Implicatures and Discourse Structure

Nicholas Asher

Implicatures and discourse structure

Nicholas Asher

CNRS, Laboratoire IRIT UMR 5505, Univ. de Toulouse, 118 route de Narbonne, 31062 Toulouse Cedex 9, France

Abstract

One of the characteristic marks of Gricean implicatures in general, and scalar implicatures in particular, examples of which are given in (1), is that they are the result of a defeasible inference.

(1a) John had some of the cookies
(1b) John had some of the cookies. In fact he had them all.

(1a) invites the inference that John didn't have all the cookies, an inference that can be defeated by additional information, as in (1b). Scalar inferences like that in (1a) thus depend upon some sort of nonmonotonic reasoning over semantic contents. They share this characteristic of defeasiblility with inferences that result in the presence of discourse relations that link discourse segments together into a discourse structure for a coherent text or dialogue—call these inferences discourse or D inferences. I have studied these inferences about discourse structure, their effects on content and how they are computed in the theory known as Segmented Discourse Representation Theory or SDRT. In this paper I investigate how the tools used to infer discourse relations apply to what Griceans and others call scalar or quantity implicatures. The benefits of this investigation are three fold: at the theoretical level, we have a unified and relatively simple framework for computing defeasible inferences both of the quantity and discourse structure varieties; further, we can capture what's right about the intuitions of so called "localist" views about scalar implicatures; finally, this framework permits us to investigate how D-inferences and scalar inferences might interact, in particular how discourse structure might trigger scalar inferences, thus explaining the variability (Chemla, 2008) or even non-existence of embedded implicatures noted recently (e.g., Geurts and Pouscoulous, 2009), and their occasional noncancellability. The view of scalar inferences that emerges from this study is also rather different from the way both localists and Neo- Griceans conceive of them. Both localists and Neo-Griceans view implicatures as emerging from pragmatic reasoning processes that are strictly separated from the calculation of semantic values; where they differ is at what level the pragmatic implicatures are calculated. Localists take them to be calculated in parallel with semantic composition, whereas Neo-Griceans take them to have as input the complete semantic content of the assertion. My view is that scalar inferences depend on discourse structure and large view of semantic content in which semantics and pragmatics interact in a complex way to produce an interpretation of an utterance or a discourse.

Keywords: Embedded implicatures; Discourse structure; Defeasible reasoning

1. Introduction

A characteristic mark of Gricean implicatures in general, and scalar or quantity implicatures in particular is that they are the result of a defeasible inference.[1] (1) illustrates an example of a scalar implicature and its defeasibility.

(1) a. John had some of the cookies
     b. John had some of the cookies. In fact he had them all.

---

[1] An earlier version of this paper will appear in Asher (in press).

(1a) invites the inference that John didn't have all the cookies, an inference that can be defeated by additional information, as in (1b). Scalar inferences like that in (1a) thus depend upon some sort of nonmonotonic reasoning over semantic contents. They share this characteristic of defeasiblility with inferences that result in the presence of discourse relations that link discourse segments together into a discourse structure for a coherent text or dialogue—call these inferences *discourse* or D-inferences. In developing the theory known as *Segmented Discourse Representation Theory* or SDRT, I have studied the effects of D-inferences on content and how they are computed. In this paper I apply the tools used to infer discourse relations to the derivation of quantity implicatures This yields a three advantages: at the theoretical level, we have a unified and relatively simple framework for computing defeasible inferences both of the quantity and discourse structure varieties; we can capture what's right about the intuitions of so called "localist" views about scalar implicatures; finally, this framework permits us to demonstrate how D-inferences and scalar inferences interact, in particular how discourse structure trigger scalar inferences. In turn this leads to a natural explanation of the variability (Chemla, 2008) or even non-existence of embedded implicatures noted recently (e.g., Geurts and Pouscoulous, 2009), as well as their occasional non-cancellability. The view of scalar inferences that emerges from this study is quite different from the way both localists and Neo-Griceans conceive of them. Both localists and Neo-Griceans view implicatures as emerging from pragmatic reasoning processes that are strictly separated from the calculation of semantic values; where they differ is at what level the pragmatic implicatures are calculated. Localists take them to be calculated in parallel with semantic composition, whereas Neo-Griceans take them to have as input the complete semantic content of the assertion. My view is that scalar inferences depend on discourse structure and an inclusive notion of semantic content; semantics and pragmatics interact in a complex way, sub sententially and supra-sententially, to produce an interpretation of an utterance or a discourse in which both scalar or quantity implicatures and discourse structure.

2. D-inferences

Scalar inferences are familiar to most linguists who work on pragmatics, but D-inferences are less so. Let us look at some examples of them in texts.

(2) a.   John walked in. / He poured himself a cup of coffee.
     b.   John fell. / Mary pushed him.
     c.   We bought the apartment, /but we've rented it.
     d.   Il commence à dessiner et peindre en 1943, / fréquente les ateliers de sculpture / puis de peinture de l' école des Beaux-Arts d' Oran, / où il rencontre Guermaz (ANNODIS corpus).
     e.   Julie had an excellent meal, / beginning with an elegant and inventive truffes du Périgord en première cuisson comme un petit déjeuner, / followed by some wonderful scallops, /then sweetbreads, /a sumptuous cheese plate, /and ending with a scrumptious dessert.

A presumption of relevance leads us to infer some link between *elementary discourse units* or *EDU*s (clauses or subclausal units whose boundaries are either sentence initial or marked by / in the examples above). These links involve relations that are familiar even to the non-linguist: some units elaborate or go into more detail concerning something introduced in another constituents (these are Elaboration type relations) as in (2e); some units form a parallel or a contrast with other units (such units are linked by Parallel or Contrast), as in (2c); some units furnish explanations why something described in another unit happened (Explanation) as in (2b); and some units constitute a narrative sequence of events (Narration) (2a) or (2d). Other discourse relations of interest for our purposes are indirect question answer pairs (IQAP), which link responses to a prior question, Correction, where a second discourse move revises the content of a first, and Alternation, which is linked to certain uses of disjunction.

Some D-inferences are encoded grammatically through the use of certain grammatical constructions (like adverbial or purposive clauses, or left fronted temporal or spatial adverbials)[2] or through discourse connectors like *as a result, puis* or the choice and sequencing of lexical items. An example of a set of discourse relations triggered by the choice of verb and complement comes in (2e), with the use of *beginning with*, *followed by* and *ending with*. Sometimes, it is less clear what linguistic source triggers the inference of the discourse relation as in (2a–b)—most likely, an as yet not fully understood mix of lexical semantics and world knowledge. The discourse relations implicated by these devices have impose structural constraints on the discourse context and have truth conditional effects that a number of researchers have explored.[3]

To see the defeasibility of D-inferences in context, consider this variant of (2b)

(3)   John fell. Mary pushed him. He rolled off the edge of the cliff.

---

[2] For a discussion of these, see for instance, Vieu et al. (2005).

[3] With regards to temporal structure, see Lascarides and Asher (1993); on pronominal anaphora, see Asher (1993), Asher and Lascarides (1998), Kehler et al. (2008), on sluicing and ellipsis see Asher (1993), Hardt et al. (2001), Romero and Hardt (2004).

Although (2b) is part of the discourse given in (3), the inference to Explanation between the two clauses of (2) is not so readily available in (3). Rather, we tend to read the first two clauses in (3) as providing a narrative sequence that results in his rolling off the cliff.

While scalar inferences occur embedded under quantifiers and other operators, many people have noted that these implicatures are less robust than unembedded cases of scalar implicatures like (1a). D-inferences, on the other hand, robustly embed under quantifiers and other operators (as well as other discourse relations).

(4) a. If it was late, John took off his shoes and went to bed.
     b. If it was late, John went to bed and took off his shoes.
     c. If John drank and drove, he put his passengers in danger.
     d. The CEO of Widgets &Co. doubts that the company will make a profit this year and that (as a result) there will be much in the way of dividends for shareholders this year.

