The role of probability in the accessibility of scalar inferences

Emmanuel Chemla
ENS

Jacopo Romoli
Ulster University

1 The over generation problem of sentences with multiple scalar terms

Our starting puzzle concerns sentences with more than one scalar term. It has been observed that theories of alternatives and scalar implicatures sometimes there predict inferences which are intuitively absent. For instance, given fairly standard assumptions about scalar implicatures and alternatives, (1-a) should implicate (1-b), while intuitively it does not (Fox 2007, Magri 2009; see also Chemla 2010, Romoli 2012, Trinh & Haida 2015).[1]

(1) a. Some students read all of the books.
b. $\not\rightarrow$ *some students read no book* (= ¬ *all students read some*, schematically)

This inference is predicted given the assumption about alternatives in (2) (Horn 1972, Sauerland 2004 a.o.) and that about implicatures in (3) (Fox 2007, Spector 2007 a.o.).[2]

(2) The set Alt(S) contains all sentences and only those sentences that can be obtained by replacing one or more scalar items in S with their scale-mates.

(3) Scalar inferences arise from the negation of alternatives which are not weaker than the assertion (alternatives that are either stronger than the assertion or logically independent from it).[3]

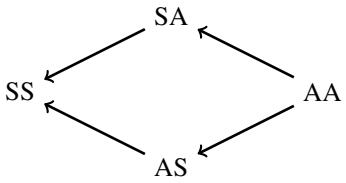
To see how the prediction comes about, consider the set of alternatives of (1-a) and their logical relations in (4), where the arrows represent asymmetric entailment. The problem for (1-a) arises because the alternative AS is logically independent from the assertion SA and therefore, if relevant, will end up being negated, giving rise to the unwanted inference that it is not true that all of the students read some books (= some student didn't read any).

1 Magri (2009) also discusses the case of *Few students read some of the books*, which does not seem to implicate the negation of *no student read all the books*. However here the sentence is judged to be unnatural by some speakers, arguably because it involves a positive polarity item, *some*, in a downward entailing context. Notice that this problem does not extend to (1-a).

2 Motivations for considering not only stronger alternatives but also non-weaker alternatives come for all types of accounts from scalar implicatures arising in non-monotonic contexts (see Chierchia et al. 2012 for discussion).

3 Notice that the alternatives that can be negated is a more restricted set than the set of all non-weaker alternatives. In particular, at least a condition that disallows alternatives to be excluded if their exclusion entails the inclusion of another alternative needs to be added, see Sauerland (2004) and Fox (2007) among others. This complication is irrelevant for our purposes, so we will stick to the simpler version in the text.

(4)
Some student read **A**ll of the books	(SA)
Some student read **S**ome of the books	(SS)
All students read **A**ll of the books	(AA)
All students read **S**ome of the books	(AS)



Fox and Magri's response and its problems

In response to the data above, Fox (2007) and Magri (2009) propose to amend the theory of alternatives, in such a way that the problematic alternatives cannot be constructed. In their idea, AS can never be an alternative to SA, because getting the former from the latter by replacing scalar items one at a time requires *weakening* steps (from AS to SS or from AA to AS), which are ruled out by hypothesis. Visually, there is no way to 'walk' the diagram in (4) from SA to AS in steps involving single replacements and without going in the direction of some entailment arrow. More formally, they replace the definition in (2) with:

(5) The set $\mathrm{Alt}(S)$ contains S and is recursively defined as follows:
$A \in \mathrm{Alt}(S)$ iff there is $A' \in \mathrm{Alt}(S)$ such that: A is obtained from A' by replacing a single scalar item in A' with a scale-mate and A is not weaker than A'.

The novelty here is the filtering constraint in the underlined part. The alternatives of (1-a) given (5) are now as in (6); (6) does not contain the problematic alternative AS and therefore the inference in (1) is predicted to never arise.

