# *'Only'* and Monotonicity in Conditionals

Matthew Berends
Dept. of Linguistics
Northwestern University

matthewberends@gmail.com

Stefan Kaufmann
Dept. of Linguistics
Northwestern University

kaufmann@northwestern.edu

**Abstract**

This paper proposes an account of the interpretation of *'only'* in the antecedents of indicative conditionals. Our concern lies with the implication from a conditional of the form *if (only* $\phi$*),* $\psi$ to its *'only'*-less counterpart *if* $\phi$, $\psi$: when and why it is warranted. We argue that the pragmatic relationship of *scalar upward monotonicity* determines its availability. Two factors serve as license. First, it may arise by virtue of a language user's pre-existing world knowledge. Second, it may manifest when it constitutes the most informative reading of the conditional available. We discuss one case in point; namely, its appearance when the consequent is desirable.

## 1 Introduction

Some, but not all conditionals with *'only'* in the antecedent license the inference to their *'only'*-less counterparts. Thus (1) implies that doing his homework will ensure that Chris passes the class. By contrast, (2) does not convey that Chris's doing his homework will ensure that he fails the class.[1]

(1) a. If Chris only [does his homework]$_F$, he will pass the class.
b. $\leadsto$ If Chris does his homework, he will pass the class.

(2) a. If Chris only [does his homework]$_F$, he will fail the class.
b. $\not\leadsto$ If Chris does his homework, he will fail the class.

The goal of this paper is to explain when and why the implication from *'if only* $\phi$, $\psi$*'* to *'if* $\phi$, $\psi$*'* holds. Our analysis relies on three main ingredients. The first is the *scalarity* of *'only'* in cases in which the inference is licensed. Intuitively, both the availability of the inference in (1) and its absence in (2) rely on a construal of the situation in which doing his homework and nothing else is at the "low end," in terms of effort or cost, of the list of things Chris might do to secure his success in the class. Formally, we argue that the inference, where available, relies in part on an ordering relation between the alternatives invoked by *'only'*. The second ingredient is a relation between this order

[1] The bracketing indicates the location of focus; see Section 4 below for details.

on the one hand, and the truth of the consequent of the conditional, on the other. Simply put, the question is whether, assuming that the alternative antecedents are ranked in terms of effort, the consequent becomes more likely by moving up or down the scale. In the above examples, it is reasonable to assume that expending more effort will make the consequent more likely in (1) but not in (2).[2] We refer to this relationship as *scalar (upward or downward) monotonicity*. It is distinct from the familiar up/downward monotonicity in terms of entailment: Although we assume that *'Chris only does his homework'* entails *'Chris does his homework'* (see Section 4), we will whos in the next section that the effects of scalar monotonicity are not limited to such cases.

It is evident from the above remarks that both of the first two ingredients – the scalarity of *'only'* and the scalar monotonicity between the alternative antecedents and the consequent – crucially depend on the content of the conditional's constituents. Thus not surprisingly, background world knowledge plays an important role in explaining when and why the inference goes through. The third ingredient of our account goes beyond the immediate goal of explaining the inference in cases in which these relations are common knowledge. Here we show that one can run the account "backwards," so to speak, and infer from the fact that the speaker chose to use the conditional with *'only'* in the antecedent, rather than its *'only'*-less counterpart, that the requisite scalar relations hold. This reasoning crucially relies on two elements: the assumption that the speaker is being helpful, and common knowledge as to whether the consequent is desirable or not.

We discuss the data in some more detail in Section 2. Following that, we prepare the ground for our account by laying out our assumption about the semantics of the main ingredients: conditionals (Section 3) and *'only'* (Section 4). We then spell out the main ingredients, specifically scalar monotonicity, in Section 5. Section 6 is devoted to the role of the desirability of the consequent. We conclude with Section 7.

## 2 Data

We first point out that the conditionals we are dealing with here are ones in which *'only'* is embedded inside the antecedent. This is the case for (1) and (2) above, as well as for the examples below. Our analysis does not apply to conditionals with *'only'* in other positions, such as (3).

(3) Only if Chris does his homework will he pass the class.

The analysis does apply to sentences like (4), however. We realize that an utterance of (4), in addition to making an assertion about Chris's situation, carries an "evaluative" connotation on the part of the speaker to the effect that Chris's success is important to him or her. Our analysis does not explain this connotation, but we believe that crucial parts of our analysis are prerequisites for an understanding of it.

(4) If only Chris does his homework, he will pass the class.

[2]In fact, is seems plausible that more effort makes the consequent *less* likely in (2).

