Unless, exceptionality, and the pragmatics of conditional statements

Prerna Nadathur

November 25, 2014

Contents

1 Introduction

The connective *unless* is often cited as a potential counterexample to a compositional theory of semantics, on the grounds that it contributes a different meaning when embedded under positive quantifiers than it does under negative ones (Janssen 1997, Szabó 2008). The original argument, due to Higginbotham (1986), relies on the problematic classical identification of *unless* with the negative material conditional *if not* (as in, q *unless* $p := \neg p \rightarrow q$);[1] however, the strong intuition that there is a difference between negatively- and positively-quantified *unless*-statements persists, even in work that eschews material implication. In its most current form, the problem centers around the empirical claim that *unless* apparently contributes a (negative) biconditional in positive quantificational contexts, but only a one-directional conditional in negative contexts (see Leslie 2008).

(1) a. Every student will succeed unless he goofs off.
$\rightsquigarrow$ *Every student who does not goof off will succeed and every student who does goof off will not succeed.*

b. No student will succeed unless he works hard.
$\rightsquigarrow$ *No student who does not work had will succeed [but hard work is not a guarantee of success].*

In addition to being an apparently propositional operator that resists a truth-functional account, *unless* is therefore of some interest with respect to the status of compositionality. The challenge has been to develop a semantic treatment which not only reflects the perceived strength of *unless* (as compared to *if not*), but also captures a host of other intuitions, such as the contextual split in (1).

The best available approach to the problem treats *unless* as an exceptive operator on quantifier domains, likening it to *exception phases* such as *but* or *except for* (in, e.g. "Everyone but/except for John attended the party"; von Fintel 1992, 1994). In its most current form, due to Leslie (2008), the exceptive account handles the compositionality question by exploiting formal differences between positive and negative universal quantifiers to build the biconditionality/unidirectionality split directly into the semantics attributed to *unless*. This treatment makes a number of predictions which have yet to be fully examined in the literature, in particular regarding the use and range of *unless* as compared to *if not*.

This paper explores these predictions, as part of a broader aim of assimilating and evaluating the status of a number of interacting (and sometimes conflicting) intuitions that seem to accompany *unless*. In section 1, I discuss Higginbotham's original problem and set out the obstacles for any attempt to account for *unless* on purely truth-functional grounds. I next consider the improvements made by the exceptive framework of von Fintel and Leslie, but argue that even this move fails to capture several important aspects regarding the use of *unless*. The third section reports on the results of two experimental studies designed to test the predictions made by the exceptive account against a number of possible points of contention: the results support the view that the exceptive account is in need of revision, and in particular suggest a significant role for pragmatic factors (specifically, presupposition and implicature) in the interpretation of *unless*-statements.

Ultimately, based on both the theoretical considerations laid out here, and the new empirical data presented, I put forward a new proposal for the meaning of *unless*. I argue that the classical approach is correct insofar as *unless* shares asserted content with *if not*, and show how a careful formulation of the latter (due again to Leslie) can circumvent Higginbotham-style embedding problems. Second, I propose that the major difference between *unless* and *if not* is located in a condition on the use of conditional statements regarding the circumstances under which they are "best" asserted; this accompanies *if not* as an implicature, but *unless* as a presupposition. Finally, I argue that the apparent contextual split in (1) is a pragmatic one, rather than a semantic one, and propose that it results from the interaction of the appropriateness presupposition with a conditional-perfection-type implicature to biconditionality (see Geis and Zwicky 1971).

[1] The inadequacy of material implication as an interpretation of natural language conditionals has long been noted: see Égré and Cozic (2012) for a recent overview of arguments against this approach.

2 Truth conditions and Higginbotham's problem

Classically, *unless* has been treated as equivalent to the negative material conditional *if not*, which reduces logically to regular disjunction (Reichenbach 1947, Quine 1959):

(2) q *unless* $p := \neg p \rightarrow q \qquad (\leftrightarrow p \vee q)$

Higginbotham (1986) points out that, while this seems (at least) acceptable in "bare" or positively-quantified contexts such as (3a) and (3b), it supplies the wrong truth conditions when *unless* is embedded under a negative quantifier as in (3c). Here it would require that each individual in the domain neither succeeds nor works hard, when the statement should in fact not be falsified by the existence of individuals who both succeed and work hard.

(3) a. John will succeed unless he goofs off.
$= \text{succeed}(j) \vee \text{goof}(j)$

b. Everyone will succeed unless he goofs off.
$= \forall x[\text{succeed}(x) \vee \text{goof}(x)]$

c. No one will succeed unless he works hard.
$= \neg\exists x[\text{succeed}(x) \vee \text{work}(x)]$
$= \forall x[\neg\text{succeed}(x) \wedge \neg\text{work}(x)]$

Higginbotham argues that an intuitively acceptable interpretation for (3c) is given by instead replacing *unless* with *and not* to yield $\neg\exists x[\text{succeed}(x) \wedge \neg\text{work}(x)]$. On this basis, he concludes that *unless* varies its contribution according to the quantificational context in which it appears, and attributes to it a noncompositional semantics.