(6)
Some students read **A**ll of the books	(SA)
All students read **A**ll of the books	(AS)

Fox/Magri's solution thus blocks the problematic inference above. There are reasons to believe, however, that their constraint is too strict and ends up blocking too many alternatives. In particular, there appears to be an unpredicted asymmetry. The inference in (7-b) from (7-a) (that is inferring the negation of SA from the assertion of AS) is intuitively possible. Indeed, this inference has been experimentally detected (see, e.g., Chemla & Spector 2011, Clifton & Dube 2010, Gotzner & Romoli 2015, Potts et al. 2015).

(7) a. All students read some of the books
b. $\rightsquigarrow$ *no student read all of the books*

This is problematic for Fox/Magri's idea because the situation with regard to alternatives is symmetrical: getting from AS to SA also involves a weakening step (see diagram in (4) again). As a consequence, according to Fox and Magri's constraint, SA can never be an alternative of AS (no more than AS is an alternative to SA) and therefore the inference in (7-b) should never arise from (7-a).
The problem related to (7-a) could be solved by Fox and Magri if they assume that scalar implicatures can appear in embedded positions (Chierchia et al. 2012, Chierchia 2013, Chemla & Spector 2011 for discussion). This is because (7-b) could also be obtained locally,

by enriching *some*, so that the sentence ends up being equivalent to *All student read [some but not all] of the books*, which entails (7-b). Therefore (7) is not problematic if we assume that embedded scalar implicatures are possible.
However, this resort to local implicatures doesn't generalize. Reproducing an argument from Chemla (2009) (also developed with free choice inferences), we will show that the same problem as above applies to (8-a), but there local implicatures cannot help. First note that (8-a) has a reading giving rise to the inference (8-b) that all students read some of the books. The existence of this inference is controversial (Romoli 2012, Trinh & Haida 2015), but Gotzner & Romoli (2015) have found parallel experimental evidence as that of Chemla & Spector (2011), suggesting that the inference is indeed there.[4]

(8) a. No student read all of the books.
b. $\rightsquigarrow$ *all student read some of the books*

In this new case, the inference cannot be obtained by local enrichments (contrary to (7-a)).[5] The inference can be obtained as the negation of the logically independent alternative "not every student read some of the books" (NeS), but, as in the case of (7-a) above, it is easy to verify that obtaining NeS from NA involves a weakening step. All in all, the inference can be explained, but crucially not within Fox/Magri's constraint.

2 The proposal

We propose the following alternative perspective to Fox/Magri's hard constraint: the standard theory of alternatives in (2) is correct and provides the whole set of possible alternatives to a sentence. However, we submit that there could be general reasons why inferences like not-AS are not made part of the intended message, given what SA and not-AS mean. The intuition is that if we learn that some students read all of the books (SA), the odds that some students didn't read any of the books (not-AS) substantially decrease (simply because there are less candidates for not having read books). In this section, we propose a specific implementation, test of this idea and show how Fox/Magri's specific, hard constraint can be replaced with a more general, soft constraint.

2.1 A possible implementation

Consider a sentence S, with propositional meaning A_S and a (potential) inference with meaning I_S. In parsing S, the hearer will ask herself whether the odds that I_S is true decrease *substantially* given that A_S is true. If the odds substantially decrease, there is conflicting evidence as to whether I_S is true: on the one hand it is a possible scalar implicature of S, on the other hand the bare interpretation of S makes it an implausible candidate inference by

4 More precisely, Gotzner & Romoli (2015) extend Chemla & Spector's (2011) experiment by comparing the inference in (8-b) and that in (7-b) and they find no evidence of a difference between the two inferences.

5 The inference could be obtained locally by assuming a decomposition of *no* in terms of *every ...not*. This decomposition, however, appears *ad hoc* and problematic (see Chemla 2009 for discussion).

reducing its odds. Going by the information provided by the mere assertion of S, the hearer may thus discard I_S. Formally, we can write (9). Also, it will be useful for later to set the definition in (10) of the *attractiveness of an inference*: according to principle above, if the attractiveness is low, the inference should not arise.

