

## “UNLESS”: AN EXPERIMENTAL APPROACH

Perna Nadathur & Daniel Lassiter

Stanford University

The connective *unless* is often suggested as a challenge to compositionality, on the grounds that it combines differently with positive and negative quantifiers [H86]. In particular, it has been argued that *unless* is biconditional in positive contexts but unidirectional in negative ones. We report on an experiment which demonstrates that *unless* is not biconditional under *every* as claimed, and in which a majority of participants endorse sentences predicted false by the **exceptional** solution proposed by [vF92,Le08]. We identify three additional, theoretically interesting new patterns in our results, propose a new semantics and pragmatics which accounts for the first two, and consider some directions for explaining the third, gradient pattern.

**Theories and predictions.** Both accounts assume that *if* and *unless* can restrict modals, quantificational adverbs, and nominal quantifiers alike. For the less troublesome *if...not* construction we assume (with  $Q[C]$  representing quantifier  $Q$  restricted to set  $C$ ,  $M$  the predicate of the main clause, and  $R$  the translation of the clause embedded under *if not/unless*):

$$(1) Q[C] M \text{ if not } R \Leftrightarrow Q[C \wedge \neg R] M$$

$$(2) \text{Every marble has a dot if it is not blue.} \Leftrightarrow \text{All non-blue marbles have dots.}$$

$$(3) \text{No marble has a dot if it is not blue.} \Leftrightarrow \text{No non-blue marbles have dots.}$$

Von Fintel’s interpretation of quantified *unless*-sentences is sketched in (4), and Leslie’s in (5).

$$(4) Q[C] M \text{ unless } R \Leftrightarrow Q[C \wedge \neg R] M \wedge \forall S \subseteq C : Q[C \wedge \neg S] M \rightarrow R \subseteq S \quad (\text{von Fintel})$$

$$(5) Q[C] M \text{ unless } R \Leftrightarrow Q[C \wedge \neg R] M \wedge Q[M \wedge C] \neg R \quad (\text{Leslie})$$

For our critical items below involving *unless*, these accounts predict the following.

$$(5) \text{Every marble has a dot unless it is blue.}$$

**von Fintel/Leslie:** All non-blue marbles have dots, and no blue marbles do.

$$(6) \text{No marble has a dot unless it is blue.}$$

**von Fintel:** No non-blue marbles have dots, and all blue marbles do.

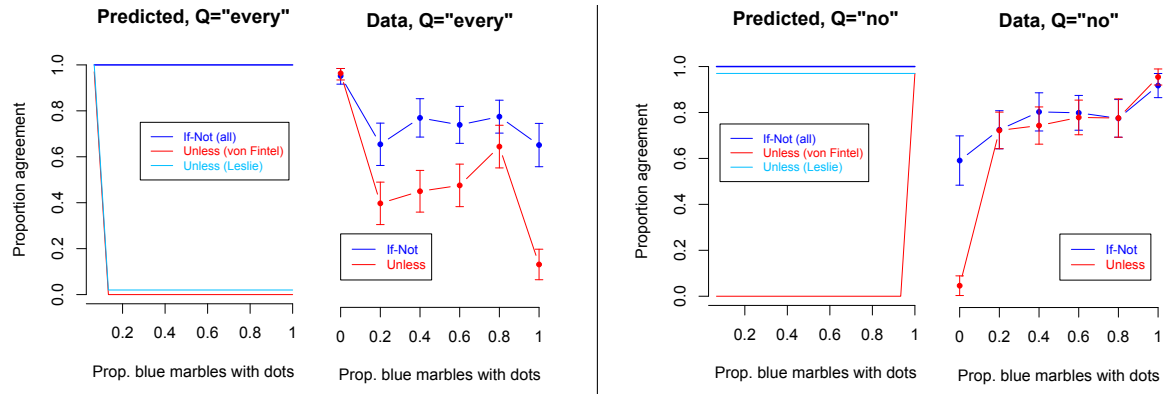
**Leslie:** No non-blue marbles have dots.

**Methods.** 155 participants were recruited using Amazon’s Mechanical Turk platform, filtered for native English status. Participants saw 24 critical quantificational trials interspersed randomly with 24 fillers, each paired with a visual display: 20 marbles, with a randomly selected red/blue distribution of 15/5, 10/10, or 5/15. The proportion of blue marbles with dots varied between 0, .2, .4, .6, .8, and 1. In each critical trial participants were presented with sentences like those in (5)/(6) or matched *if...not* sentences, and were asked to judge if the sentence was true or false of the display. The colour of the marble in the prompt also varied (we use ‘red’ and ‘blue’ as labels for continuity with examples above). In order to avoid overwhelming participants with false sentences, we ensured that the minimal truth conditions on all theories were satisfied: all red marbles had dots in *every*-sentences, and no red marbles had dots in *no*-sentences.