In both (4a–b), the D-implicature that there is a narrative sequence between the two clauses in the consequent of the conditional survives under embedding,and (4c) shows that this holds in the antecedent of a conditional as well. (4d) shows that the causal relation of result holds when embedded under a downward entailing attitude verb.

So D-inferences and S-implicatures share at least one feature—defeasibility. Below, I sketch the account of how D-inferences are inferred. I then try to answer two questions: how does the mechanism for inferring D-inferences apply to S-implicatures? How do D-inferences and S-implicatures interact? I argue that the mechanism for inferring D-inferences readily adapts to the computation of S-implicatures, regardless of what exact view of S-implicatures one adopts. I also show that there is a tight interaction between D-inferences and S-implicatures. I show how to derive S-implicatures from D-inferences and I show why sometimes the S-implicatures aren't defeasible. The mechanism for deriving D-inferences shows that all S-implicatures depend on discourse structure, which may be implicit or inferred in out of the blue contexts. Thus, I offer a way of understanding S-implicatures that is different both from the localist one (e.g., Chierchia, 2004, *inter alia*) and a Gricean one (e.g., van Rooij and Schulz, 2004, *inter alia*).

3. SDRT, a theory about D-inferences

To address the questions I have set myself, I must say more about D-inferences and the theory for investigating them. A theory of D-inferences must provide a logic or means for computing these inferences and detail how they interact with semantic content. While there are several theories of discourse structure that researchers in linguistics and computer science investigate, SDRT is an appropriate theory to use here because it spells out a very clear picture of the structure, the construction and the *semantics* of discourse structures with an emphasis on the computation of D-inferences and their interaction with semantic content. The theory answers to three tasks:

- it segments a text into EDUs;
- it computes attachment points of EDUs in a discourse structure;
- it computes one or more discourse relations between an EDU and its attachment point(s).

EDUs are discourse units containing elementary predications involving some sort of eventuality (event or state). All clauses give rise to EDUs, but appositives, parentheticals, non-restrictive relative clauses, and adverbials that are detached to the left of the main syntactic structure of a clause also introduce EDUs. Coordinated verb phrases that use recognized discourse connectors like *but* in *John went to the store but didn't get any milk* also give rise to two EDUs. The other element involved in the tasks above that is perhaps unfamiliar to linguists who work primarily on sentential semantics and syntax are discourse relations. In general all researchers working on discourse agree that there are relations that are causal, thematic (e.g., elaboration or background) and narrative. The philosophical background for this work goes back to Hume's taxonomy of ideas and to Kant's categories of relation. Within these general categories, researchers and different theories of discourse structure differ as to the number of finer-grained relations. SDRT defines relations as distinct just in case they make a difference to the content of the discourse (but not distinguish relations based on, e.g., speaker intentions).

An SDRT discourse structure or SDRS is the result of the computations outlined above. It may contain complex constituents where several EDUs combine together to make one larger constituent. An SDRS is a logical form for discourse with a well-defined dynamic semantics that has many equivalent formulations—as a first order model like structure consisting of a set of labels and assignments of formulas to labels (Asher and Lascarides, 2003), as a DRS like structure (Asher, 1993) or as a λ term in intensional logic (Asher and Pogodalla, 2010).

To get an idea of what SDRSs look like consider the following text (5) discussed at length in (Asher and Lascarides, 2003). The model-like SDRS is given in (5′)

(5) π_1 John had a great evening last night.
       π_2 He had a great meal.
       π_3 He ate salmon.
       π_4 He devoured lots of cheese.
       π_5 He then won a dancing competition.

(5′) $\langle A, \mathcal{F}, Last \rangle$, where:

       $A = \{\pi_0, \pi_1, \pi_2, \pi_3, \pi_4, \pi_5, \pi_6, \pi_7\}$

       That is, in addition to the EDUs $\pi_1, \ldots, \pi_5$, we have the complex constituents π_0, π_6 and π_7.

       $\mathcal{F}(\pi_0) = Elaboration(\pi_1, \pi_6)$
       $\mathcal{F}(\pi_6) = Narration(\pi_2, \pi_5) \wedge Elaboration(\pi_2, \pi_7)$
       $\mathcal{F}(\pi_7) = Narration(\pi_3, \pi_4)$

In SDRT we abstract away from the details of the structure to get a graph representation relevant to computing discourse accessibility for anaphoric antecedents and sites for presupposition accommodation or binding (again for details see Asher and Lascarides, 2003 and Asher, 2008).

3.1. Inferring D-inferences

Inferring D-inferences is a matter of defeasible and uncertain inference. Many of the features used to infer discourse relations are only good indications of a particular discourse relation or particular discourse structure; very few are in and of themselves sufficient to deductively infer the relation or structure. Many discourse connectives are for example ambiguous. In addition, many segments may bear discourse relations to other segments despite the lack of discourse connectives or known structural or lexical cues, as in (2a,b) or (5). To solve this problem, my colleagues and I developed a nonmonotonic logic, a logic for defeasible inference, tailored to inferring D-inferences.

The task of building such a logic is not completely trivial. Integrating nonmonotonicity in discourse interpretation is problematic, especially if this integration occurs at the level of contents or what is said. Reasoning over contents non-monotonically requires finding a class P of preferred models, those with the intended discourse relations, and computing validity or logical consequence with respect to P. Given that the language of information content is at least that of first order logic, where the complexity of the computation of validity and logical consequence is only recursively enumerable and that almost all nonmonotonic logics require some sort of consistency test over the formulas one is using for the inference, the complexity of computing logical consequence with respect to the class P of preferred models is not recursively enumerable—i.e., computationally hopeless. This is not just a matter of implementation but one of principle. We cannot assume that agents, with their limited computational capacities, are able to solve a problem reliably which we can show mathematically to be incapable of having anything like what we would call an algorithm. Attributing such computational capacities to agents shows that we have mischaracterized the problem: they are not computing logical consequence over formulas of information content; either the language in which the computation is done is somehow a simplification of the language of information content, or they are computing something other than logical consequence, perhaps using some sort of heuristic.

One could develop a heuristic for computing D-inferences. In fact, Fox (2007) as well as others have proceeded to do this for scalar and free choice implicatures. The problem with taking this tack, however, lies in its verification. In order to make sure that the heuristic is doing what it is supposed to, we must check it in the way that computer programs are checked, via program verification. In most if not all instances, this means translating the problem into a logic and then checking that the result desired is in fact a logical consequence of the program's translation and the input data. This leads us back to the task of building a nonmonotonic logic for D-inferences.

SDRT's solution to this problem is to look at non-monotonic reasoning not over contents but over logical forms. Roughly, instead of trying to compute the nonmonotonic consequences of a bunch of facts about the world, facts which may be quantificationally complex, we try to compute the nonmonotonic consequences of a discourse logical form's having a certain shape and of a segment's having the lexical choices and structure that it does. This means that we are trying to solve a logical consequence problem, not in the language of information content, but in a language for describing discourse logical forms and features of discourse constituents. Asher and Lascarides (2003) develop such a language, which they call the *glue language*. The nonmonotonic logic adapted to this language is known as the *glue logic* or *GL*. Asher and Lascarides (2003) show that the problem of logical consequence for formulas of this language is in fact decidable.

GL uses axioms exploiting various resources to get the intended discourse relations to hold between discourse constituents. The general form of such axioms is this:

- General Form: $(? (\alpha, \beta, \lambda) \wedge \text{some stuff}) > R(\alpha, \beta, \lambda)$

In this general form, '?' represents an underspecified relation linking β to λ within the constituent α, while $>$ is a weak conditional whose semantics supplies the non monotonic consequence relation; "some stuff" is information about α, β and λ that's transferred into the glue language from more expressive languages for other information sources like: compositional semantics, lexical semantics, pragmatic maxims of conversation, generalizations about agent behaviour in conversation, and domain knowledge. To the right hand side of the $>$ is a formula that specifies the underspecified discourse relation to the left.