(9) **Constraint on inferences:**
Do not infer I_S from S if $\text{odds}(I_S|A_S)$ are substantially lower than $\text{odds}(I_S)$

(10) **Attractiveness**: The *attractiveness* of an inference from S to I is defined as:

$$\mathscr{A}(S \rightsquigarrow I_S) = \frac{\text{odds}(I_S|A_S)}{\text{odds}(I_S)}$$

2.2 Working, practical hypotheses for computations

In order to estimate probabilities and the attractiveness $\mathscr{A}$ of inferences as defined in (10), we make use of the following simplified framework: a situation is described as $\langle n, b, p\rangle$, with n the number of students, b the number of books, p the constant probability that a given student reads a given book.[6] One can then compute the probabilities for the sentences above to be true, for their negations, for conjunctions of these elements, for these elements assuming the truth of some other sentence, etc. For instance, we can answer questions such as: what's the odds that not-NA holds in a given $\langle n, b, p\rangle$ situation? What if we also know that SA is true? With such numbers in hand, we will be able to evaluate the current proposal.

2.3 Test and results

To evaluate our proposal, we first randomly picked 200 situations, by sampling 200 values of $\langle n, b, p\rangle$ uniformly in $[5, 60] \times [3, 20] \times [.1, .9] \times [.1, .5]$.[7] We then computed the attractiveness of various inferences $\mathscr{A}(S \rightsquigarrow I_S)$ for each of these situations. The idea is that when trying to evaluate a possible inference 'out of the blue,' we cannot rely on the specifics of the situation of utterance. Here we use sampling and aggregation over a range of hypothetical situations as a mean to model this out of the blue evaluation process.[8]
The R script we used to compute these attractiveness measures is available at http://semanticsarchive.net/Archive/zc4NDExN/Chemla-Romoli-AlternativesAndProbabilities.html

6 Assume that p doesn't depend on what other students did nor on what a student did with other books.

7 We adjusted the technical parts of the sampling process in two ways. First, we disregarded sampled situations for which we obtained an infinite value for some attractiveness value due to an approximation error (numbers below 10^{-16} are not handled properly). Second, we replaced sampled situations for which the semantic meaning of *few* would be equivalent to *at most 1 individual* with the same situation but where the semantic meaning of *few* would be *at most 2 individuals*, which seems more sensible.

8 Notice that under appropriate hypotheses, one could replace sampling simulations with analytical computations of probability distributions.

2.3.1 Solving the initial problems

Consider the case of SA and AS. First, we want to explain why not-AS is not an inference associated with an assertion of SA. Intuitively, our proposal relies on the fact that if we learn that some students read all of the books (SA), the odds that some students didn't read any of the books (not-AS) decrease. Our first result confirms this intuition: the attractiveness $\mathscr{A}(\text{SA} \rightsquigarrow \text{not-AS})$ is rather low indeed, providing reasons not to derive the not-AS inference from SA (median=.005, see Fig. 1).