As we mentioned above, one main ingredient of our account is what we call *scalar monotonicity*. We first take a look at its upward version – scalar upward monotonicity, or 'SUM' – in more detail, previewing the main ideas of our analysis along the way. We construe the SUM relationship as resting on two components: a scale consisting of a set of alternative antecedents and an ordering on this set; and an upward monotonic relationship between the values on the scale and the consequent. Given these components, SUM holds if and only if: if the consequent is true for some element of the scale, then it is also true for all higher-ranked elements. The case of scalar downward monotonicity is similar.

*'Only'* facilitates the inference by making a scale salient, but is not itself essential. A scale may also arise implicitly from the context. Thus the inferences in (5) and (6) go through despite the absence of *'only'*.

(5) a. If Chris passes the final, he will pass the class.
  b. $\rightsquigarrow$ If Chris gets a "B+" on the final, he will pass the class.

(6) a. If Chris gets a "B+" on the final, he will fail the class.
  b. $\rightsquigarrow$ If Chris fails the final, he will fail the class.

On occasion, particular lexical items used in the conditionals in question draw attention to a scale. Thus in (7) and (8), it is the cardinals and gradable adjectives, respectively, that invoke the scale in virtue of their lexical semantics.

(7) a. If you have five dollars, you can buy a medium coffee.
  b. $\rightsquigarrow$ If you have more than five dollars, you can buy a medium coffee.

(8) a. If you are a lazy student, you will pass the class.
  b. $\rightsquigarrow$ If you are a hard-working student, you will pass the class.

That the availability of the inference indeed depends on the content of the conditional rather than its logical form is indicated by the reversal of the pattern in (1-a) when the antecedent belongs to a scale of activities which cumulatively lead to failing rather than passing the class.

(9) a. If Chris only [skips the exam]$_F$, he will pass the class.
  b. $\not\rightsquigarrow$ If Chris skips the exam, he will pass the class.

(10) a. If Chris only [skips the exam]$_F$, he will fail the class.
  b. $\rightsquigarrow$ If Chris skips the exam, he will fail the class.

As mentioned above, we identify two general reasons that scalar monotonicity might manifest. One may know that the relationship holds by virtue of one's pre-existing knowledge of how the world works. Alternatively, an attempt to derive as informative an interpretation of the conditional as possible may lead one to infer that it holds even in the absence of any supporting world knowledge. The role of informativeness is made plain by its interaction with the desirability consequent. It is possible to infer from the mere fact that the speaker chose to assert *'if (only $\phi$), $\psi$'* that a SUM relationship holds, hence *'if $\phi$, $\psi$'* can be inferred, without knowing so in advance. We observe this with the following conditionals, which fail to provide, at first glance, either a notion of the relevant alternatives to *'florp'*, nor a dimension along which to rank those alternatives, nor a sense of how the resulting scale relates to the consequent. However, the reader may verify for herself that the second-person subject greatly facilitates the availability of these judgments.

(11) a. If you only [florp]$_F$, you will win a hundred dollars.
b. $\rightsquigarrow$ If you florps, you will win a hundred dollars.

(12) a. If you only [florp]$_F$, you will lose a hundred dollars.
b. $\not\rightsquigarrow$ If you florp, you will lose a hundred dollars.

# 3 Conditionals and Modality

We follow Lewis (1975) and Kratzer (1981) in treating conditionals of the form *if* $\phi$, $\psi$ as modal expressions. Our assumptions for this paper are very simple: Syntactically, a conditional is composed of a matrix clause and an adverbial headed by *'if'*. We refer to the adverbial clause as the *antecedent* and to the matrix clause as the *consequent*. Semantically, *'if'* introduces a modal operator, restricted by the antecedent, that scopes over the consequent.[3]

The semantic analysis of modal expressions that we adopt is that of Kratzer (1981, 1991; see also Kaufmann, 2005b; Kaufmann et al, 2006; Portner, 2009). Modal operators are interpreted relative to two parameters, a *modal base* and an *ordering source*. Both are *conversational backgrounds* – formally, functions from possible worlds to sets of propositions (i.e., sets of sets of possible worlds.) The modal base provides the domain of modal quantification and determines the kind of modality (e.g., epistemic, doxastic, deontic etc.), analogously to the *accessibility relations* familiar from modal logic. More specifically, from the perspective of a world $w$ of evaluation and a modal base $f$, the worlds in the intersection $\bigcap f(w)$ play the same role as those accessible from $w$ via the corresponding accessibility relation. However, within this set of accessible worlds, some may be more salient, likely, or otherwise relevant to the truth of the conditional than others. This is incorporated by ranking worlds according to their "goodness" with respect to the propositions provided by the ordering source.

**Definition 1 (Frame)** A *frame* is a structure $\langle W, f, g\rangle$ where $W$ is a non-empty set of worlds and $f$ and $g$ are *conversational backgrounds*, i.e., functions from worlds to sets of propositions.

