local reading is true and **all:** all three readings – literal, global and local – are true.

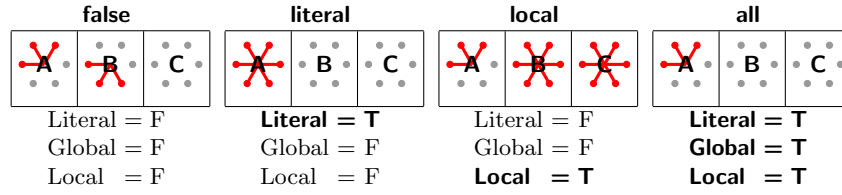| false	literal	local	all
Literal = F	**Literal = T**	Literal = F	**Literal = T**
Global = F	Global = F	Global = F	**Global = T**
Local = F	Local = F	**Local = T**	**Local = T**

Fig. 5: Illustrative examples of the images used to illustrate the different conditions **false**, **literal**, **local** and **all** for the test sentence (9). We also reported below each image whether the literal, global and local readings are true (T) or false (F).

5.2 Results

Main result: the local reading exists. Fig. 6 reports the mean ratings of the target conditions.[10] All 2 by 2 differences are significant, except for the **local** vs. **literal** conditions in the case of 'or'.[11]
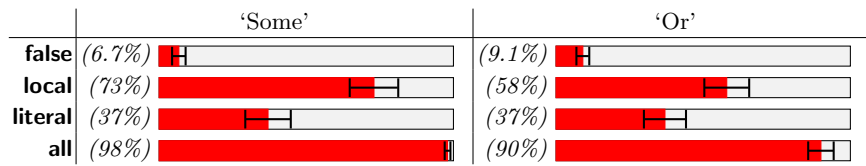

Fig. 6: Mean responses in the target conditions of experiment 2.

---

[10] This experiment involved 16 native speakers of French, with no prior exposure to linguistics, ranging in age from 18 to 35 years (9 women). 14% of the responses had to be excluded for various technical reasons. All statistical analyses presented below are computed per subject; per item analyses yielded similar results.

[11] SOME: **false** vs. **literal**: $F(1,15) = 12$, $p < .01$, **literal** vs. **local**: $F(1,15) = 6.7$, $p < .05$, **local** vs. **all**: $F(1,15) = 10$, $p < .01$. OR: **false** vs. **literal**: $F(1,15) = 11$, $p < .01$, **literal** vs. **local**: $F(1,15) = 2.3$, $p = .15$, **local** vs. **all**: $F(1,15) = 18$, $p < .001$.

This first set of data qualifies the local reading as a possible interpretation for our target sentences (involving non-monotonic operators), since ı) the **local** condition is rated much higher than in the **false** condition and ii) the **local** reading is rated significantly higher than the **literal** condition, a fact which is totally unexpected under the globalist approach, but can be understood within the localist approach.

Control result: downward entailing environments. Fig. 7 reports the results for the DE control conditions (which were the same as in Exp. 1). All 2 by 2 differences are statistically significant with both 'some' and 'or'.[12]

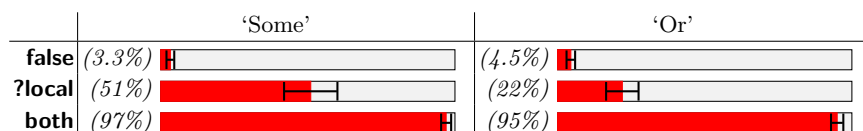|	'Some'	'Or'
false	*(3.3%)*	*(4.5%)*
?local	*(51%)*	*(22%)*
both	*(97%)*	*(95%)*

Fig. 7: Mean responses for the control conditions when administered at the end of experiment 2.

Surprisingly, the rates for the **?local** condition are higher than they were in the first experiment (compare Fig. 7 to Fig. 4), which calls for an explanation. A possible hypothesis is the following: subjects become much better at perceiving 'local' readings even in cases where they are normally dispreferred *once* they have experienced cases in which the local reading is salient. The target conditions of the second experiment seem to have precisely this property, given the results we have just presented.