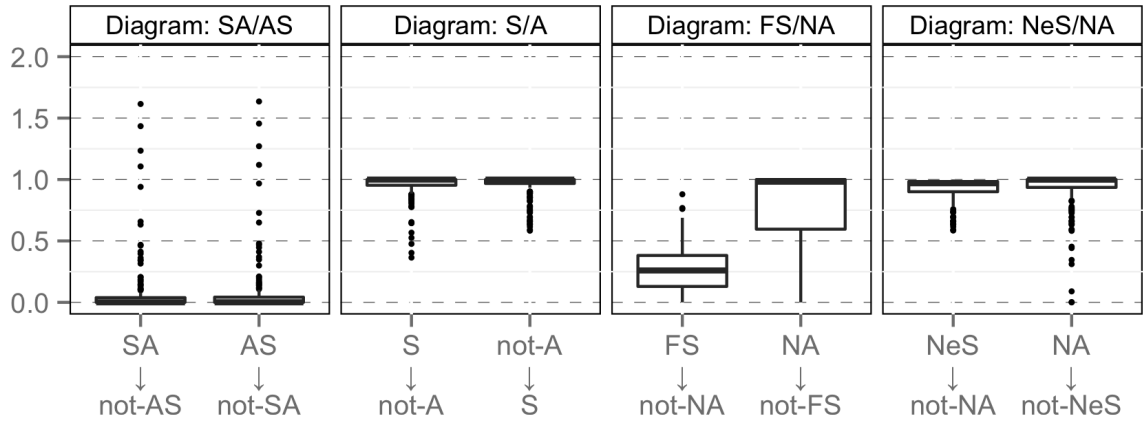


Figure 1 Attractiveness scores of various inferences

Remember however that we assume that inferences are blocked only if their attractiveness is *substantially* below 1. In practice, we will be interested in *relatively* low attractiveness values. The fact that the SA⇝not-AS inference is not attested thus follows for us from its low attractiveness compared to that of other inferences. Fig. 1 is sufficient to assess relative attractiveness values and, for instance the blocking of the SA⇝not-AS inference is visible (and confirmed by various statistics, e.g., it is significantly lower than the classic cases from diagram A/S according to Wilcoxon Signed-Rank tests, $V > 5000, p < 10^{-5}$).
Notice that, by the same token, the inference AS⇝not-SA, which is (11-b) from (11-a), repeated from above, is also predicted to be ruled out. This is not a problem, however, because, as mentioned above, there is another way to derive the inference; a way that does not involve complex alternatives, e.g., through local exhaustification (Chierchia et al. 2012 among others). This alternative route is not sensitive to the current considerations: no 'global' propositional I_S inference can be isolated from local exhaustification processes, and the current constraint would therefore be silent. Hence, the proposal here predicts that there is no global way of obtaining (11-b) from (11-a). This, however, does not exclude the possibility that (11-b) is an inference of (11-a), but it predicts that this will not be obtained by global mechanisms involving the SA alternative.[9]

9 See Chemla et al. in press for an entirely different type of evidence pointing in the same direction.

(11) a. All students read some of the books.
b. ⇝ *no student read all of the books*

Our proposal, therefore, is on a par with Fox and Magri with respect to (11-a), in that we also have to rely on a local solution to obtain the inference in (11-b). Importantly, however, our proposal improves on Fox and Magri with respect to the corresponding inference NA⇝not-NeS repeated below in (12-b) from (12-a).

(12) a. No student read all of the books.
b. ⇝ *all student read some of the books*

This inference is blocked in a Fox/Magri framework because the logical relations are as before and unlike (12-a) there is no local way of obtaining this inference. As can be seen in Fig. 1, on the other hand, its attractiveness is high and it is therefore not blocked by the current framework. In particular, under this view, universal inferences for "No all" may be obtained by multiple replacements, even if SA⇝not-AS are blocked, in line with the results in Gotzner & Romoli (2015) mentioned earlier. The third part of Fig. 1 provides the results for the inference from *Few students read some of the books* to Some (=not no) students read all the books, as discussed in fn. 1. The interested reader can check that this case too gives rise to a parallel asymmetry correctly predicted in our system.

2.3.2 A potential difficulty

Consider the (unattested) inference, from NeS to not-NA in (13): this inference receives a decent attractiveness score in our simulation (see Fig. 1). So it is not blocked and this makes our proposal incomplete. We leave a resolution of it for future work.[10] Let us point, however, to a possible solution. Given that (13-a) is a negative sentence, it leads us to accommodate that the positive version of the sentence would have been unsurprising (Gualmini 2004, Crain et al. 1996, Gualmini et al. 2008 a.o.), and this might then lead to interpreting (13-a) as equivalent to 'fewer students than expected ' or, in short, 'few students.' If that informal reasoning is right, then we would be looking at the inference *Few students read some of the books* ⇝ *some student read all of the books*. This inference is intuitively absent (see fn. 1) and the current framework predicts that, see above. Therefore, if our intuitions about (13-a) are influenced by this possible reading, the problem posed by (13-a) could be solved.