These examples show that material *if not* cannot provide an invariant semantics for *unless*; however, there is reason to question the adequacy of the equivalence in (2) as a definition. Uncertainty on this point is reflected by the fact that a number of alternative truth-functional proposals appear in the literature, including *p only if not q* from Clark and Clark (1977), which is based on the perceived similarity between sentences like (4a) and (4b).

(4) a. Twain liked people unless they were hypocrites.

b. Twain liked people only if they were not hypocrites.

Formally, this proposal gives q *unless* $p := p \rightarrow \neg q$, which is the converse of (2) and seems to capture what is missing from the translations in (3). Putting the two together, then, we get a negative biconditional $\neg p \leftrightarrow q$, which reinterprets the examples in (3) as follows:

(5) a. John will succeed unless he goofs off.
$= \neg\text{goof}(j) \leftrightarrow \text{succeed}(j)$

b. Everyone will succeed unless he goofs off.
$= \forall x[\neg\text{goof}(x) \leftrightarrow \text{succeed}(x)]$

Unfortunately, while the biconditional embeds under negative quantifiers in a well-behaved fashion (shown in (6)), it simply seems to enforce too strong a meaning for negatively-quantified *unless*-statements.

(6) No one will succeed unless he works hard.
$= \neg\exists x[\neg\text{work}(x) \leftrightarrow \text{succeed}(x)]$
$= \neg\exists x[(\neg\text{work}(x) \rightarrow \text{succeed}(x)) \wedge (\neg\text{work}(x) \leftarrow \text{succeed}(x))]$
$= \neg\exists x[\neg(\text{work}(x) \wedge \text{succeed}(x)) \wedge (\text{work}(x) \vee \text{succeed}(x))]$
$= \forall x[(\text{work}(x) \wedge \text{succeed}(x)) \vee \neg(\text{work}(x) \vee \text{succeed}(x))]$
$= \forall x[(\text{work}(x) \wedge \text{succeed}(x)) \vee (\neg\text{work}(x) \wedge \neg\text{succeed}(x))]$

The biconditional interpretation stipulates that working hard is not only a necessary condition but also a sufficient one for success. Intuitively, however, the English statement (6) only makes the claim that one cannot succeed without working hard – and this should not be falsified by someone who works hard and still does not succeed. In particular, it seems that we want a biconditional interpretation for *unless* in the positive cases but a one-directional interpretation in the negative cases. This difference cannot be captured by any purely truth-functional proposal for *unless*; as long as we are committed to a truth-functional account, then, Higginbotham's problem remains.

3 The exceptive account

3.1 Exceptionality and uniqueness

In addition to the motivation the biconditionality/unidrectionality split provides for considering non-truth-functional approaches, an independent line of inquiry in the literature provides a number of reasons to reconsider classifying *unless* as a propositional operator.

Against the equation of *unless* with *if not*, Dancygier (1985) proposes that, in the statement q *unless* p, "what *unless* really negates is q, not p" (p.68). She argues that someone asserting q *unless* p intends essentially to assert q but is also considering the circumstances under which q might not hold and so provides a qualification (to the effect that $\neg q$, if p). In other words, the speaker is noting that p represents a marked exception to the generalization that she primarily wishes to communicate.

The idea that *unless* has to do with exceptions also appears in Geis (1973), who offers a systematic comparison of *unless* with *except if*, and Zuber (1999), who compares *unless* to **exclusion phrases**, such as "No/every student *except* Leo." As conveyed by Dancygier's proposal, the central notion here is that *unless*-statements mark some particular set as an exception from the generalization the speaker wishes to assert (although the three authors do not necessarily agree on the precise formulation that this should take).

The first apparent formalization of the exceptive view comes from von Fintel (1992) (see also von Fintel 1994). He argues for a view of *unless* as an **exceptive operator** on quantifier domains, in the sense that *but* takes in "Everyone but John attended the party," and represents *unless* as a conjunction of two assertions. The first makes a (quantified) generalization over a domain from which the complement of *unless* (p, in q *unless* p) is subtracted. The second states that this complement is the unique smallest set over which the generalization q fails. This proposal is stated in (7).

(7) **Von Fintel's proposal:**
$Q[C]M \textit{ unless } R := (Q[C \wedge \neg R]M) \wedge (\forall S \subseteq C : Q[C \wedge \neg S]M \rightarrow R \subseteq S)$
where Q is the interpretation of the quantifier (or quantificational adverb),[2] C is its restriction, M its nuclear scope, and R is the *unless*-complement, or excepted set.

This provides the interpretation in (8) for (1a). I have used the abbreviation STU to represent the set of all (relevant) students and GOOF for the set of all individuals who goof off.