The main item of interest in this general form is the semantics of $>$ and the notion of defeasible consequence that issues from it. The semantics of $>$ was developed by Asher and Morreau (1991) in a first order non-monotonic logic known as *common sense entailment*. This is a logic for nonmonotonic or defeasible reasoning based on a weak conditional $>$. Originally devised to treat generics, I have used a version of it restricted to a quantifier free description language, the glue language, now for many years to calculate D-inferences, and it is a relatively adaptable non monotonic logic. It has two parts: a basic, monotonic, conditional logic with a standard proof theory $\vdash$ and consequence relation $\vDash$, and then a defeasible inference relation $\vdash\!\!\sim$ and a nonmonotonic consequence relation $\vDash\!\!\sim$ that make use of the basic logic. I will use the glue logic version of common sense entailment to model both D- and S-implicatures.

Let me briefly recapitulate the basics of common sense entailment restricted to a propositional language.

- A modal generic frame $\mathfrak{F} = \langle W, * \rangle$ where W is a non empty set of worlds and $* : W \times \mathfrak{P}(W) \to \mathfrak{P}(W)$ is a selection function. A model $\mathfrak{A}$ is constructed by adding a valuation function.
- $\mathfrak{A}, w \vDash A > B$ iff $*(w, \|A\|) \subseteq \|B\|$
- the standard clauses for the quantifiers and connectives, though the glue language itself has only a quantifier free logical structure.

This semantics for $>$ will be familiar to those accustomed to conditional logics; it comes with a standard monotonic validity predicate and complete axiomatization.[4]

---

[4] It validates the following axioms and rules, some of which depend on frame constraints that I give below.
- classical tautologies
- closure on the right

$$((\phi > \psi) \wedge (\phi > \chi)) \to (\phi > (\psi \wedge \chi))$$

- right weakening:

$$\frac{\vdash \phi \to \psi}{\chi > \phi \vdash \chi > \psi}$$

- idempotence:
follows with the frame constraint $*(w, p) \subseteq p$
- supraclassicality:

$$\frac{\vdash A \to B}{\vdash A > B}$$

- substitution of logically equivalent formulae in the antecedent of a $>$ formula.

$$\frac{\vdash A \leftrightarrow B}{\vdash (A > C) \leftrightarrow (B > C)}$$

- the "or" principle: $*(w, p \cup q) \subseteq *(w, p) \cup *(w, q)$

$$A > B, C > B \vdash (A \vee C) > B$$

- the "specificity" principle: $(p \subseteq q \wedge *(w, p) \cap *(w, q) = 0) \to *(w, q) \cap p = 0$

$$\frac{\vdash A \to B}{\vdash (A > \neg C \wedge B > C) \to (B > \neg A)}$$

How do we pass from a notion of monotonic consequence to a nonmonotonic one? The idea of the nonmonotonic consequence relation is to assume that matters are as normal as "possible" given the information in the premises and to then see what follows. Assuming matters are as normal as possible means making > as much like the material conditional as "possible"—moving from *if p then normally q* to *if p then q*. There is both a "proof theoretic method" for defining nonmonotonic consequence ($\vdash$) and a model theoretic method ($\models$) for which a correspondence theorem is given in Asher (1995).[5] For the proof theoretic method, we define first an $A\rightarrow$ extension of Γ, $\Gamma^{A\rightarrow}$:

- $\Gamma^{A\rightarrow} = \Gamma \cup \{A \rightarrow \phi : \Gamma \vdash A > \phi\}$, if consistent.
- $\quad\quad = \Gamma$, if not

For each antecedent A of a > conditional derivable from Γ, define an $A\rightarrow$ extension of Γ inductively relative to an ordering ρ over antecedents of > statements with $\Gamma \rightarrow_{\rho,0} = \Gamma$. Every such sequence has a fixed point. We now have the definition:

Definition 1. $\Gamma \vdash \phi$ iff for all orderings ρ the fixed point $\Gamma*$ of each $\Gamma \rightarrow_\rho$ extension sequence is such that $\Gamma* \vdash \phi$.

4. Gricean scalar implicature

With this sketch of commonsense entailment, let us now turn to Grice's picture about implicatures. Grice's view of implicatures is that they are calculated after compositional semantics has finished its job via his famous general maxims of conversation, quality, quantity and relevance. In principle the Gricean picture tells the beginning of an attractive story for computing scalar implicatures (Horn, 1972, 2005; Schulz and van Rooij, 2006; Schulz, 2007; Spector, 2006). But this is only the beginning of a story, since Grice's maxims of quality and quantity are not precise enough really to license any inferences. In addition, these maxims should clearly be formalized within a nonmonotonic logic. GL provides a simple yet precise reconstruction.

Consider (6) and its implicature signaled by $\leadsto$:

(6)   John or Susan came to the party $(j \vee s) \leadsto \neg(j \wedge s)$.

The weak Gricean implicature that the speaker does not believe that $j \wedge s$ follows in GL if we incorporate the defeasible principle of *Sincerity*—Say$\phi > \mathcal{B}\phi$—together with a principle of ignorance about more informative alternatives to what was said. Since GL works with descriptions of logical forms, we have the means to write down the fact that a certain relation holds between two formulas, namely that ψ is an alternative that could have been said in the given discourse context instead of ϕ and that it is an alternative. Being able to express and to define this set of alternatives is crucial to the enterprise of formalizing S-implicatures. To pick the right alternatives, we need a measure of informativeness. Let's suppose it's logical entailment that is provided from the background logic of information content (it is not stateable in GL itself, since its consequence relation is much weaker than that of the background logic). I will define Alt(ϕ, ψ) iff ψ is a strictly more informative formula that is an alternative to ϕ. I assume that the set of relevant alternatives is finite and abbreviate the fact that Alt(ϕ, ψ_1) $\wedge \ldots \wedge$ Alt(ϕ, ψ_n) as Alt(ϕ, Ψ), the relevant axiom to draw ignorance implicatures is now:

(7) $\text{Alt}(\phi, \Psi) \rightarrow ((\text{Say}\phi \wedge \neg\text{Say}\psi_1 \wedge \ldots \wedge \neg\text{Say}\psi_n) > (\mathcal{B}\phi \wedge \neg\mathcal{B}\psi_1 \wedge \ldots \wedge \neg\mathcal{B}\psi_n))$

Provided Alt($j \vee s$) = $\{j, s, j \wedge s\}$, we get ignorance implicatures from (6) that the speaker doesn't believe that s, j or that $j \wedge s$, by simply turning the relevant > instance of the schema (7) into a $\rightarrow$ statement. Using (7) alone we have only one fixed point in which the speaker believes $j \vee s$ and that she believes neither j, s nor $j \wedge s$.

To get the stronger Gricean implicature, that the speaker believes $\neg(j \wedge s)$, I adopt the following axiom:

(8) $\text{Alt}(\phi, \psi) \rightarrow (\neg\text{Say}\psi > \mathcal{B}\neg\psi)$

---

[5] Define a *normalization* of Γ relative to A for a model $\mathfrak{A}$

- $\|\Gamma\|_{\mathcal{N},A}^{\mathfrak{A}} = \cap((W - \|A\|^{\mathfrak{A}}) \cup \bigcup_{w \in \|\Gamma\|^{\mathfrak{A}} *}(w, \|A\|))$ if this is non 0.
- $= \|\Gamma\|^{\mathfrak{A}}$, otherwise.

For each antecedent A of a > conditional derivable from Γ, define inductively a Γ normalization sequence relative a model $\mathfrak{A}$ and an ordering ρ over the antecedents of > statements. $\Gamma \models \phi$ iff for every ordering ρ the fixpoints of the Γ normalization sequence relative to ρ and the canonical model provided by completeness verify ϕ.