**Definition 2 (Accessibility relation)** Given a frame $\langle W, f, g\rangle$, the *accessibility relation* $R_f$ determined by $f$ defined as follows: For all $w, w' \in W$, $wR_f w'$ iff $w' \in \bigcap f(w)$.

**Definition 3 (Relative likelihood)** Given a frame $\langle W, f, g\rangle$, the *relative likelihood* order determined by $g$ is a three-place relation $\leq$ on $W$ defined as follows: For all $w, w', w'' \in W$,

$$w' \leq_{g(w)} w'' \text{ iff } \{p|w'' \in p \wedge p \in g(w)\} \subseteq \{p|w' \in p \wedge p \in g(w)\}$$

[3] This treatment is too simplistic to account for certain facts. For instance, there is good evidence that the modal operator is not introduced by *'if'*, but by tense. For this and related ideas, see Kaufmann (2005a).

It is easy to show that $\leq_{g(w)}$ is transitive and reflexive.

For simplicity, we illustrate with a language $\mathcal{L}_{\mathcal{A}}$, which for now is simply the language of standard propositional logic built on a set $\mathcal{A}$ of propositional variables. We extend it with a modal operator below.

**Definition 4 (Model)** A *model* for a language $\mathcal{L}_{\mathcal{A}}$ is a structure $\langle W, f, g, [\![]\!]\rangle$ such that $\langle W, f, g\rangle$ is a frame and $[\![\cdot]\!] : \mathcal{L}_{\mathcal{A}} \mapsto (W \mapsto \{0,1\})$ is a valuation function mapping propositional variables and their Boolean compounds to (characteristic functions of) subsets of $W$.

As mentioned above, we treat conditionals as modal expressions and assume, following Kratzer, that their interpretation depends not only on the modal base, but also on the ordering source. The main role of the ordering source here is to provide the formal basis for a weaker notion of necessity (and a stronger notion of possibility) than that afforded by standard necessity and possibility operators. Kratzer refers to these modal forces as *human necessity* and *human possibility*, respectively. The idea is to make certain worlds in the modal base irrelevant to the interpretation of modal expressions. The worlds made irrelevant are those that are strictly "outranked" by others with respect to some contexually salient criterion (plausibility, likelihood, normalcy). The relevant ranking is given by the ordering source, formally represented as a world-dependent preorder on the set of possible worlds, defined in Definition 3 above. Human necessity is defined with reference to this order as in Definition 5. Human possibility is its dual.

**Definition 5 (Human necessity)** The notion $\boxdot$ of *human necessity* is defined as follows: For all $w \in W$ and $\phi, \psi \in \mathcal{L}_{\mathcal{A}}$: $[\![\boxdot(\phi)(\psi)]\!]^{M,w} = 1$ iff for all $w'$ in $\bigcap f(w)$ such that $[\![\phi]\!]^{M,w'} = 1$, there is some $w''$ in $\bigcap f(w)$ such that $[\![\phi]\!]^{M,w''} = 1$ and $w'' \leq_{g(w)} w'$ and for all $w'''$ in $\bigcap f(w)$ such that $[\![\phi]\!]^{M,w'''} = 1$ and $w''' \leq_{g(w)} w''$), $[\![\psi]\!]^{M,w'''} = 1$.

Notice that the syntactic form employed in Definition 5 assumes that $\boxdot$ is a binary operator. We follow a convention common in linguistics and refer to its two arguments $\phi$ and $\psi$ as its *restrictor* and its *scope*, respectively. Typically, for simple sentences involving human necessity modals, the restrictor is semantically inert, and the relevant set of worlds is simply $\bigcap f(w)$. We may account for this by assuming that in the absence of any overt information about the restrictor, the constant function $\lambda w.1$ is inserted by default.

In the case of conditionals, we follow Kratzer's assumption that a covert human-necessity modal is present by default (although that modal force can be overridden by overt modal expressions). Here, the *'if'*-clause does contribute explicit information about the restrictor, while the consequent serves as the scope. The relevant domain of modal quantification, then, consists of those worlds in the modal base where the antecedent holds. Thus we interpret a conditional *'if* $\phi$, $\psi$*'* as $\boxdot(\phi)(\psi)$.

(13) $\quad \mathbf{if'} = \lambda\phi\lambda\psi \boxdot (\phi)(\psi)$

The reason why we use human necessity in the semantics of *'if'* rather than strict necessity is that human necessity better captures our intuitions about the inferences available with conditionals. In particular, we avoid licensing strengthening of the antecedent.