(8) Every student will succeed unless he goofs off.
$= (\text{ALL } x[\text{stu}(x) \wedge \neg\text{goof}(x)]\text{succ}(x)) \wedge (\forall S \subseteq \text{STU} : \text{ALL } x[\text{stu}(x) \wedge \neg(x \in S)]\text{succ}(x) \rightarrow \text{GOOF} \subseteq S)$

The first conjunct asserts that all students who do not goof off will succeed. This carries the meaning of "Every student will succeed if he does not goof off," but circumvents the problems of the material conditional

[2] It is worth noting that von Fintel (1992) does not explicitly provide a formula for interpreting *unless*-statements involving a nominal quantifier – the formula given above is intended specifically to capture the restriction as applied to quantificational adverbs, or a covert universal modal quantifier *à la* Kratzer (1986). He notes that *unless* clauses are also able to restrict nominal quantifiers, but does not say anything further on the subject; this leads to a certain amount of ambiguity as to the extent to which his proposal can be extended to cover these cases. Leslie (2008) shows that simply embedding von Fintel's *unless* under wide-scope nominal quantifiers leads to a number of undesirable conclusions due to the second conjunct. I have adopted what seems to be a more plausible extension of his proposal here, and assume that the quantifier Q in (7) can be taken to be either modal/adverbial or nominal. These assumptions are reflected in (8).

by switching to the restrictive Kratzer conditional (Lewis 1975, Kratzer 1986). The second conjunct (which von Fintel calls the "uniqueness" clause) asserts a qualification to the main generalization. In particular, it produces the effect of the reverse conditional (i.e. *not q if p*), by requiring that all of the students who do goof off are necessarily excluded from any arbitrary set that contains only successful individuals. This entails that none of the students who goof off are successful; thus, by stipulating the unique nature of the excepted set (R), von Fintel's proposal provides the desired biconditional interpretation for example (8).

3.2 Capturing the negative contexts

Consider what happens when we use proposal (7) to interpret negatively-quantified *unless*-statements.

(9) No student will succeed unless he works hard.
$= \text{NO } x[\text{stu}(x) \wedge \neg\text{work}(x)]\text{succ}(x)) \wedge (\forall S \subseteq \text{STU} : \text{NO } x[\text{stu}(x) \wedge \neg(x \in S)]\text{succ}(x) \rightarrow \text{WORK} \subseteq S)$

Example (9), interpretated according to (7), holds that no student who does not work hard will succeed, and no student who works hard can ever belong to a set which contains only unsuccessful individuals. Consequently, working hard is both a necessary and sufficient condition for success, and we see that von Fintel's account predicts biconditionality for *unless* in the negative cases as well as the positive ones.

This contradicts the intuitions outlined in section 2, and Leslie (2008) also notices this problem. To take an example that parallels hers, suppose we are discussing a university class which is notoriously difficult. We expect that students who take this class have to work very hard in order to pass, but it is also the case that only the smartest students pass (for some unfortunate students, even working hard is not enough). In this context, "No student will succeed unless he works hard" does not seem to be either an invalid nor an infelicitous statement, which suggests, as noted earlier, that it only requires working hard to be a necessary condition for success, but does not hold it to be sufficient. The interpretation of (9) given above fails to capture this.

Leslie offers a clever solution to this problem, which involves a simple modification of von Fintel's uniqueness claus. Her proposal is given in (10), with the updated clause marked in boldface.

(10) **Leslie's proposal:**
$Q[C]M \textit{ unless } R := Q[C \wedge \neg R]M \wedge \mathbf{Q[C \wedge M]\neg R}$

Leslie's uniqueness clause states that Q-many Cs that are M are not R; this retains the similarity to the reverse conditional *not q if p* enforced by von Fintel's version. However, the precise formulation that Leslie employs exploits the fact that the negative universal quantifier is **symmetric**: in general, the statement *No Xs are Ys* is logically equivalent to the statement that *No Ys are Xs*. Thus, when $Q = \text{NO}$, the uniqueness clause reduces to a statement equivalent to the first conjunct. The positive universal quantifier *every*, on the other hand, is not symmetric, so Leslie's proposal preserves biconditionality in the positive cases.

(11) a. Every student will succeed unless he goofs off.
$= \text{ALL } x[\text{stu}(x) \wedge \neg\text{goof}(x)]\text{succ}(x) \wedge \text{ALL } x[\text{stu}(x) \wedge \text{succ}(x)]\neg\text{goof}(x)$
All students who do not goof off succeed and all students who succeed do not goof off.

b. No student will succeed unless he works hard.
$= \text{NO } x[\text{stu}(x) \wedge \neg\text{work}(x)]\text{succ}(x) \wedge \text{NO } x[\text{stu}(x) \wedge \text{succ}(x)]\neg\text{work}(x)$
No students who do not work hard succeed.

For ease of presentation, I have so far glossed over some of the details of Leslie's proposal. The notion that *unless*-statements can restrict nominal quantifiers as well as quantificational adverbs (where C is then interpreted as picking out the set of contextually-relevant worlds or situations) is central to her account. She also argues that nominally-quantified *unless*-statements such as (11a) and (11b) incorporate a wide-scope universal modal quantifier, but the motivation for this (and the complications it introduces) will be discussed only in section 5.1.

3.3 Some potential issues for the exceptive account

Despite the elegance of Leslie's solution to the biconditionality/unidirectionality split, further examination reveals a number of issues with the exceptive account, and in particular with the predictions made by uniqueness. These issues can be divided into two classes: the interaction of uniqueness with non-universal quantifiers, such as *most*, and the stipulation of uniqueness as an entailed consequence of *unless*.