The form in (7) has a more specific antecedent than (8) and so will override (8)'s incompatible consequences. Let's see what happens when we compute the fixed point of (7) and (8). We have the atomic facts:

(9) $\text{Say}(j \lor s), \neg\, \text{Say}\, j, \text{Alt}(j \lor s, \{j, s, j \land s\}), \neg\, \text{Say}s, \neg\, \text{Say}(j \land s)$

All fixed points of the set of premises contain the$\rightarrow$ transformations of the relevant instance of (7):

(10) $(\text{Say}(j \lor s) \land \text{Alt}(j \lor s, \{j, s, j \land s\}) \land \neg\, \text{Say}(j \land s)) \rightarrow (\mathcal{B}(j \lor s) \land \neg\mathcal{B}(j \land s) \land \neg\mathcal{B}j \land \neg\mathcal{B}s)$

(7)'s antecedent entails (8)'s antecedent, so in cases of conflicting defeasible consequences where Alt(ϕ, ψ), the specificity principle entails:

(11) $\neg\text{Say}\psi > \neg\, (\text{Say}\phi \land \neg\, \text{Say}\psi_1 \land \ldots \land \neg\, \text{Say}\psi_n)$

From the facts at hand, we see that (8) and (7) conflict when ψ in (8) is j or s (but not $j \land s$), because, for instance, instantiating ψ in (8) to j we would have $\mathcal{B}\neg j$. But since $\mathcal{B}$ is closed under logical consequence and $\mathcal{B}(j \lor s)$ follows from (7), we would get $\mathcal{B}s$, which contradicts another consequence of (7). This means that Specificity adds the following additional premises in the derivation:

(12) $\neg\text{Say}j > \neg\, (\text{Say}(j \lor s) \land \neg\, \text{Say}j \land \neg\, \text{Say}s \land \neg\, \text{Say}(j \land s));$
(13) $\neg\text{Say}s > \neg\, (\text{Say}(j \lor s) \land \neg\, \text{Say}j \land \neg\, \text{Say}s \land \neg\, \text{Say}(j \land s));$

Once we add (12) and (13) to our premises, attempting to transform the instances of (8) involving j and s into $\rightarrow$ statements fails, because in transforming the relevant instances of (8), we must turn all > statements involving the antecedents $\neg\text{Say}j$ and $\neg\text{Say}s$ into $\rightarrow$ formulas. And this renders our set of formulas inconsistent. On the other hand, there is no inconsistency between (7) and the instance of (8) using $j \land s$; the defeasible consequence of (8) in that case is $\mathcal{B}(\neg j \lor \neg s)$. So we get in the final fixed point for the computation using (7) and (8) the desired implicatures:[6]

(14) $\mathcal{B}(j \lor s) \land \neg\mathcal{B}(j \land s) \land \mathcal{B}\neg(j \land s)$

Like many other derivations of scalar implicatures for (6), the GL derivation is valid only if the alternativeness relation Alt is restricted as above. If the set of alternatives contains other strictly stronger logical entailments, then GL predicts no implicature can be drawn.[7]

The Gricean programme itself says nothing about relevant sets of alternatives (Block, 2009). We would need an additional component of the theory to add this. Alternatives might be specified lexically and be part of the grammar (see Fox and Katzir, 2011). For example, we might say that alternatives are defined lexically for scalar items, for adjectives *beautiful, stupendous, gorgeous, . . .*, for nouns *genius, idiot, . . .*, for connectives ($\lor$, $\land$) (here it depends as to whether >, $\rightarrow$, etc. are also considered connectives), and for quantifiers (*no, some, many, most, all*). Such lexically specified alternatives seem reasonable but we need to know how alternatives compose together. One might also try to fix the alternatives for an utterance via a question under discussion approach, but this has many difficulties of its own, not the least of which is determining what the question under discussion is.[8] Another approach is to reason directly over preferred

---

[6] In particular Sauerland's (2004) prioritization of defaults (his Expertness assumption) is not needed; the underlying logic and the Specificity Principle of common sense entailment handles this. This is desirable, since this allows for a straightforward cancellation of the implicatures if the explicit semantic content contains for example an explicit denial of one of the implicatures. Further, note that because GL calculates defeasible implicatures over a language describing logical forms in which quantifiers are eliminable, the process is decidable, in fact PSPACE.

[7] This is a formalization of the so called *symmetry problem* (Kroch, 1972).

[8] One option is to take the question under discussion (QUD) to be the polar question induced from the utterance, which for which (6) is:

(15)   Did John or Susan go to the party?

Unfortunately, the answers to this question doesn't specify the right alternatives. Alternatively we could take the QUD to be

(16)   Who went to the party?

But this now has too many partial or linguistic answers like $(s \lor j) \land \neg (j \land s)$.

models of information content (van Rooij and Schulz, 2004; Schulz and van Rooij, 2006), but this generates another set of problems, the problem of side effects, which leads again to an introduction of relevant alternatives.[9]

5. Embedded implicatures and the localist challenge

So far, I have detailed a general logic, GL, for computing D-inferences that also serves to capture S-inferences given a set of alternatives. We need only axioms (7) and (8), together with a specification of alternatives.[10] So what about the set of alternatives? It seems reasonable that the set of alternatives is at least partially conventionally determined. But at what level are these alternatives computed? Griceans claim that the alternatives are computed pragmatically, after truth conditional semantics has finished its job. Given a standard view of what truth conditions are in semantics (sets of possible worlds), this means that Griceans do not have access to the "fine structure" of a sentence's meaning; the particular way the truth conditions have been derived from lexical meanings has been erased in the semantic value. A contrasting, "localist" theory claims that implicatures are, like presuppositions in the theory of Karttunen and Peters (1979), conventionally determined by the lexicon and computed in parallel to compositional semantic interpretation. In contrast to Griceans, localists like (Chierchia, 2004) have access to the fine structure of meaning in computing implicatures, in particular in computing the set of alternatives upon which the derivation of implicatures depends.

The principal motivation for the localist approach is the presence of embedded implicatures, which can present problems for Neo-Gricean approaches. An embedded implicature is an implicature that is generated by an expression within the scope of a truth functional, modal operator or quantifier.[11]

(18)   John did the reading or some of the homework

This implicates John didn't do the reading and some of the homework. It also implicates that he didn't do all of the homework. Since Griceans compute implicatures only on whole utterances or full sentences, it's not clear how to get the second implicature, which is based on the meaning of a subsentential component of (18). (18) and similar examples have impressed some linguists as a decisive argument against a Gricean view (for example Fox, 2007). Nevertheless, the difficulty for Griceans with (18) depends once again on the set of alternatives chosen (Spector, 2007). In Asher (in press) I show how to adapt the basic rules (7) and (8) to exploit the recursive structure of the logical form and lexically based scales for scalar predicates, disjunction and determiners (the latter has been acknowledge to be a source of implicatures since the early work of Horn, 1972). By reasoning over logical forms GL has access to the alternatives of subsentential constituents and thus allows us derive implicatures of the sort that Chierchia claims to hold, if we want them, all within a broadly Gricean framework.

6. Interactions between S-implicatures and discourse context

The localist vision, however, is as flawed as the Neo-Griean one. They both overlook the effects of discourse context. The heart of the localist position is that implicatures are calculated on the basis of lexical semantics and syntactic structure. I show in this section that so called embedded S-implicatures are sensitive to discourse structure; in fact they are triggered by it. Sometimes these embedded implicatures, when necessary for discourse coherence, lose their characteristic as a defeasible implicature. They are, if you will, still pragmatic inferences, insofar as discourse structure is a matter of pragmatics as well as semantics. But these uncancellable "S-inferences" bear little resemblance to their intuitive characterization as implicatures given at the beginning of this paper. First, I want to examine how S-inferences and discourse structure interact. I will then propose an account of that interaction that makes S-inferences parasitic or at least codependent upon discourse structure. This account will open the way to a more radical position: even the standard examples of S-implicatures depend on assumptions about implicit discourse context, and in effect both localists and Griceans have gotten it wrong about how such inferences are to be computed.

---

[9] In Asher (in press) I look at this solution in more detail. To see the problem of side effects consider

(17)   a.   (Assume a problem set with 10 problems on it.) How many problems did John do on the problem set?
         b.   John did some of the problem set $\leadsto \neg \mathcal{B}$ (John did more than 1 problem of the problem set).

But (17b) does not have any such implicature.