With strict necessity, we would incorrectly predict that *if* $\phi$, $\psi$ entails *if* $(\phi \wedge \gamma)$, $\psi$, for any $\gamma$. Human necessity avoids this problem because the ideal $(\phi \wedge \gamma)$-worlds relevant for the latter need not be a subset of the $\phi$-worlds relevant for the former. Accordingly, the latter may $\psi$-worlds while the former are not.

Most relevant for our purposes is that human necessity blocks the entailment from *if* $\phi$, $\psi$ to *if (only* $\phi$*)*, $\psi$. The intuitive notion is that the "stereotypical" ways of doing $\phi$ may make propositions true that are denied by *only* $\phi$. For instance, an utterer of (1-b) might take for granted that typically, students who complete their homework also attend class. However, a speaker of (1-a) may have in mind that doing one's homework is sufficient for passing without any further effort. Clearly (1-b) does not entail (1-a) under these conditions. Nevertheless, a classical account of indicative conditional meaning would predict the entailment from *if* $\phi$, $\psi$ to *if (only* $\phi$*)*, $\psi$ under the assumption (which we make – see below) that (only $\phi$) entails $\phi$. Using human necessity as the modal operator correctly allows the implication to fail.

# 4 Focus and *'Only'*

With a semantics for conditionals and modality in hand, we can now discuss *'only'*. We begin with a bare-bones and intuitive account of *'only'*, and proceed to refine this by appeal to notions of information structure and then finally scalarity.

*'Only'* is commonly analyzed as bearing two distinct semantic components, its *positive* and its *negative* contribution. The *prejacent* of a sentence containing *'only'* is what remains after removing *'only'*. Thus in (14), the prejacent is (14-a). The positive contribution is the proposition denoted by the prejacent. The negative contribution is the negation of a number of alternative propositions derived in a certain systematic way which relies on a bipartition of the prejacent into two parts, typically referred to as *focus* and *background*.

(14) Only Bill slept.
  a. Bill slept.
  b. {Mary slept, Sue slept, Bill and Mary slept, ...}

Simply put, the alternatives are derived by substituting alternative values for the focus (subject to pragmatic factors such as domain restriction) while holding the background constant. In (14), the focus is *'Bill'* and the background is *'slept'*. Thus the alternatives are propositions which assert of various individual(s) that it/they slept. It is usually assumed that the prejacent is itself one of the alternatives (Rooth, 1992). The negative contribution is the denial of all alternatives other than the prejacent (Horn, 1996).

It is generally agreed that the negative contribution is an entailment, but there is some debate over whether the positive contribution is implicated, presupposed, or entailed. Without committing ourselves irrevocably to either of these positions, we see good reasons for taking the latter route and assuming that the positive contribution, like the negative one, is an entailment. If it were an implicature, then *'only p'* would literally mean that among the alternatives, *at most p* is true, but it would be consistent with the falsehood of $p$. This would result in the wrong predictions about the conditionals we are concerned with: *'Chris only does his homework'* would be equivalent to *'Chris does at most his homework'*, leading to truth conditions for (1-a) and (2-a) that are too strong.[4] On the other hand, if the positive contribution were a presupposition, we would expect it to project out of the antecedent, but we see no evidence for that. Thus we assume, at least for the purposes of this paper, that both the positive and the negative contribution are entailed.

We note in passing that while the focus is generally marked by accent placement, the relationship between prosodic accent and semantic focus is not one-to-one and mediated in the standard theoretical approach by an abstract syntactic feature 'F'. Thus for instance, while the focus marking in (15) would be unambiguously expressed by placing the nuclear pitch accent on *'walked'* and *'dog'*, respectively, an accent placement on *'dog'* would be compatible with all three focus markings in (16).

(15) a. John only [walked]$_F$ his dog
b. John only walked [his]$_F$ dog

(16) a. John only [walked his dog]$_F$
b. John only walked [his dog]$_F$
c. John only walked his [dog]$_F$

The relationship between accent placement and F-marking is the object of a long-standing line of investigation and continues to be debated (Schwarzschild, 1999; Selkirk, 2001; German et al, 2006, among others). An exploration of this topic would lead us too far afield. Instead, for the purposes of this paper, we simply assume that the F-marking in the antecedents in question is given. F-marked constituents are treated as the focus, while the remainder of the clause constitutes the background.

The interpretation of *'only'* sketched above can be made precise within a *Structured Meanings* representation (Krifka, 1991, 1995). We assume that *'only'* syntactically takes sentential scope and applies to an entire structured meaning. Formally, we may represent it as in (17).[5]

(17) (Preliminary version)
$\mathbf{only}' = \lambda\langle F,B,A\rangle[B(F) \wedge \forall X[(X \in A \wedge B(X)) \rightarrow X = F]]$

To illustrate, the sentence in (18-a), with F-marking as indicated, is interpreted as (18-b), which given the denotation in (17) simplifies to (18-c).