3.3.1 Non-universal quantifiers and across-the-board contexts

Von Fintel (1992) claims that *unless* can only occur with universal quantifiers (pg.144). He argues that uniqueness provides a natural explanation for this restriction: a quantifier such as *most* already incorporates the notion that there is a set over which the stated generalization fails to hold, and it does not seem particularly reasonable to discuss unique exceptions from *most*-statements. Leslie (2008) disagrees with this restriction on empirical grounds, and her view is supported by a wealth of naturally-occurring examples, including the statements given in (12).

(12) a. "Most livestock are fed GMO grains unless you buy organic pasture-raised animals."
b. "Some diners won't get water unless they ask."
c. "You cannot be certain how to pronounce some words unless you know their pre-history."
d. "Smoking kills half of smokers unless they quit."

Since Leslie's proposal incorporates the uniqueness clause, statements of this sort ultimately pose a problem for her account as well as von Fintel's. Consider, for example, the interpretation she assigns to (13).

(13) Most students will succeed unless they goof off.
$= \textsc{most}\ x[\text{stu}(x) \wedge \neg\text{goof}(x)]\text{succ}(x) \wedge\ \textsc{most}\ x[\text{stu}(x) \wedge \text{succ}(x)]\neg\text{goof}(x)]$

The first conjunct seems straightforward: most students who do not goof off succeed. At first glance, the second conjunct (most students who succeed do not goof off) seems reasonable as well, but things go wrong if the domain of students under consideration is dominated by students who do goof off. To illustrate the problem, suppose we have a domain consisting of 110 students, only 10 of whom do not goof off. If 9 of these students succeed, satisfying the first conjunct, the second conjunct then requires that these 9 comprise most of the students who succeed over the whole domain of 110. Thus, at most 8 (and possibly fewer) of the 100 students who do goof off can succeed without falsifying Leslie's interpretation of (13). This seems much too strong a condition to impose on this sentence. In particular, (13) would appear to be a reasonable description of a parallel situation in which 10, or 15 (or any number up to at least half) of the 100 students who goof off still manage to succeed. Leslie's formula rules these situations out.

This is not the end of the trouble. The scenario described above suggests that Leslie's truth conditions, and in particular those enforced by her second conjunct, are too strong. However, it is also possible to construct scenarios in which the truth conditions of (13) seem to be too general.

Suppose now that our domain of quantification is a class of 12 students, only 4 of whom goof off. 6 of the students who do not goof off succeed, and 3 of the others do as well. Then we have 6 successful students out of the 8 who do not goof off, which satisfies the first conjunct of (13). We also have 6 students who do not goof off from the total of 9 students who succeed, so the second conjunct is also satisfied. In this situation, however, most of the students who do goof off also succeed. In particular, goofing off makes absolutely no different to the student success rate: 75% of students succeed, whether they goof off or not. Here, (13) seems to be a highly inappropriate (if not false) description of the situation, but Leslie's formula makes no provision for this.

Truth conditions that seem too general are not, of course, inherently problematic: it might be possible to rescue (10) from the last objection by appealing to pragmatics to rule out situations such as the one constructed above. It is harder to make this appeal, however, for a quantifier like *half*, where Leslie's truth conditions actually logically enforce the problematic scenario. (14) provides Leslie's predictions for (12d).

(14) $\textsc{half}\ x[\text{smoker}(x) \land \neg\text{quit}(x)]\text{die}(x) \land \textsc{half}\ x[\text{smoker}(x) \land \text{die}(x)]\neg\text{quit}(x)$

That is, half of the smokers who do not quit die, and half of the smokers who die do not quit. On the assumption that HALF can be interpreted as *exactly half*,[3] this logically entails that the other half of smokers who die are erstwhile smokers. In particular, half of all (one-time) smokers die, whether or not they quit. Again, since the proportion of smokers affected does not seem conditional on their quitting status, this seems to be an inappropriate situation in which to use a statement like (12d): it is, however, the situation required by Leslie's semantics for *unless*.

More generally, the problem illustrated by (14) and by the second scenario provided for (13) can be stated as the fact that *unless*-statement seem to be inappropriate (if not false) in **across-the-board** contexts: situations where the main generalization ($Q[C]M$) holds across the entire domain, including the set (R) marked as an exception. Uniqueness enforces this (as a logical consequence of biconditionality) for universally-quantified *unless*-statements, but fails to capture it in the non-universal quantificational contexts examined here. Insofar as the naturally-occurring data in (12) support the view that these are both grammatically acceptable and interpretable occurrences of *unless*, this represents a serious gap in the exceptive account as it stands.

3.3.2 Uniqueness as a pragmatic inference

Having observed that uniqueness causes problems in non-universal contexts as well as negative ones, it seems reasonable to ask whether it ought to have been included in the semantics of *unless* in the first place. Its inclusion was motivated by the sense that "bare" (e.g. (3a)) and positively-quantified *unless*-statements seem to impart biconditionality, but pushing farther on this intuition suggests that biconditionality may not hold the status of an entailment even in these cases.

Consider the naturally-occurring example (15).

(15) a. "Mantou is always late unless she's already out before we meet."
b. "Mantou is always late unless she's already out before we meet, but she's often just less late then."