(13) a. Not every student read some of the books.
b. ⇝ *some student read all the books*

10 Notice that Fox/Magri's constraint correctly blocks this case. This is, however, at the cost of over-blocking in other cases, as we saw above; a problem from which it is harder to recover than underblocking (since blocking can come from other sources).

2.4 Discussion

The soft probabilistic constraint proposed here makes the unwanted inference in (1-b) very unlikely to arise, while, at the same time, allowing other inferences such as (12-b), therefore improving over Fox and Magri's hard constraint. In general, the idea is that the unenriched version of a sentence sets the stage for additional inferences one could decide to attach to that very sentence. Before concluding, let us briefly discuss two issues.

First, some mechanisms to derive scalar implicatures do not fundamentally individuate negations of each alternative as potential inferences. Instead, so-called exhaustification systems propose that alternatives are all considered at once, a subset of them are elected and the conjunction of their negations together provides the piece of meaning to be added to the meaning of the original phrase (see also fn. 3). The current proposal however is phrased in terms of the status of 'individual inferences.' Applying this proposal without individual inferences requires a more careful investigation of how to treat composed inferences, or how to recover them from the (final) output of exhaustivity mechanisms. This point connects with the point mentioned above about embedded implicatures: the proposal is a comparison between a bare meaning and a well-identified potential inference, which local exhaustification does not yield as such.[11] Second, one could ask why we reasoned about out of the blue evaluation and sets of situations rather than on a situation per situation basis. Indeed, since attractiveness of an inference is sensitive to the specific of a situation, a stronger and natural hypothesis would be that it could directly tell us *in which* situations a possible inference may come through. Given that in some cases we observe variability of the attractiveness measure, this should indeed yield finer-grained predictions. There are two reasons not to go in that direction, at least as a first approximation. First, various aspects of scalar implicatures have been argued to be blind to the contexts (see Magri 2009). What this would mean here is that what matters is not so much what the facts are actually like, but rather what they may be like. Second, this finer-grained proposal would be hard to assess: it would amount to evaluating whether native speakers judge that I is a more appealing conclusion after an utterance of S in situations which make the odds of I high than in situations which make them low. This will most likely come out right, but there is an obvious and essential confound: I is indeed more likely in the former case, by construction.

3 Tentative extensions

One could develop the current tools and ideas to investigate other constraints, where potential inferences and bare meaning are weighted against one another. Consider for instance:

(14) Do not utter S if odds(A_S and $I_S|I_S$) are significantly higher than odds(I_S).

Simplifying a bit, this principle compares the relative informativity of an inference I_S and of the assertive part A_S (as cumbersomely measured here by odds(A_S and $I_S|I_S$), for

11 One may try however to extend the current proposal using local contexts as introduced by Schlenker (2009) for presuppositions and used in Mayr & Romoli (2014) for phenomena precisely involving embedded implicatures.

immediate applicability to the later discussion of presuppositions). It allows both assertions and inferences to be very informative or not informative at all: the flexibility here will be driven by general informativity constraints on different conversational contexts. But it crucially puts a 'packaging' or 'balancing' constraint: the assertion can be used as a threshold of informativity, and potential inferences should remain below that threshold. In other words, the most important part of the information of a message should be in the assertion, not in its (potential) inferences. This idea has the potential to explain the following pattern of judgments, where oddness occurs when the main part of the message (i.e. where surprise/unlikelihood is) is shifted from the assertion to potential scalar implicatures.[12]

(15) a. John killed some of his students. b.# John didn't kill all of his students.