[4]One way to avoid this undesirable result while maintaining that the prejacent is an implicature would be to claim that this implicature is compiled into the literal meaning as part of the interpretation of the antecedent. Such a view would not be without precedent (Chierchia, 2004), but we will not explore it any futher here.

[5]Notice that while we assume that *'only'* composes with a structured meaning, the result is not a structured meaning again. Thus the dentotation in (17) does not allow other focus-sensitive expressions "higher up" to associate with a focus in the same clause. See Krifka (1991) for a solution to this problem, which we sidestep here because it is orthogonal to our concerns.

(18) a. John only [did his homework]$_F$

b. $$\textbf{only}'\left(\left\langle \lambda x.\text{do-hw}(x), \lambda P.P(j), \begin{Bmatrix} \lambda x.\text{attend}(x), \\ \lambda x.\text{pass-final}(x), \\ \lambda x.\text{do-hw}(x), \ldots \end{Bmatrix} \right\rangle\right)$$

c. $$\text{do-hw}(j) \wedge \forall X \left[ \left( \begin{array}{l} X \in \begin{Bmatrix} \lambda x.\text{attend}(x), \\ \lambda x.\text{pass-final}(x), \\ \lambda x.\text{do-hw}(x), \ldots \end{Bmatrix} \\ \wedge X(j) \end{array} \right) \rightarrow X = \lambda x.\text{do-hw}(x) \right]$$

The above denotation for *'only'* denies that alternatives other than the focus truthfully combine with the background. However, not all sentences with *'only'* seem to make so strong a claim. On occasion, the background may truthfully apply to other alternatives. This happens when *'only'* bears a scalar tinge: In this case, the alternatives whose truth is not denied are "lesser" in some sense than the focus. For illustration, imagine a selection of lunch choices ordered by size, as in (19-c). A truthful utterance of (19-a) assuredly conveys that John did not eat a hamburger, but need not deny that he had a handful of raisins. There are two distinct readings: on the first, John ate nothing *other* than an apple; on the second, John ate nothing *more* than an apple. The data is clearer if we imagine (19-a) as a response to either *'Did John eat any items from the refridgerator for lunch?'* or *'Why does John look so famished?'*.

(19) a. John only ate [an apple]$_F$ for lunch

b. Focus alternatives to *'apple'*: {apple, raisins, pear, cheeseburger}

c. $\langle$raisins $\preceq$ {pear, apple} $\preceq$ cheeseburger$\rangle$

Debate lingers over the nature of this scalar reading. One possible approach is to explain it as a restriction on the set of alternatives, stipulating that *raisins* fails to count as an alternative. However, this misses the fact that even among potential alternatives that are excluded for pragmatic reasons, there is an asymmetry between higher-ranked ones and lower-ranked ones: The former are still denied (counterfactually, as it were) in a way the latter are not. For instance, one does not typically think of *steak* as a luncheon choice, but (19-c) certainly seems to deny that John ate steak for lunch. Alternatively, the optionality of the scalar reading might suggest that *'only'* is ambiguous. This approach gains traction from the presence of similar accounts for *'even'*. However, although there may be good reasons for taking this approach in the case of *'even'*, parsimony demands that we consider this a last resort. In the case of *'only'*, several authors (e.g., Bonomi and Casalegno, 1992; Beaver, 2004; van Rooij, 2002) have demonstrated that a uniform scalar denotation for *'only'* is sufficient. On such an implementation, *'only'* conveys that the combination of the background and higher-ranked alternatives is false, but crucially makes no claims about the combination of the background with lower-ranked alternatives. The appropriate scale is determined by a contextually given dimension for ranking the alternatives. In the default case, the relevant scale is the semi-lattice of the alternatives with conjunction, thus the relevant "scale" in that case is that of entailment.

Under this approach, we can always assume that a scale ranked by a pre-order $\preceq$ is available. This allows for a unified account of *'only'* along the lines of (20).

(20) $\textbf{only}' = \lambda\langle F,B,A\rangle[B(F) \wedge \forall X[(X \in A \wedge B(X)) \rightarrow B(X) \preceq B(F)]]$

Notice that here we compare $B(X)$ and $B(F)$, rather than just $X$ and $F$ as in (17) above. The reason for this is that we assume that the scale in question is uniformly one of propositions. In (17), the difference did not matter since if $X = F$, then $B(X) = B(F)$. However, since $B$ may itself contribute to the scalar ordering (e.g., the scale gets reversed if $B$ contains a negation), from $X \prec F$ it does not follow that $B(X) \prec B(F)$, and it is the latter that counts for the truth of the sentence.