(15a) is assigned the following interpretations by von Fintel and Leslie, respectively (where s is a variable over worlds or situations, and $\text{Rel}(s)$ means that s belongs to the set of currently relevant situations):

(16) a. **Von Fintel:**
$(\textsc{all}\ s[\text{Rel}(s) \land \neg\text{already-out}(m,s)]\text{late}(m,s)) \land (\forall S \subseteq \textsc{rel} : \textsc{all}\ s[\text{Rel}(s) \land \neg(s \in S)]\text{late}(m,s) \rightarrow \textsc{already-out} \subseteq S)$
b. **Leslie:**
$\textsc{all}\ s[\text{Rel}(s) \land \neg\text{already-out}(m,s)]\text{late}(m,s) \land \textsc{all}\ s[\text{Rel}(s) \land \text{late}(m,s)]\neg\text{already-out}(m,s)$

Both (16a) and (16b) stipulate biconditionality: that all relevant situations are ones in which Mantou is not late if she's already out. Consequently, they predict that (15b), the original context of the example, is a contradictory statement, since it specifies that at least some of the situations in which Mantou is already out are ones in which she is late (albeit less late than other situations). (15b), however, does not present as contradictory, especially when compared to an obvious case of entailment-denial:

(17) #Roses are always red and violets are always blue, but sometimes violets are not blue.

The acceptability of (15b) suggests that uniqueness (or biconditionality) is defeasible, and is therefore not part of the entailed content of *unless*. Crucially, however, the biconditional interpretation of the truncated version in (15a) remains privileged, which suggests that an inference to uniqueness is attached by some pragmatic means at least to positive *unless*-statements.

[3] Similar problems arise if *at most half* or *at least half* is used instead.

Two additional observations support the claim that biconditionality is pragmatically derived: that is, the claim that the reverse (*not if*) conditional is pragmatically added to *unless*-statements. First, example (18) (also naturally-occurring) shows that the *not if* direction can be reinforced without apparent redundancy. Entailments do not have this property, as shown by the comparison example (18b).

(18) a. "Always be yourself, unless you are Fernando Torres. Then always be someone else."
b. *Compare:* Always be yourself, unless you are Fernando Torres. ?Otherwise always be yourself.

Here, uniqueness would produce something to the effect of "Always be not yourself, if you are Fernando Torres." On the assumption that *be not yourself* is substituted for accurately by *be someone else*, the second sentence in (18a) reinforces uniqueness. For entailments, this type of reinforcement appears redundant, as shown by (18b), where the entailed *if not* direction is restated in the second sentence.

Secondly, uniqueness can be questioned without creating contradiction. The example in (19) provides a natural case of this.

(19) a. "The answer is no unless you ask. If you do ask the answer might be no."
b. *Compare:* The answer is no unless you ask. #If you don't ask the answer might be yes.

Uniqueness, for (19a), holds that in any relevant situation where you do ask, the answer will be yes. This is questioned felicitously by the second sentence. On the other hand, questioning the *if not* entailment in (19b) poses a contradiction and is highly infelicitous. This contrast also supports pragmatic status for uniqueness.

Descriptively speaking, uniqueness can be classified as a generalized conversational implicature (GCI; see e.g. Levinson 2000). Example (20) suggests that it is not presupposed by *unless*, since we can felicitously use an *unless*-statement where uniqueness has been explicitly suspended. In addition, (21) goes against a conventional implicature account, since it shows that uniqueness is backgroundable, where conventional implicatures typically are not (see Potts 2005).

(20) a. The student might not fail if he studies, but he will fail unless he studies.
b. *Compare:* ?There might not be a student, but he will fail unless he studies.

(21) a. John won't fail if he studies. He will fail unless he studies.
b. *Compare:* John is a student. John, ?the student, will fail unless he studies.

Finally, the regularity of the inference to uniqueness (as attested by attempts to capture it in a semantic treatment; see Comrie (1986) as well as von Fintel and Leslie) stands against classifying it as a particularized conversational implicature. It is better described, alongside GCIs *à la* Levinson, as a default inference that "captures our intuitions about preferred or normal interpretations" (p.11).

This classification is supported additionally by the striking resemblance that uniqueness/bicondionality bears to the well-known GCI of **conditional perfection** (Geis and Zwicky 1971). Conditional statements such as (22) are known to often be interpreted as if they were biconditional, and this inference also shows the properties of defeasibility, questionability, and reinforceability.

(22) I'll give you five dollars if you mow the lawn.
$\rightsquigarrow$ *I'll give you five dollars if and only if you mow the lawn.*

In addition, both inferences are *nonconventional*, in the sense that they are not encoded and do not attach in all circumstances – consider, for instance, negative contexts for *unless* and example (23) for *if*-conditionals.

(23) If this cactus grows native to Idaho, it's not an *Astrophytum.*
$\not\rightsquigarrow$ *If and only if this cactus grows native to Idaho, it's not an* Astrophytum.

Levinson (2000) treats perfection as an instantiation of his *I-heuristic*, which is tied to the second maxim of quantity ("do not make your contribution more informative than is required"; Grice 1975) – roughly, the biconditional interpretation is drawn as the "most informative" of possible interpretations for a conditional. This particular classification is debated (for an overview, see van der Auwera 1997), but the emphasis it places on the default nature of the perfecting inference seems to capture the attachment properties of uniqueness as well.