(16) a.# John loves some of his children. b. John doesn't love all of his children.

Notice that the examples above do not allow us to fully test the relation between assertion and inferences, because in the case of scalar implicatures, the content of the assertion and that of the inference are inherently connected and hard to manipulate independently. Let us take this opportunity to sketch how we could apply this principle with other inference types, presuppositional inferences in particular. It is well-known that less controversial or surprising facts are easier to accommodate (in the sense of Lewis 1979). The current principle reveals the possibility that accommodation fails when the *balance* between assertion and presupposition is not right. In a toy illustration of this, the well-known observation is that (17-a) is more natural than (17-b), presumably because its presupposition is easier to accommodate. Principle (14) suggests that the crucial factor could be the relative likelihood of the presupposition and that of the assertion. This predicts that if we lower the likelihood of the assertion, accommodation should become easier. The presupposition of (17-c) is thus predicted to be easier to accommodate than that of (17-b). The judgments are difficult because the sentence is overall highly implausible, but the claim is that the deviance due to the linguistic form of the sentence or to its presuppositional aspect should be reduced in (17-c). We leave a detailed explorations of these possible extensions for future work.

(17) a. Sorry I am late, my bike was stolen. $\rightsquigarrow$ *I have a bike*
b. Sorry I am late, my spaceship was stolen. $\rightsquigarrow$ *I have a spaceship*
c. Sorry I am late, a dog broke my spaceship. $\rightsquigarrow$ *I have a spaceship*

4 Summary and conclusion

It is clear that the decision to draw an inference relies on the *a priori* likelihood of this inference. We argue, however, that the literal content of the uttered sentence should be

12 Notice that tautologies such as 'War is War' or 'John is only a human being' (Michael Franke p.c.) escape this generalization in that meaning is entirely conveyed through (other kinds of) implicatures (see Snider 2015 for discussion). It is not clear to us at this point how to evaluate the content of the implicatures in such cases, which probably count as obvious to all participants if they are to be understood, nor whether such extreme cases fall under other competing pragmatic principles.

factored in this decision. One may wonder whether this idea is a subcomponent of recent proposals which try to reduce further scalar implicatures to probabilistic reasoning (e.g., Russell 2012, Potts et al. 2015).[13] Our goal here is less ambitious. We do not propose a reduction of scalar inferences to probabilistic reasoning and we do not advance ideas about what scalar implicatures are and how they arise. Likelihood considerations (and in our cases *relativized* or *relative* likelihood considerations) ought to play a role in all approaches and we investigated how this could contribute to explaining why some candidate implicatures never arise. The core intuition also connects directly to discussions in philosophy and logic, in particular with attempts to formalize 'relevance' in ways that ressemble our attractiveness measure (Merin 1999, van Rooij & Schulz 2004 and Benz 2006 a.o.). Our results may be used to show how these discussions can yield concrete progress in very specific cases: we suggested how to explain the surprising behavior of sentences with multiple scalar items without the need to amend the classic theory of alternatives. Finally, we proposed to test these ideas using a novel sampling and aggregation method.

References

Benz, Anton. 2006. on relevance scale approaches. In E. Puig-Waldmüller (ed.), *Proceedings of sinn und bedeutung 11*, 91–105. Universitat Pompeu Fabra.

Chemla, Emmanuel. 2009. Universal implicatures and free choice effects: Experimental data. *Semantics and Pragmatics* 2(2).

Chemla, Emmanuel. 2010. Similarity: towards a unified account of scalar implicatures, free choice permission and presupposition projection. Unpublished manuscript.

Chemla, Emmanuel, Chris Cummins & Raj Singh. in press. Training and timing local scalar enrichments under global pragmatic pressures. *Journal of Semantics* .

Chemla, Emmanuel & Benjamin Spector. 2011. Experimental evidence for embedded scalar implicatures. *Journal of Semantics* 28(3). 359–400.