### 4.1 *'Only'* in conditional antecedents

The composition of *'only'* within the antecedent of a conditional offers no surprises. Intuitively, the result is an interpretation of *'if only* $\phi$, $\psi$*'* as *'if* $\phi$ *and no more than* $\phi$, $\psi$*'*. Importantly, $\psi$ is not ruled out if *'more than* $\phi$*'* is true. This means that, borrowing from (1), Chris may do other things and still pass the class.

(21) $\textbf{if}'(\text{only } \langle F,B,A\rangle)(\psi) = \boxdot(\textbf{only}'\langle F,B,A\rangle)(\psi)$
$= \boxdot(B(F) \wedge \forall X[(X \in A \wedge B(X)) \rightarrow B(X) \preceq B(F)])(\psi)$

## 5 Scalar Monotonicity and *'Only'*

The implication from *if (only* $\phi$*),* $\psi$ to *if* $\phi$*,* $\psi$ prompted our excursion into the interpretation of *'only'* conditional antecedents. The implication is present for some conditionals, but absent for others. Our paper explores why the particular asymmetry exists and what might explain it. The answer we arrive at is that the implication is governed by a pragmatically-determined scalar relationship between the antecedent and the consequent of a conditional. Before we go into nature of this relationship, we briefly mention why we discard entailment and ambiguity in *'only'* as possible explanations. Then, in the remainder of the paper we discuss the role of scalar monotonicity. There are four parts to our discussion. First, we offer a description of what it is to be SUM. Second, we discuss how scalar monotoniciy accounts for the implication arising in (1). Third, we supply an explanation of the conditions under which the SUM relationship arises. Lastly, we discuss how the SUM reading may be informative.

### 5.1 Against ambiguity and entailment as explanations

Here we consider two possible alternative explanations before proferring our SUM explanation. These are first that *'only'* is ambiguous between a reading that permits the implication and one that does not, and second that the relationship is adequately characterized by entailment.

The most straightforward-seeming hypothesis regarding the asymmetry is that it arises from an ambiguity in the meaning of *'only'* itself. However, we can account for our data with a single denotation for *'only'*, as we argue below. Parsimony, then, demands that we explain the asymmetry as arising from facts about conditionals rather than an ambiguity with regard to *'only'* itself.

As for entailment, we find that it offers an inadequate description of the relationship between the sentences in (1). Given that the implication in (1) is from a stronger antecedent to one that it is weaker, it is an instance of an upward monotonic inference that has never been claimed valid for conditionals.

## 5.2 Scalar monotonicity characterized

The above observations are made concrete in the examples below. If *'if Chris only does his homework, he will pass the class'* is true, then so is the conditional *'if* $\phi$*, Chris will pass the class'*, for all alternatives $\phi$ that are higher than than *'Chris only does his homework'* on the scale, such as the antecedents in (22-b) and (22-c):

(22) a. If Chris only does his homework, he will pass the class.
b. $\rightsquigarrow$ If Chris does his homework, he will pass the class.
c. $\rightsquigarrow$ If Chris does his homework and attends class, he will pass the class.
d. $\not\rightsquigarrow$ If Chris does his homework and gets caught cheating, he will pass the class.

Working out a formal implementation of this idea is not trivial. Suppose the scale in question is $\langle\Phi,\preceq\rangle$, and the proposition is $\psi$. Intuitively, what one would want is a condition on the distribution of $\psi$-worlds in the various propositions in $\Phi$, stating in effect that if $\psi$ is a human necessity relative to some $\phi\in\Phi$, then it is also a human necessity relative to all $\phi'$ such that $\phi\preceq\phi'$.

To see why this is not trivial, notice first that in order for $\psi$ to be a human necessity relative to both $\phi$ and $\phi'$, it is not required that $\psi$ be true at all worlds in either $\phi$ or $\phi'$. For as we saw above, it is the very point of human necessity that some worlds in the restrictor are made irrelevant to the truth of the modal expression. Without this ability, we would predict that (22-b) entails (22-d) and (22-c), since both instantiate strengthenings of the antecedent. We also saw that this ability is important for our purposes because otherwise we could not account for the fact that even though *'only p'* entails $p$, the conditional *'if p, q'* does not entail *'if only p, q'*.

So we need a more restricted statement roughly to the effect that if we inevitably and inescapably end up in $\psi$-territory by inspecting less and less far-fetched worlds in $\phi$, then the same is bound to happen when we inspect $\phi'$-worlds in the same manner. Somewhat more simplified, assuming that there is a set of "best" worlds (technically, local minima) under $\leq_{g(w)}$ within each $\phi\in\Phi$, if all the "best" worlds in $\phi$ are $\psi$-worlds, then all the "best" worlds in $\phi'$ are $\psi$-worlds, too. Stating it in this way is easy enough, but we would like to go deeper than that by capturing the conditions under which this outcome is guaranteed in terms of the worlds in $\phi$, $\phi'$, and $\psi$.