4 Marbles and dots: an experimental investigation of *unless*

Taking stock, we have seen that *unless*-statements invoke a number of intuitive consequences, some of which apparently contradict one another. The idea that *unless* is a biconditional in positive contexts is challenged by its acceptable use in contexts where the *not if* direction is not intended, and the evidence this provides that *not if* behaves like a GCI. Similarly, while intuition strongly suggests that *unless* only makes stipulations about necessary conditions in negative contexts, we have seen that *unless*-statements seem at best infelicitous in across-the-board contexts. Broadly, while the idea underlying the exceptive account – that *unless* draws attention to the truth value of its consequent over the excepted set – is a running theme, it is extremely unclear what precisely is communicated in this regard, and whether it is communicated by semantic or by other means.

This highlights the inadequacy of relying solely on intuition in developing an account of *unless*. The comparative status of the various intuitions detailed above cannot be assessed in the absence of data on a large scale, and thus an experimental investigation seems warranted. In this section, I report on the results of two such investigations. The first, which is limited to an investigation of *unless* under universal quantifiers, sets out to compare the empirical reality of the predictions made by (both versions of) the exceptive account against the concerns raised in section 3.3. This study also provides some insight into the nature of the perceived biconditionality/unidirectionality split, and the original question of how *unless* does (and does not) diverge from *if not*-conditionals. These last two questions are explored further in the second experiment, which examines the behavior of *unless* under the non-universal quantifiers *most, some* and *few*. The less precise nature of these quantifiers allows a better assessment of the role played by pragmatics in certain aspects of *unless*.[4]

4.1 Experiment 1: Universal quantifiers

4.1.1 Experiment design: materials, methods, and participants

This experiment used a forced-choice true/false paradigm. Participants were shown a picture display of twenty marbles, each of which was either red or blue, and asked to decide whether a statement about the display was true or false (Figure 1 is the sample display shown during instructions). During the experiment itself, failure to select either "true" or "false" option prompted a message which informed participants that they would have to choose before being allowed to proceed to the next trial.

The design incorporated three independent variables. First, both *unless* and *if not* conditionals were used. Second, these appeared in each of three quantificational contexts: *every*, *no*, and a "bare" context, which, on both versions of the exceptive account, is taken to include a covert universal modal quantifier. The test statements are given in (24)-(26).

(24) a. Every marble has a dot unless it is [target colour].
b. Every marble has a dot if it is not [target colour].

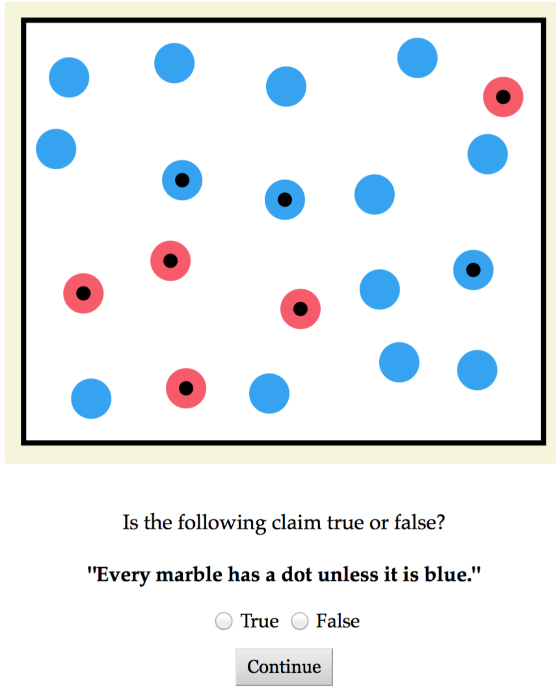
(25) a. No marble has a dot unless it is [target colour].
b. No marble has a dot if it is not [target colour].

(26) a. The selected marble has a dot unless it is [target colour].
b. The selected marble has a dot if it is not [target colour].

For the "bare" condition, the display was accompanied by this additional statement: "A marble is selected at random from the collection pictured above. Before you can look at the selected marble, or the remaining marbles in the box, you hear a claim about the marble that has been chosen." This appeared between the display and the statement to be evaluated as true or false. In this context, the covert universal quantifier over situations represents quantification over possible selections of marble, and consequently should result in

[4] Experiment 1, and the interpretation of the results presented here, are essentially as presented in Nadathur and Lassiter (2014); Experiment 2 and its results are an extension from that work.

Figure 1: Sample display



the same judgements as the positively-quantfied statements in (24). This condition was included to control for the possibility that this prediction would not be upheld.

The last independent variable was the proportion of marbles in the target colour which had dots in the display. This was varied evenly from amongst $\{0, 0.2, 0.4, 0.6, 0.8, 1\}$. In total, these three variables (conditional, quantifier, and target dot proportion) resulted in a set of $2 \times 3 \times 6 = 36$ test items.