**Results.** The figures below show the experimental data plotted alongside idealized (noise- and pragmatics-free) predictions of the two theories. The latter are jittered slightly for visibility. The x-axis gives the proportion of blue marbles with dots, and the y-axis gives the proportion of ‘True’ responses for the relevant sentence. Error bars are 95% binomial confidence intervals.

**Interpretation.** Relevant aspects of pragmatics (e.g. conversational implicature, presupposition) could readily lead some participants to reject true sentences, but parallel mechanisms salvaging false sentences are unavailable. A theory which predicts a sentence  $S$  as false – so zero acceptance *modulo* noise – is thus incorrect if a nontrivial proportion of participants endorse  $S$ .

By this logic, the *every* condition’s results falsify both theories. Biconditionality of *unless* under *every* means that *Every marble has a dot unless it is blue* should be false above 0, but a sizable fraction of participants endorsed this sentence at proportions .2-.8, reserving categorical



rejection for 1. The *no* results are also problematic. Von Fintel predicts zero acceptance away from 1, but a majority of participants endorsed *No marble has a dot unless it is blue* on .2-.8, falling off to categorical rejection only at 0. Leslie's account, on the other hand, is unidirectional with *no* and so predicts no effect from proportion manipulation in the *no* condition. In particular, it is surprising for this theory that participants categorically judge the sentence as false at 0.

The one respect in which *if ... not* and *unless* differ categorically is at 1 (for *every*) and 0 (for *no*). The results reveal an additional surprise: there is a systematic but non-categorical divergence between *unless* and *if ... not* in the 0.2-0.8 range of the *every* condition.

**A new semantic and pragmatic account.** The problem is to find a semantics and pragmatics for *if ... not* and *unless* that accounts for (a) their categorical divergence at 1/0 in the respective conditions, (b) the fact that both are less than fully acceptable in the intermediate range, and (c) the fact that *unless* is worse than *if ... not* in the intermediate range, but only under *every*. Puzzles (a) and (b) boil down to the contrast between merely degraded (7) and unacceptable (8).

- (7)  $p$  unless  $q$ , but sometimes  $p \wedge q$       (8)  $p$  unless  $q$ , but always  $p$

We propose that *if ... not* has the interpretation in (1), while *unless* is revised to:

- (9)  $Q[C] M$  **unless**  $R \Leftrightarrow Q[C \wedge \neg R] M \wedge \neg Q[C] M$ .

Since *unless*-sentences do not readily embed, it is difficult to discern whether the boldfaced addition is an entailment or a presupposition; we remain agnostic for present purposes. Either would predict (a), the categorical rejection at 0/1 with *unless*, but not with *if ... not*.

Phenomenon (b), the reduced acceptability in the .2-.8 range is related to the degraded status of (7). We predict both if there is a weak "never  $p \wedge q$ " inference associated both *unless* and *if ... not*. This inference can be identified as a **conditional perfection** implicature [GZ71], which applies equally to *unless* and *if ... not*. We do not take a stand here on how conditional perfection is explained or how broadly applicable it is, requiring only that it is a pragmatic inference which applies to both *unless* and *if ... not*.

The remaining puzzle is to explain why the *unless* examples are reliably less acceptable than *if ... not* examples in the *every* condition, but not in the *no* condition. One possibility is that there is some additional pragmatic pressure, in addition to conditional perfection, which further degrades the intermediate *every ... unless* examples; it remains unclear precisely what this inference is, however. Another possibility is that conditional perfection is itself explained as a strengthening inference from a more basic **conditional strengthening** inference [vF01], and that this pragmatic reasoning interacts differently with *if ... not* and *unless* under *every*.

While puzzle (c) remains unexplained, the present work contributes to the debate by sharpening the empirical issues and by motivating the position that important aspects of the *unless/if ... not* distinction reside in their pragmatics, not their semantics. In addition, our work reinforces the broader point that *unless*, while puzzling, is no serious threat to compositionality.

**References:**

[vF92] **von Fintel**, Exceptive conditionals: the meaning of “unless”. NELS22: 135-151. [vF01] **von Fintel**, Conditional strengthening: a case study in implicature. Ms. MIT. [GZ71] **Geis & Zwicky**, On invited inferences. *Linguistic Inquiry* 2: 561-566. [H86] **Higginbotham**, Linguistic theory and Davidson’s program in semantics. Lepore (ed), *Truth and Interpretation: Perspectives on the Philosophy of Donald Davidson*: 29-48. [Le08] **Leslie 2008**. “If,” “unless,” and quantification. Stainton & Viger (eds), *Compositionality, Context, and Semantic Value: Essays in honor of Ernie Lepore*: 3-30.