The intuition now is that even though the worlds in $\phi$ and $\phi'$ may be distinct, there is nevertheless a "correspondence" of sorts which determines, for a given world $v$ in $\phi$, which worlds in $\phi'$ are "at least as good" as $v$ with respect to the ordering source. This seems straightforward enough, but we cannot be sure that any two worlds in $\phi$ and $\phi'$ are even comparable under $\leq_{g(w)}$. For instance, if $g(w)$ contains $\phi$ and $\phi'$, then no two worlds are comparable in terms of "goodness" across the two propositions.[6] As an easy way to avoid this problem, we simply stipulate that for all worlds $w$, $g(w)$ consists entirely of propositions that have non-empty intersections with *all* propositions in the alternative set $\Phi$.[7] It is important to keep this in mind in reading the following definition:

**Definition 6 (Scalar Upward Monotonicity)** Let $M = \langle W, f, g \rangle$ be a model and $\langle \Phi, \preceq \rangle$ a scale of propositions such that for all $w \in W$ and $\phi \in \Phi$, all propositions in $g(w)$ have a non-empty intersection with $\phi$. A proposition $\psi$ is *scalar upward monotone* at a world $w \in W$ relative to $\langle \Phi, \preceq \rangle$ if and only if for all $\phi, \phi' \in \Phi$ such that $\phi \preceq \phi'$ and all worlds $v \in \phi, v' \in \phi'$ such that $v' \leq_{g(w)} v$, if $z \in \psi$ for all $z \leq_{g(w)} v$, then $z' \in \psi$ for all $z' \leq_{g(w)} v'$.

# 6 Scalar monotonicity and desirability

One factor affecting the availability of the inference concerns the interplay between scalar monotonicity and the interlocutors' goals. The relevant examples are repeated here. The failure of the inference from (2-a) to (2-b), as well as from (9-a) to (9-b) tends to get strengthened to the conclusion that the (b)-conditional is false.

(1) a. If Chris only [does his homework]$_F$, he will pass the class.
 b. $\rightsquigarrow$ If Chris does his homework, he will pass the class.

(2) a. If Chris only [does his homework]$_F$, he will fail the class.
 b. $\not\rightsquigarrow$ If Chris does his homework, he will fail the class.

(9) a. If Chris only [skips the exam]$_F$, he will pass the class.
 b. $\not\rightsquigarrow$ If Chris skips the exam, he will pass the class.

(10) a. If Chris only [skips the exam]$_F$, he will fail the class.
 b. $\rightsquigarrow$ If Chris skips the exam, he will fail the class.

Assuming competence on the part of the speaker, this follows from two assumptions: First, it is common knowledge that listeners strive to make choices which lead to desirable outcomes, and to avoid negative ones, both with minimal effort; and second, speakers try to impart information that will help listeners in doing so.

For given a scale $\langle \Phi, \leq \rangle$ of alternative antecedents and a consequent $q$, let $A = \{p \in \Phi | \textit{If } p, q \text{ is true}\}$. Then cooperative speakers will choose '*If* min$(A)$, $q$' if $q$ is scalar increasing in $A$, as in (1-a) and (10-a), and '*If* max$(A)$, $q$' if $q$ is scalar monotone decreasing in $A$, as in (2-a) and (9-a) (choosing at random if min$(A)/\max(A)$ is not unique).

Together with the fact that $q$ is desirable in (1-a) and (9-a), the listener expects the speaker to choose an antecedent that is *minimal* on its respective scale among those alternatives for which the conditional is true – for knowing the least costly way to guarantee the truth of the consequent is useful both in securing and in preventing the latter.

[6] Problems arise under weaker conditions as well. For instance, if $g(w)$ contains non-empty subsets of $\phi$ and $\phi'$, then there are two worlds $v \in \phi, v' \in \phi'$ such that neither $v \leq_{g(w)} v'$ nor $v' \leq_{g(w)} v$. We do not attempt a complete characterization of the problematic cases in this paper.

[7] Ultimately the most elegant solution might be to manipulate $g$ "online" in the course of the interpretation, filtering our propositions from the ordering source that imposes a ranking among the alternatives.

Antecedents higher on the scale than the minimal ensure the truth of the conditional, too. Likewise, since $q$ is undesirable in (2-a) and (10-a), the listener expects the speaker to choose an antecedent that is *maximal* on its scale among those alternatives that ensure the truth of the conditional – the listener's interests are the same way as before, but since the consequent is decreasing in $A$, knowing the most costly way to ensure its truth is more useful to him. Antecedents higher than the maximal one do not ensure the truth of the conditional. The preceding argument rests on the assumption that *only* is scalar, such that for each of the two scales, *only* $p \leq p$.