To increase the variety of displays shown to participants, the target colour was also varied at random between red and blue, and the ratio of target to non-target marbles was also randomly selected from {5:15, 10:10, 15:5}. In each display accompanying a test statement, the dot-bearing proportion of non-target marbles was set to satisfy the minimal truth conditions shared by *unless* and *if not*, as represented by $Q[C \wedge \neg R]M$ (the first conjunct of both exceptive formulas). This was done to avoid collecting uninformative "false" responses, and the possibility that it would result in an unbalanced proportion of true sentences was offset by including fillers that were necessarily false. In total, there were six possible (but truth-conditionally identical) possible displays for each of the 36 test items.

A number of filler statements were also included. These varied according to three parameters: quantifier condition (*every, no* and bare), whether they mentioned red, blue, or only dots, and the type of construction used. The last parameter was varied between *if*-conditionals, single-clause quantified statements, and existential statements using *there*. This resulted in a space of $3 \times 3 \times 3 = 27$ possible filler sentences; examples of each variable are given in (27).

(27) a. [bare, red, *if*]: The selected marble has a dot if it is red.
b. [*every*, dot, single-clause]: Every marble has a dot.
c. [*no*, blue, *there*]: There are dots on no blue marbles.

Filler displays also had randomly-varied red:blue marble ratios (5:15, 10:10, or 15:5), and the dot-bearing proportion for both colours was selected at random from the set $\{0, 0.2, 0.4, 0.6, 0.8, 1\}$. Consequently, it was possible for a given filler statement to occur with any one of a total of $3 \times 6 \times 6 = 108$ possible displays.

Participants were recruited using Amazon's Mechanical Turk platform. At the end of the experiment, they were asked about their linguistic background, including what languages they spoke natively. After removing those who were not native speakers of English, 155 participant responses were taken as data. Each of these responses incorporated 48 trials: 24 randomly-selected test trials, interspersed with 24 randomly-selected fillers. As a result of randomization, the test sentences were not seen by a uniform number of people, but each was seen by a minimum of 80 participants (the average was 103.33).

4.1.2 Predictions

Both von Fintel's and Leslie's formulations of the exceptive account make clear predictions about the truth-value of the *unless*-statements in each test condition. These are given (informally) by quantifier in (28)-(30). Table 1 gives the predictions by experimental condition (F = von Fintel 1992; L = Leslie 2008).

(28) Every marble has a dot unless it is blue.
F/L: TRUE just in case *all red marbles have dots and no blue marbles have dots.*

(29) No marble has a dot unless it is blue.
F: TRUE just in case *no red marbles have dots and all blue marbles have dots.*
L: TRUE just in case *no red marbles have dots.*

(30) The selected marble has a dot unless it is blue.
F/L: TRUE just in case *all red marbles have dots and no blue marbles have dots.*

Table 1: Predictions for *unless* by condition

		Target dot proportion					
		0	0.2	0.4	0.6	0.8	1
every	**F**	T	F	F	F	F	F
	L	T	F	F	F	F	F
no	**F**	F	F	F	F	F	T
	L	T	T	T	T	T	T
bare	**F**	T	F	F	F	F	F
	L	T	F	F	F	F	F

As noted earlier, I take *if not* to have the representation in (31), and this is expected to be true in all test conditions (see (32)).

(31) $Q[C]M$ *if not* $R := Q[C \wedge \neg R]M$

(32) Every/no/the selected marble has a dot if it is not blue.
TRUE just in case *all/no/all red marbles have dots.*

One of the objectives of this experiment was to compare the predictions of the exceptive account against the challenges raised in section 3.3. In that section, I suggested first that *unless*-statements will be rejected in across-the-board contexts, and secondly that *unless*-statements are never biconditional, including in the positive (*every*, bare) contexts.[5] If both of these suggestions are correct, then *unless*-statements will be judged false if and only if their consequents hold for all of the marbles in the display, and not just those that do not meet the description of the excepted set. These predictions are given by quantifier in (33)-(35), and by condition in Table 2.

[5]Strictly speaking, I have suggested that biconditionality is a strong pragmatic inference associated with *unless*, and would predict some experimental effect of this. For clarity, however, the predictions given in section 4.1.2 just reflect the proposed truth-conditional content. This will be revisited in the discussion and in section 5.

(33) Every marble has a dot unless it is blue.
TRUE just in case *all red marbles have dots and not all blue marbles have dots.*

(34) No marble has a dot unless it is blue.
TRUE just in case *no red marbles have dots and at least one blue marble has a dot.*

(35) The selected marble has a dot unless it is blue.
TRUE just in case *all red marbles have dots and not all blue marbles do.*

Table 2: Alternative predictions for *unless* by condition

	Target dot proportion					
	0	0.2	0.4	0.6	0.8	1
every	T	T	T	T	T	#/F
no	#/F	T	T	T	T	T
bare	T	T	T	T	T	#/F

4.1.3 Results

The graphs in Figures 2-4 represent the experimental results by quantifier: Figure 2 shows the results across the *every* conditions, Figure 3 across the *no* conditions, and Figure 4 across the bare conditions. In each graph, the x-axis gives the proportion of target marbles with dots, and the y-axis gives the fraction of participants who selected "true" for a given test item. The red line gives the results for *unless*, and the blue line *if not*. Error bars show 95% binomial confidence intervals.