6 Conclusions

Our first experiment showed that sentences in which a scalar item is embedded under a universal quantifier can be interpreted according to what we called the 'local' reading, contrary to Geurts and Pouscoulous' (2009) conclusions. We pointed out that this result is nevertheless not sufficient to establish the existence of embedded scalar implicatures (because the local reading in such a case can be predicted by a *globalist* account). In our second experiment, we focussed on a case where the local reading cannot be derived by globalist means – sentences where a scalar item occurs in a non-monotonic environment –, and we were able to detect experimentally genuinely local readings. The existence of embedded scalar implicatures is unexpected from a Gricean perspective. The grammatical approach to SIs provides one possible way of making sense of these data.

---

[12] SOME: **false** vs. **?local**: $F(1, 14) = 20$, $p < .001$; **?local** vs. **literal**: $F(1, 14) = 28$, $p < .001$. OR: $F(1, 15) = 6.1$, $p < .05$ and $F(1, 15) = 190$, $p < .001$, respectively.

Bibliography

Abusch, D. (1993). The scope of indefinites. *Natural Language Semantics 2*(2), 83–135.

Chemla, E. (2008). *Présuppositions et implicatures scalaires: études formelles et expérimentales.* Ph. D. thesis, ENS.

Chemla, E. (2009a). Presuppositions of quantified sentences: experimental data. *Natural Language Semantics 17*(4), 299–340.

Chemla, E. (2009b). Similarity: towards a unified account of scalar implicatures, free choice permission and presupposition projection. Under revision for *Semantics and Pragmatics.*

Chemla, E. (2009c). Universal implicatures and free choice effects: Experimental data. *Semantics and Pragmatics 2*(2), 1–33.

Chemla, E. and B. Spector (2009). Experimental evidence for embedded implicatures. Ms. IJN & LSCP.
http://www.emmanuel.chemla.free.fr/Material/Chemla-Spector-eSI.pdf.

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In A. Belletti (Ed.), *Structures and Beyond.* Oxford University Press.

Chierchia, G. (2006). Broaden Your Views: Implicatures of Domain Widening and the 'Logicality' of Language. *Linguistic Inquiry 37*(4), 535–590.

Chierchia, G., D. Fox, and B. Spector (in press). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics.

Crain, S. and R. Thornton (2000). *Investigations in Universal Grammar: A guide to experiments on the acquisition of syntax and semantics.* MIT Press.

Fox, D. (2007). Free Choice and the theory of Scalar Implicatures. In U. Sauerland and P. Stateva (Eds.), *Presupposition and Implicature in Compositional Semantics*, pp. 537–586. New York, Palgrave Macmillan.

Geurts, B. and N. Pouscoulous (2009, July). Embedded implicatures?!? *Semantics and Pragmatics 2*(4), 1–34.

Landman, F. (1998). Plurals and Maximalization. In S. Rothstein (Ed.), *Events and Grammar*, pp. 237–271. Kluwer, Dordrecht.

Meyer, M.-C. and U. Sauerland (2009). A pragmatic constraint on ambiguity detection. *Natural Language & Linguistic Theory 27*(1), 139–150.

Reinhart, T. (1997). Quantifier scope: How labor is divided between QR and choice functions. *Linguistics and Philosophy 20*(4), 335–397.

van Rooij, R. and K. Schulz (2004). Exhaustive Interpretation of Complex Sentences. *Journal of Logic, Language and Information 13*(4), 491–519.

Spector, B. (2003). Scalar implicatures: Exhaustivity and Gricean reasoning. In B. ten Cate (Ed.), *Proceedings of the Eigth ESSLLI Student Session*, Vienna, Austria. Revised version in *Questions in Dynamic Semantics*, eds. M. Aloni, P. Dekker & A. Butler, Elsevier, 2007.

Spector, B. (2006). *Aspects de la pragmatique des opérateurs logiques.* Ph. D. thesis, Université Paris 7.