## 7 Conclusions and Future Directions

We investigated the implication from *if (only* $\phi$*),* $\psi$ to *if* $\phi$*,* $\psi$, arguing that its presence hinges on the availability of a scalar relationship between antecedent and consequent. The pragmatic relationship of Scalar Upward Monotonicity that governs the implication is characterized by three components. SUM first requires a set of alternatives; second, a ranking of the alternatives; and third a relationship between the ranking and the consequent. SUM holds if the consequent remains true when higher-valued propositions are substitued into the antecedent.

Two factors appear to license SUM. First, language users may know of it simply by virtue of their world knowledge. Second, pragmatic considerations may lead them to conclude that it motivates the speaker's choice of asserting *'if only* $\phi$, $\psi$*'* rather than just *'if* $\phi$, $\psi$*'*. Future work will explore this link between desirability and scalar monotonicity in greater detail.

## Acknowledgements

We are grateful to Brady Clark, the audiences who commented on presentations of various versions of this material, especially at the 2008 Annual Meeting of the LSA and at Sinn und Bedeutung 13, and to the editors of this volume for their assistance and patience. The second author gratefully acknowledges the support of the Japan Society for the Promotion of Sciences (JSPS, Project "Inferentical mechanisms and their linguistic manifestations").

## References

Beaver, D. 2004. Five *only* pieces. *Theoretical Linguistics* **30**, 45-64.

Bonomi, A. and P. Casalegno. 1992. Only: Association with focus in event semantics. *Natural Language Semantics* **2**, 1-45.

Chierchia, G. 2004. Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In Belletti, A. (ed.) *Structures and Beyond*, pages 39-103. Oxford University Press.

German, J.B., J. Pierrehumbert, and S. Kaufmann. 2006. Evidence for phonological constraints on nuclear accent placement. *Language* **82**(1), 151-168.

Grice, H.P. 1989. *Studies in the Way of Words*, Cambridge, MA: Harvard University Press.

Hajicová, E., B. Partee, and P. Sgall. 1998. *Topic-focus Articulation, Tripartite Structures, and Semantic Content*, Dordrecht: Kluwer Publishers.

Horn, L. 1996. Presupposition and implicature. In S. Lappin (ed.), *Handbook of Contemporary Semantic Theory*, 299-319. Oxford: Blackwell.

Jackendoff, R. 1972. *Semantic Interpretation in Generative Grammar*, MIT Press.

Kaufmann, S. 2005. Conditional truth and future reference. *Journal of Semantics* **22**(3), 231-280.

Kaufmann, S. 2005. Conditionals. In Brown, Keith, editor-in-chief, *Encyclopedia of Language and Linguistics*, 2nd edition, Volume 3, 6-9. Elsevier.

Kaufmann, S, C. Condoravdi, and V. Harizanov. 2006. Formal approaches to modality. In Frawley, William (ed.) *The Expression of Modality*, pages 71-106. Mouton de Gruyter

Kratzer, A. 1981. The notional category of modality. In Eikmeyer, J. and H. Riesner, editors, *Words, Worlds, and Contexts*, 38-74. Walter de Gruyter.

Kratzer, A.: 1991. 'Modality/Conditionals'. In: von Stechow, A. and D. Wunderlich (eds.): *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung. [=Semantics]*, 639-656. Mouton de Gruyter.

Krifka, M. 1991. A compositional semantics for multiple focus constructions. Ms. University of Texas, Austin.

Krifka, M. 1995. The semantics and pragmatics of polarity items. *Linguistic Analysis* **25**, 1-49.

van Kuppevelt, J. 1996. Inferring from Topics: Scalar Implicatures as Topic-Dependent Inferences. *Linguistics and Philosophy* **19**, 393-443.

Lewis, D. 1975. Adverbs of quantification. In Keenan, E.L. (ed.), *Formal Semantics of Natural Language*. Cambridge U.P.

Portner, P. 2009. *Modality.*p Oxford University Press.

van Rooij, R. 2002. Relevance only. in Bos J., M.E. Foster, and C. Matheson (eds): *Proceedings of the sixth workshop on the semantics and pragmatics of dialogue (EDILOG 2002)*, 155-160.

Rooth, M. 1992. A theory of focus interpretation. *Natural Language Semantics* **1**, 75-116.

Schwarzschild, R. 1999. GIVENness, AvoidF, and other constraints on the placement of focus. *Natural Language Semantics* **7**(2), 141-177.

Selkirk, E. 2001. The syntax-phonology interface. In Smelser, N.J. and P.B. Baltes (eds.) *International Encyclopedia of the Social and Behavioral Sciences*, 15407-15412. Oxford: Pergamon.