[10] Fox (2007), Kratzer and Shimoyama (2002) and Alfonso-Ovalle (2005) argue that free choice implcatures should be treated with the mechanism for scalar implicatures. I have a rather different take on free choice implicatures, but that would take us too far afield here. See Asher and Bonevac (2005).

[11] This means of course that implicatures may be embedded even when there is no syntactically embedded sentence.

Sometimes the discourse context can license S-inferences that aren't normally considered by either localists or Griceans. Consider

(19)   a.   A: Do you like him?
        b.   B: Why, yes, I do.
        c.   A: I mean, do you like him or do you LIKE him like him?
        d.   B: I like him.

In the first question answer pair with a yes/no question, the lexical scale associated with *like* isn't really operative or needed to understand the exchange to an alternative question where a scale is explicitly invoked, and where the response implies the negation of the stronger emotion (liking him liking him is stronger than liking him).[12] What is going on in these examples? While there are, most likely, scales lexically associated with adjectives like *full*, *bald*, etc. and determiners like *some*, *most*, etc., the actual values and perhaps even the presence or the activation of the scale for the purposes of calculating implicatures is dependent on discourse context. It is the discourse context that suggests the relevant alternatives, and it is the requirement of computing a particular discourse relation that triggers the S inference.

There are in fact many effects of discourse structure on S-inferences. A localist computation of alternatives predicts that one should not get the standard implicatures (the exhaustivity implicature for *or*, for instance) inside the scope of downward entailing operators.[13] Most Gricean accounts would agree. But this is wrong for the antecedents of conditionals in some cases, which are downward entailing. Consider (20), and its paraphrase (21):

(20)   If you take cheese or dessert, you pay 20 euros; but if you take both there is a surcharge.
(21)   If you take only a cheese dish or only a dessert, the menu is 20 euros; but if you take both, there is a surcharge.

or

(22)   If John owns two cars, then the third one outside his house must be his girlfriend's.
(23)   If one person reads my book, I'll be happy.

Here the interpretation of the consequent forces the antecedent to have the exclusive reading. This is a further indication that the calculation of the alternative set relevant to deriving S-implicatures cannot be purely localist but depends rather on the global discourse context in which it is embedded.

The dependence of implicatures on discourse contexts also shows us that we must take care how to calculate them. Both Griceans and localists consider implicatures to be calculated after semantics. An implication of this view (made explicit in Chierchia, 2004) is that implicatures that aren't inconsistent with established facts in the common ground or the narrow semantic content continue to be operative as the discourse content is computed. This insensitivity of implicatures to a more nuanced view of discourse content gives us the wrong results. Consider the example (24a) with its embedded implicature in (24b)

(24)   a.   John either did some of the reading or some of the homework $\rightsquigarrow$
        b.   John didn't do all of the reading; John didn't do all of the homework; and he didn't do some of the homework and some of the reading.

And now consider the following dialogue.

(25)   a.   A: John either did some of the reading or some of the homework.
        b.   B: John did all of the reading, but you're right, he didn't do all of the homework.

What happens to the implicature (24b) in this discourse context? To my ears, (25a–b) has the implicature that John did all of the reading and some of the homework. In fact this is an implicature of (25b). But this implicature cancels the implicature of (25a) that John didn't do some of the homework and some of the reading. Thus, implicatures interact in important ways with discourse moves like the corrective move given by B. We need to integrate S-implicatures into discourse content and

---

[12] This exchange invites an extended evaluation of how views on contrastive topic impinge on this enterprise (Lee, 2010a,b). For lack of space, I cannot do justice to the intriguing parallels between the observations here and that work.
[13] A downward entailing operator O is one where if ϕ implies ψ, then $O(\psi)$ implies $O(\phi)$.

structure and then compute the appropriate update of that content after taking into account B's corrective move. This implies that we cannot calculate implicatures merely at a sentential level, as most Griceans and localists have assumed. We need a much more finegrained view of the discourse context to calculate implicatures properly.[14]

Another problem with the current (localist and Gricean) accounts of S-inferences is that these inferences are not always cancellable. In particular, when S-inferences are required for discourse coherence, they are not cancellable.

(27) a. John has an even number of children. He has four. (implicature is that he has exactly 4)
      b. #John has an even number of children (π_1). He has three (children) (π_2).
      c. John has an even number of children (π_1'), and he has at least three (children) (π_2').

A Gricean or a localist should predict that (27b) is OK, since the implicature to the stronger, "exactly" meaning of *three* should be blocked. However, it is not, and (27b) is infelicitous. I hypothesize this stems again from an interaction of discourse structure and implicatures: there is a particular sort of elaborative move going on in the second clauses of (27a–b), which accounts for the "freezing of the implicature".

A final indication that something is amiss with current accounts of S-inferences is their well-known fragility—i.e., their dependence on contextual effects and on their logical form. As Chemla (2008) notes, localist theories predict that (28b) should not have the implicature below, making a stark contrast between (28a) and (28b).

(28) a. John didn't read all of the books. $\rightsquigarrow$ John read some of the books.
      b. No student read all of the books. $?\rightsquigarrow$ All of the students read some of the books.

For localists, the predicted implicature of (28b) is $\neg$ No student read some book—or *some students read some of the books*, which is weaker than the implicature tested by Chemla. However, (28b) is equivalent to:

(28c) All the students didn't read all of the books.

and this intuitively (and on a localist theory) implicates that all the students read some of the books. So equivalent meanings seem to yield distinct implicatures!

Interestingly, D-inferences are not closed under arbitrary first order equivalences either. Consider the logical equivalence in (29a). If D-inferences were computed on deep semantic content and hence closed under first order equivalences, we would predict no difference between (29b) and (29c) since (29b), where the relation of Explanation linking the two clauses is inferred, is perfectly coherent in contrast to (29c), where no discourse relation is inferred:

(29) a. Some one pushed him $\leftrightarrow_{CL}$ Not everyone didn't push him.
      b. John fell. Someone pushed him (Explanation inferred)
      c. # John fell. Not everyone didn't push him

Few people think that D-inferences are compositionally determined at the syntax semantics interface so this failure of substitutivity, which is equivalent to a failure of compositionality, is not much of a surprise. In GL this failure comes about because D-inferences are dependent on information about logical form and about the global discourse context, rather than just semantic content.

6.1. Combining D-inferences and S-implicatures

I have given arguments that call into question the empirical adequacy both of Gricean and localist accounts of implicature. Such accounts don't explain the fragility of S-inferences, their occasional uncancelability, and their sensitivity to discourse context. My simple, alternative hypothesis is that S-inferences not only use the general logic underlying GL's computation of D-inferences, but, like D-inferences, are also dependent the structure of the discourse context.[15] Making S-implicatures

---

[14] We can continue this pattern with more complex embedded examples.

(26) a. A: Some of the students did some of the reading or some of the homework.
      b. B: At least one student did all of the reading, but otherwise you're right.

It would seem that B's correction still leaves many of the implicatures of his original statement intact; he's still committed to the implicature that some of the students didn't do all of the homework and that some of the students didn't do all of the reading and some of the homework. We need a more finegrained notion of implicature revision in the face of corrections.

[15] See Geurts (2009) for a similar point.

dependent on discourse context should not come as a surprise. The behavior of other sorts of non-assertoric content, like presupposed content, also has complex interactions with discourse structure (Asher and Lascarides, 1998). As over 30 years of work on presupposition has shown, it is unwise to try to compute presuppositions without examining how the surrounding discourse context might affect these presuppositions. But unlike presuppositions, a theory of S-implicatures needs, not a theory of accommodation or of binding, but a theory of *triggering*. More specifically, we must answer the question: when does a discourse context license or induce an appropriate alternative set over which to compute (scalar) inferences?

Here in a nutshell is the approach: discourse relations like Contrast, Correction, Parallel, QAP, and various species of Elaboration induce structural constraints on the discourse in the form of structure preserving maps between constituents. Together with prosodic stress, which typically signals a lexical choice, these structural maps provide a set of alternatives relevant to calculating an S inference. When the requisite elements for constructing alternatives are not present, the alternatives aren't generated and neither is the inference. But there is also a co-dependence between S-inferences and discourse structure. Sometimes, inferred S-inferences are required to establish discourse relations. When the latter is the case, then I predict that the S-inferences are not cancellable without affecting the coherence of the discourse. I will thus be able to account for the discourse sensitivity of S-inferences, their fragility and their occasional uncancelability.