Chierchia, Gennaro. 2013. *Logic in grammar: Polarity, free choice, and intervention*. Oxford University Press.

Chierchia, Gennaro, Danny Fox & Benjamin Spector. 2012. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In Claudia Maienborn, Klaus von Heusinger & Paul Portner (eds.), *Semantics: An international handbook of natural language meaning volume 3*, Berlin: Mouton de Gruyter.

Clifton, Jr., Charles & Chad Dube. 2010. Embedded implicatures observed: a reply to geurts and pouscoulous (2009). *Semantics and Pragmatics* .

Crain, Stephen, Rosalind Thornton, Carole Boster, Laura Conway, Diane Lilo-Martin &

13 In particular, Russell (2012: p.58) discusses (18) as a possible *definition* of scalar implicatures, which is very close to our principle in (9). However, Russell considers (17) as a candidate definition of what scalar inferences are. For us, instead, the probabilistic constraint in (16) selects among possible scalar inferences independently generated by a theory of scalar implicatures.

(18) The utterance of a sentence S conversationally implicates I iff $P(I|S) > P(I)$.

Elaine Woodams. 1996. Quantification without qualification. *Language Acquisition* 5(2). 83–153.

Fox, Danny. 2007. Free choice and the theory of scalar implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presupposition and Implicature in Compositional Semantics*, 71–120. Palgrave.

Gotzner, Nicole & Jacopo Romoli. 2015. Testing theories of alterntives on sentenes with multiple scalar terms. Talk to be given at Xprag Chicago.

Gualmini, Andrea. 2004. *The ups and downs of child language*. New York: Routledge.

Gualmini, Andrea, Sarah Hulsey, Valentine Hacquard & Danny Fox. 2008. The Question-Answer Requirement for scope assignment. *Natural Language Semantics* 16. 205–237.

Horn, Lawrence. 1972. *On the semantic properties of logical operators in English*: UCLA dissertation.

Lewis, David. 1979. Scorekeeping in a language game. *Journal of Philosophical Logic* .

Magri, Giorgio. 2009. *A theory of individual-level predicates based on blind mandatory scalar implicatures*: Massachusetts Institute of Technology dissertation.

Mayr, Clemens & Jacopo Romoli. 2014. Redundancies and the notion of local contexts. Unpublished manuscript ZAS/University if Ulster.

Merin, Arthur. 1999. Information, relevance and social decisionmaking. Some principles and results of decision-theoretic semantics. *Logic, Language and Computation* .

Potts, Christopher, Daniel Lassiter, Roger Levy & Michael Frank. 2015. Embedded implicatures as pragmatic inferences under compositional lexical uncertainty. Manuscript Stanford.

Romoli, Jacopo. 2012. *Soft but strong: Neg-raising, soft triggers, and exhaustification*: Harvard University dissertation.

Russell, Benjamin. 2012. *Probablistic reasoning and the computation of scalar implicatures*: Brown University dissertation.

Sauerland, Uli. 2004. Scalar implicatures in complex sentences. *Linguistics and Philosophy* 27(3). 367–391.

Schlenker, Philippe. 2009. Local contexts. *Semantics and Pragmatics* 2(3). 1–78. http://dx.doi.org/10.3765/sp.2.3.

Snider, T. 2015. Using tautologies and contradictions. In E. Csipak & H. Zeijlstra (eds.), *Proceedings of sinn und bedeutung 19*, .

Spector, Benjamin. 2007. Aspects of the pragmatics of plural morphology: On higher-order implicatures. In Uli Sauerland & Penka Stateva (eds.), *Presupposition and implicature in compositional semantics*, Palgrave.

Trinh, Tue & Andreas Haida. 2015. Building alternatives. *Journal of Semantics* .

van Rooij, Robert & Katrin Schulz. 2004. Pragmatic meaning and non-monotonic reasoning: The case of exhaustive interpretation. *Linguistics and Philosophy* .