To build a case for my claim, I will look at examples that localists like Chierchia et al. (2008) have put forward to argue for the robustness of localist implicature computations. I will show that in all of those examples, it is the discourse structure that triggers the embedded S-inference, and in some cases the S-inference is required to maintain the discourse relation established. In the latter cases, the S-inferences are not cancellable except on pain of discourse incoherence.

Let's first take a look at the large collection of examples in Chierchia et al. (2008) all of which involve the discourse relation of Correction.

(30) a. Joe didn't see Mary or Sue; he saw BOTH. (only a clear exhaustive interpretation of the embedded disjunction is possible).
      b. It is not just that you CAN write a reply. You MUST.
      c. I don't expect that some students will do well, I expect that ALL students will.

(30a–c) all are only felicitous as corrections of assertions that are echoed under the scope of the negation. The observation is that the echoic use of correction in (30) makes the embedded implicatures happen. In fact, they're not subsequently cancelable either. From an SDRT perspective, all of these examples involve Correction to some prior constituent π_0, of the form *Joe saw Mary or Sue*. The correction move in the first clause of (30a) must correct something in π_0, and the second clause elaborates and in effect says what the corrected element should be. In this case the only thing that can be corrected is the implicature that Joe didn't see Mary and Sue.

Asher (2002) (written almost a decade earlier) provides an analysis of discourse relations in terms of a map from the source (the constituent to be linked to the discourse structure) to a target (a discourse constituent that serves as the other term of the Correction relation). This map exploits prosodic cues and the logical structure of the constituents. It can also be made to serve our purposes here.[16] My account of Correction involves the following constraint:

- Correction(α, β) holds only if K_β entails $\neg K_\alpha$ and there is a map $\mu : K_\beta \rightarrow K_\alpha$ such that there is at least some element x of K_β $K_\beta(\mu(x)/x) > K_\alpha$. The element x is said to be the *correcting element* and should be prosodically marked.

In (30a) the Correction move is signalled by the form of (30a) and natural prosody. There is an echoic use of *John saw Mary or Sue*, that the first clause (30a) denies. So we naturally read this example as occurring in a discourse context in which there is a constituent K_{π_0}, *John saw Mary or Sue*, that is the target of the Correction move. In a natural dialogue, you could just drop (30a.1), the first clause of (30a). (30a.2), the second clause of (30a), elaborates on or specifies the meaning that the speaker means to suggest as a replacement for the corrected content. The full discourse structure of (30a) and its target is then:

(30a′) Correction(α, π); π: Elaboration(30a.1, 30a.2)

Given (30a), especially the elaboration (30a.2), we need to investigate how the constraint on Correction is satisfied. The map $\mu : K_\pi \rightarrow K_\alpha$ (actually the sub map $\mu : K_{30a.2} \rightarrow K_{\pi_0}$) maps $\wedge$ to $\vee$. This map supplies an S-alternative for K_{π_0}: John saw M & S. But K_α, encoding the litteral meaning of *John saw Mary or Sue*, *doesn't* satisfy the constraint on Correction. Moreover, keeping K_α as the target of the Correction would not allow us to specify any sensible relation

between (30a.1) and (30a.2). In this case, however, the presence of S-inference allows a repair of the discourse context, in which we add the implicature generated using (7), (8) and the procedure outlined in Section 4, to K_{π_0}—i.e., K_{π_0} is reset to: *John saw Mary or Sue, and J. didn't see Mary and Sue*. In other words, we have a local accommodation of the implicature within K_{π_0}. The constraint on Correction is now satisfied. As Correction requires the presence of the implicature, the implicature is not cancellable, except on pain of discourse incoherence, as witnessed in (30a″):

(30a″)   #Joe didn't see Mary or Sue (or both); he saw BOTH.

(30a″) is nonsense because the sentence has strong clues indicating Correction thus precluding any other discourse relation, but the Correction constraint can't be satisfied. So no discourse relation can be inferred.

The introduction of the S-inference in the example above is a sort of discourse based coercion. Notice that the implicature calculated is relative to the map μ, *only after* GL has inferred a discourse relation. So there is no conflict between the GL computation of discourse structure, which is primary, and the computation of S-implicatures which is secondary. μ provides the relevant element in the alternative set to *or*; we then proceed to calculate the implicature as in Section 4 on the local part of the SDRS. The inference of the S implicature itself, however, is triggered by the need to satisfy the constraints imposed by the discourse move of Correction. Thus, D and S inferences are codependent; the need to calculate a D inference triggers the calculation of the S implicature, and it is the S implicature that satisfies the constraints of the D inferences.

Another set of examples of embedded S-inferences concerns the discourse relations of Parallel and Contrast.[17]

(31)   a.   [If [you take salad or dessert]$_{\pi_1}$, [you pay 20 euros]$_{\pi_2}$]$_{\pi_3}$; [but if [you take both]$_{\pi_4}$ [there is a surcharge.]$_{\pi_5}$]$_{\pi_6}$
       b.   If most of the students do well, I am happy; if all of them do well, I am even happier.
       c.   If you can fire Joe, it is your call; but if you must, then there is no choice.
       d.   Every professor who fails most of the students will receive no raise; every professor who fails all of the students will be fired.

Contrast also involves a structure preserving map μ from source β to target α and requires that at least one element x of τ_β be such that x and $\mu(x)$ be defeasibly contradictory, in the sense that they defeasibly imply, in context, contradictory propositions (Asher, 1993). Contrast also has a particular restriction when it involves two conditionals; the map μ must specify map the antecedent of the conditional in the source to the antecedent of the conditional in the target and the two antecedents must be defeasibly contradictory. In the case of (31a), for example, we have the Contrast signalled by the presence of the discourse connector *but*. Then:

- Contrast requires a map $\mu : \pi_6 \rightarrow \pi_3$ such that
- $\mu(\pi_5) = \pi_2$ and $\mu(\pi_4) = \pi_1$ where π_1 and π_4 are defeasibly contradictory.
- Given that the content of π_1 and π_4 are not as they stand defeasibly contradictory, we need to repair the situation with an S-inference if possible.
- The map μ provides a map from $A \vee B$ to $A \wedge B$.
- The presence of this alternative allows GL to compute the S-inference $\neg(A \wedge B)$ as in Section 4 relative to the logical form of π_2. The implicature is now added to π_2, allowing us to satisfy the constraint on Contrast.

Once again, it is the discourse structure and the need to support the D-implicature of Contrast introduced by the discourse particle *but* that provides the relevant alternatives and triggers the S implicature. The S implicature verifies the structural constraint on the discourse context imposed by Contrast. Once again, this is an implicature that is non-cancelable in this discourse context; it's as strong as any unembedded implicatures, stronger even. These examples lead to the following generalization:

- If the discourse structure requires the S implicature for coherence, it isn't cancellable except on pain of incoherence.

(31b) provides another example of S-implicature driven by the relation of Contrast. The juxtaposition and connection of the two complex constituents forms the basis of the implicature in the first constituent. The structures are very close; it is just the difference between the two quantifiers in the conditional that validate the two different consequents and the contrast between the two conditionals. The natural map takes the antecedent and consequent of the source and maps them onto antecedent and consequent of the target. I note that the map specifies in fact the scale: *most* vs. *all*. The

---

[17] My view implements and corroborates the observations about local implicatures and structures of parallelism and contrast in Geurts (2010).

implicature that *most* yields *most but not all* is needed to validate the structural requirement of Contrast. Once again, it is discourse structure that licenses the S implicature.

To verify that it is the juxtaposition of the two clauses that drives the implicature consider:

(32)   John loves Susan. But Sam loves her too.

John loves Susan doesn't implicate on its own that other people, and in particular Sam, don't love Susan. But to support the Contrast in (32), we get the implicature that Sam isn't expected to love Susan or that perhaps lots of other people beside John aren't expected to love Susan.

Besides Contrast, there are many other discourse relations that trigger S-implicatures. In (33b), we have plausibly a Continuation of the discourse topic (an *and* is easily insertable between the two sentences while the insertion of *but* is slightly less felicitous), what did the students do?,. As many have noted, (33b) contains what some have labelled a pair of *contrastive topics*; that is, in the two clauses the predication exploits a partition over the students and predicates properties that are incompatible. To verify this structural constraint, once again we need to trigger an exclusivity S-implicature in (33b),

(33)   a.   Two students wrote a paper or ran an experiment. (weak, or no, implicature)
        b.   Two students wrote a paper or ran an experiment. The others either did both or made a class presentation. (exclusivity of the disjunction)
        c.   Three students did most of the exercises; the rest did them all.

Particular kinds of elaborations also generate S-inferences that are uncancellable. Consider again:

(27)   a.   John has an even number of children. He has four.
        b.   John has an even number of children, four.
        c.   #John has an even number of children (π_1). He has three (π_2).
        d.   #John has an even number of children (π_1), three (π_2).
        e.   John has an even number of children (π_1), and he has at least three (π_2).

The structure of (27a–c) implies that π_2 modifies the DP just to its left. Note also that you can't introduce *and*) between the two clauses (27a,b) or (27c,d), though you can with (27e). These clues tell us that (27a–d) exemplify the discourse relation of entity-elaboration or E-elab, and the second term of the relation must identify some object mentioned in the first term. Here the object in question is the even number of John's children. To meet the identification constraint on E-elab, *four* must mean exactly four. Thus, as is by now familiar, the constraints on discourse structure make the S inference in (27a,b) uncancellable. On the other hand, while E-elab is signalled in (27c,d), the constraint on identification of the even number can't be met. So these examples are predicted to be infelicitous. While it is possible to have E-elab in (27e), it's not necessary or even preferred; inserting an *and* between the two clauses is a strong signal that a non-elaborating discourse relation is in play.

Question answer pairs also trigger S-implicatures.[18] A question induces, dynamically speaking, a partition on the information state (Groenendijk, 2008; Asher, 2007). A complete answer picks out one cell in the partition; indirect answers (which stand in the IQAP relation to the question they address) may require reasoning or additional information to infer a complete answer. Sometimes the additional information comes from an implicature given by a structure preserving map from the response to the question. IQAP can in fact give rise to an *overanswer* in which we get more than just a complete answer to the question; we get a more informative subset of the element of the partition picked out.

A particularly interesting case involves prosodically marked overanswers to polar questions like the following.

(34)   a.   Did John eat all of the cookies?
        b.   John ate SOME of the cookies.

(34b) is an overanswer to (34a). By itself (34b) doesn't provide enough information to compute an answer to the question. But the prosodic marking gives rise to a structure preserving map from the response to the question that must map the focussed element to an element of the same type in the question while preserving as the non focussed structure and content as far as possible. In this case the prosodically marked *some* is mapped to *all*, and provides the relevant alternative set. The lack of a full answer also triggers the S-inference, and including the inference that John didn't eat all of

---

[18] Zondervan, 2007 provides empirical evidence that question answer pair generates implicatures.

the cookies together with (34b) provides a complete answer to (34a). Thus, (34b) also gives more information than just a simple *no* would have.

A response that on its own fails to give a complete answer to a question can also trigger embedded implicatures and ones that wouldn't be calculated from standard lexical alternatives for *some*. Consider, for instance, the S-implicature of (35b), which is that John believes that not many of the students passed the exam, or the implicature of (36b) which is that everyone didn't read most of the books. These implicatures follow given the structure preserving mapping required by question-answer pairs that maps *some* to *many* in (35b) and *some* to *most* in (36b).

(35)　a.　Does John believe that many of the students passed the exam?
　　　b.　John believes that SOME of the students passed the exam.

(36)　a.　Did everyone read most of the books?
　　　b.　Everyone read SOME of the books.

Notice that once again it is the addition of the implicatures in these responses that gives us complete answers to the questions they are paired with.[19]

Yet another case of an implicature arising from the presence of a discourse relation concerns the case where a speaker responds to a question with another question:[20]

(38)　a.　A: Where is John?
　　　b.　B: Where is Jill?

SDRT's GL links (38b) via the relation of *Question Elaboration* or Q-elab, which has a particular semantics. Q-elab(α, β) holds iff an answer to β helps determine an answer to α. And this is indeed the "implicature" that interpreters draw from (38b)—namely that B believes that getting an answer to his question will help find an answer to A's question.

If we take the line of thought developed here to its inclusion, it implies that even unembedded S inferences are not the product of Gricean style reasoning dependent on some set of stipulated alternatives but rather largely dependent on the computing of discourse structure. Consider again unembedded disjunctions. The S-inferences conveyed by simple unembedded disjunctions are predicted not to arise when not needed to verify constraints imposed by discourse structure or to infer discourse relations. Txurruka and Asher (2008) argue that disjunctions are defeasible marks of the relation Alternation. Alternation partitions the set of discourse possibilities into non-empty sets relative to some topic, which can be introduced via a question or simple assertion.[21] When the topic, however, already contains the disjunction, the constraint that is a consequence of Alternation is not met and so neither Alternation nor the exhaustivity implicature holds, as in (40):

(40)　a.　Did John meet the Vice President or the President?
　　　b.　He must have met the Vice President or the President, since he got the job.

A similar moral holds for S inferences generated by question answer pairs. If they're not required by the constraints on discourse structure, the S inferences don't arise, even if the implicature is consistent with the information in the discourse context.

(41)　Did some students go to the party?

(42)　Yes, some students went to the party.

---

[19] The approach extends naturally to Wh-questions:

(37)　a.　Who read all of the books on the reading list?
　　　b.　JOHN read SOME of the books.

This example follows the treatment of polar questions, except there are two prosodically prominent elements in the response. It is the second that under the structure-preserving mapping generates the implicature.

[20] Thanks to Bart Geurts for bringing up this example at the Nijmegen Implicatures Workshop, Nijmegen, January 2012.

[21] Here's an example of Alternation, which also involves an auto-Correction.

(39)　Sam or Susan came, or both did.

The meaning of the first disjunct of (39) has to be strengthened to *Sam or Susan but not both came*, if the disjuncts are both to be informative and Alternation is to hold.

Here there is no prosodically marked element and no pairing of two distinct elements on a scale. I predict the implicature not to arise.

Let's now return to the problematic

(43)   None of the students answered all of the questions.

(43) does not seem to have, in an out of the blue context, the implicature:

(44)   All of the students answered some of the questions

even though (43) is equivalent to

(45)   All the students didn't answer all of the questions.

which does have the predicted implicature, because its more complex form suggests some sort of a Correction. A natural context for (45) is:

(46)   a.   A: All of the students answer all of the questions.
        b.   B: No, all of the students DIDN'T answer ALL of the questions.

There is a natural prosodic prominence to the second occurrence *all* in (45) signalling a lexical choice that is the source of the disagreement, which gives rise to a structure preserving map and the relevant implicature. (43) in this context less clearly has the implicature because the Correction is also supported by the choice of *none*.

The interaction between discourse structure and S inferences has implications on how evidence for S implicatures has often been gathered, as argued by Geurts (2009). In fact even out of the blue contexts may involve tacit questions under discussion in the way suggested by Roberts (1996), which may be responsible for generating most if not all S inferences, in the way that i have outlined for the discourse relation IQAP. I discuss this issue further in Asher (in press), which is crucial to the success of this general line of thought.

7. Conclusions

I've argued that there is a unified theory of implicatures. Implicatures, understood as defeasible implications, arise from several sources: semantics, discourse structure or prosody together with the logical form of what is said. They employ axioms in a non-monotonic logic, the Glue Logic of SDRT, which works on logical forms rather than semantic contents in order to preserve tractability and exploit structure at the level of the sentence and the discourse. But there are distinctions between D implicatures and S implicatures; the latter are to a large extent parasitic on the first. On my approach S implicatures are triggered by the requirements of the discourse context; they are calculated relative to a set of alternatives either provided lexically or by the discourse context, in large part to render consistent or to strengthen discourse relations that GL has already computed. The theory makes several new predictions: S inferences can be ''cancelled'' even if they're consistent with the purely semantic content given information in the discourse context; S inferences may also be uncancellable even in the face of inconsistency, when they are mandated by discourse structure; embedded S-inferences (both negative and positive) require a more elaborate discourse setting to be triggered and so should be harder to get without the appropriate discourse context. This line of thinking also suggests a line of empirical research: given the right discourse contexts, embedded implicatures should follow as easily as the unembedded ones, to some extent confirmed by Zondervan (2008).

References

Alfonso-Ovalle, L., 2005. Distributing the disjuncts over the modal space. In: Bateman, L., Ussery, C. (Eds.), North East Linguistics Society, vol. 35. Amherst, MA.

Asher, N., 1993. Reference to Abstract Objects in Discourse. Kluwer Academic Press.

Asher, N., 1995. Commonsense entailment: a logic for some conditionals. In: Crocco, G., del Cerro, L.F., Herzig, A. (Eds.), Conditionals: From Philosophy to Computer Science. pp. 103–147.

Asher, N., 2002. From discourse micro-structure to macro-structure and back again: the interpretation of focus. In: Kamp, H., Partee, B. (Eds.), Current Research in the Semantics/Pragmatics Interface, Vol. 11: Context Dependence in the Analysis of Linguistic Meaning. Elsevier Science Ltd.

Asher, N., 2007. Dynamic discourse semantics for embedded speech acts. In: Tsohatzidis, S. (Ed.), John Searle's Philosophy of Language. Cambridge University Press, pp. 211–244.

Asher, N., 2008. Troubles on the right frontier. In: Kühnlein, P., Benz, A. (Eds.), Proceedings of Constraints in Discourse 2005. Benjamins.

Asher, N. Implicatures in discourse. In: Grewendorf, G., Zimmerman, E. (Eds.), Discourse and Grammar. From Sentence Types to Lexical Categories. de Gruyter, in press.

Asher, N., Bonevac, D., 2005. Free choice permission is strong permission. Synthèse 145, 22–43.

Asher, N., Lascarides, A., 1998. The semantics and pragmatics of presupposition. Journal of Semantics 15, 239–299.

Asher, N., Lascarides, A., 2003. Logics of Conversation. Cambridge University Press.

Asher, N., Morreau, M., 1991. Commonsense entailment: a modal theory of nonmonotonic reasoning. In: Proceedings of the 12th International Joint Conference on Artificial Intelligence.

Asher, N., Pogodalla, S., 2010. SDRT and continuation semantics. In: LENLS 2010, Tokyo, Japan.

Block, E., 2009. Gricean implicature. Paper presented at the Michigan Pragmatics Workshop.

Chemla, E., 2008. Universal implicatures and free choice effects: experimental data. Semantics and Pragmatics 2 (2), 1–33.

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

Chierchia, G., Fox, D., & Spector, B., 2008. The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. Draft.

Fox, D., 2007. Free choice disjunction and the theory of scalar implicatures. In: Sauerland, U., Stateva, P. (Eds.), Presupposition and Implicature in Compositional Semantics. Palgrave Macmillan, New York, pp. 71–120.

Fox, D., Katzir, R., 2011. On the characterization of alternatives. Natural Language Semantics 19.1, 87–107.

Geurts, B., 2009. Scalar implicature and local pragmatics. Mind and Language.

Geurts, B., 2010. Quantity Implicatures. Cambridge University Press.

Geurts, B., Pouscoulous, N., 2009. Embedded implicatures?!? In: Egré, P., Magri, G. (Eds.), Presuppositions and Implicatures. MIT Working Papers in Linguistics.

Groenendijk, J., 2008. Inquisitive semantic. In: Proceedings of SALT XVIII, The University of Massachusetts at Amherst, Amherst, MA.

Hardt, D., Asher, N., Busquets, J., 2001. Discourse parallelism, scope and ellipsis. Journal of Semantics 18, 1–16.

Horn, L., 1972. The semantics of logical operators in English. PhD Thesis. UCLA.

Horn, L., 2005. The border wars: a neo-Gricean perspective. In: Turner, T., von Heusinger, K. (Eds.), Where Semantics Meets Pragmatics. Elsevier.

Karttunen, L., Peters, S., 1979. Conventional implicature. In: Oh, C.K., Dinneen, D.A. (Eds.), Syntax and Semantics, Vol. 11: Presupposition. Academic Press/Harcourt Brace Jovanovich, New York, pp. 1–55.

Kehler, A., Kerta, L., Rohde, H., Elman, J., 2008. Coherence and coreference revisited. Journal of Semantics (Special Issue on Processing Meaning) 25 (1), 1–44.

Kratzer, A., Shimoyama, J., 2002. Indeterminate pronouns: the view from Japanese. In: The Proceedings of the Third Tokyo Conference on Psycholinguistics. p. 125.

Kroch, A., 1972. Lexical and inferred meanings for some time adverbs. Quarterly Progress Report of the Research Laboratory of Electronics 104.

Lascarides, A., Asher, N., 1993. Temporal interpretation, discourse relations and commonsense entailment. Linguistics and Philosophy 16, 437–493.

Lee, C., 2010a. Scalar implicatures revisited. In: Kishimoto, H. (Ed.), Kotoba-no Taisyoo (Speech in Contrast). Kurosio, Tokyo, pp. 67–80.

Lee, C., 2010b. Information structure in PA/SN or descriptive/metalinguistic negation: with reference to scalar implicatures. In: Dingfang, S., Turner, K. (Eds.), Contrasting Meaning in Languages of the East and West. Peter Lang, Berne, pp. 33–73.

Roberts, R., 1996. Information structure in discourse: towards an integrated formal theory of pragmatics. In: Yoon, J.H., Kathol, A. (Eds.), OSU Working Papers in Linguistics 49: Papers in Semantics. The Ohio State University Department of Linguistics, pp. 91–136.

Romero, M., Hardt, D., 2004. Ellipsis and the structure of discourse. Journal of Semantics 21, 1–42.

Sauerland, U., 2004. Scalar implicatures in complex sentences. Linguistics and Philosophy 27, 367–391.

Schulz, K., 2007. Minimal models in semantics and pragmatics: free choice, exhaustivity, and conditionals. PhD Dissertation. University of Amsterdam, 2007.

Schulz, K., van Rooij, 2006. Pragmatic meaning and non-monotonic reasoning: the case of exhaustive interpretation. Linguistics and Philosophy 29 (2), 205–250.

Schwarzschild, R., 1999a. GIVENness, AvoidF and other constraints on the placement of accent. Natural Language Semantics 7, 141–177.

Schwarzschild, R., 1999b. GIVENness, Avoid F, and other constraints on the placement of focus. Natural Language Semantics 7 (2), 141–177.

Spector, B., 2006. Aspects de la pragmatique des opérateurs logiques. PhD Dissertation. Université Paris 7.

Spector, B., 2007. Scalar Implicatures: Exhaustivity and Gricean Reasoning. In: Aloni, M., Butler, A., Dekker, P. (Eds.), Questions in Dynamic Semantics. Elsevier, pp. 225–249.

Txurruka, I., Asher, N., 2008. A discourse-based approach to Natural Language Disjunction (revisited). In: Aunargue, M., Korta, K., Lazzarabal, J. (Eds.), Language, Representation and Reasoning. University of the Basque country Press.

van Rooij, R., Schulz, K., 2004. Exhaustive interpretation of complex sentences. Journal of Logic Language and Information 13 (4), 491–519.

Vieu, L., Bras, M., Asher, N., Aurnague, M., 2005. Locating adverbials in discourse. Journal of French Language Studies 15 (2), 173–193.

Zondervan, A., 2008. Scalar implicatures or focus: an experimental approach. PhD Dissertation. University of Utrecht.