Figure 2: Results in the *every* conditions

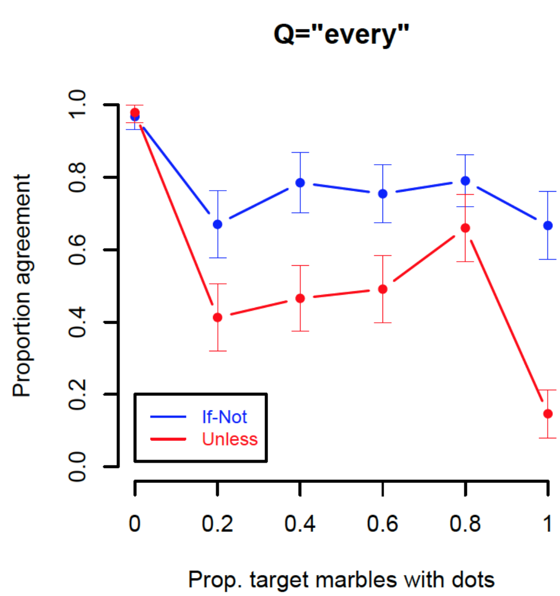


Table 3 gives the numerical data for each condition. N represents the total number of participants who responded to a given item, and given errors are again 95% binomial confidence intervals. All cells for

Figure 3: Results in the *no* conditions

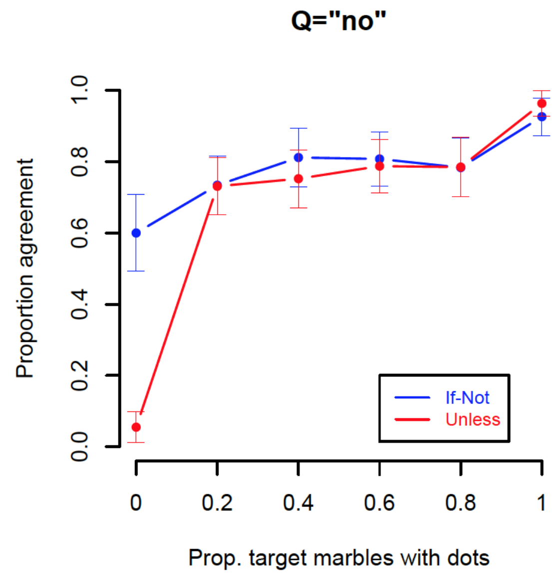
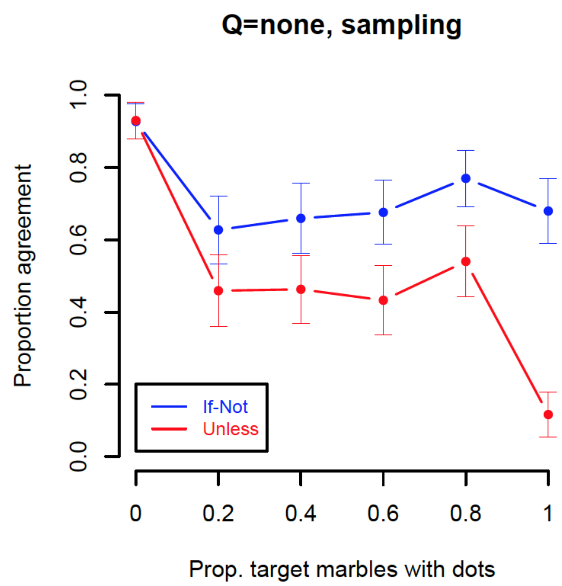


Figure 4: Results in the bare conditions



which the upper limit of the confidence interval falls below 50% agreement are marked in grey; all such cells empirically also fail to intersect 25% agreement. (*N.B.:* The results indicate that the expected parallelism between the *every* and bare test items is upheld. This is a useful result, but will not bear significantly on the discussion here, so I restrict my focus to the results of the *every* and *no* conditions.)

Table 3: Numerical data for Experiment 1

			every		*no*		bare	
			N	% agree	N	% agree	N	% agree
Prop. target marbles with dots	*If not*	*0.0*	93	96.8 ± 3.6	80	60.0 ± 10.7	110	92.7 ± 4.9
		0.2	100	67.0 ± 9.2	109	73.4 ± 8.3	102	62.7 ± 9.4
		0.4	93	78.5 ± 8.4	85	81.2 ± 8.3	91	65.9 ± 9.7
		0.6	110	75.5 ± 8.0	104	80.8 ± 7.6	108	67.6 ± 8.8
		0.8	124	79.0 ± 7.2	97	78.4 ± 8.2	113	77.0 ± 7.8
		1.0	96	66.7 ± 9.4	95	92.6 ± 5.3	103	68.0 ± 9.0
	Unless	*0.0*	95	97.9 ± 2.9	109	5.5 ± 4.3	100	93.0 ± 5.0
		0.2	109	41.3 ± 9.2	119	73.1 ± 8.0	98	45.9 ± 9.7
		0.4	116	46.6 ± 9.1	109	75.2 ± 8.1	108	46.3 ± 9.4
		0.6	112	49.1 ± 9.3	113	78.8 ± 7.5	104	43.3 ± 9.5
		0.8	100	66.0 ± 9.3	93	78.5 ± 8.4	100	54.0 ± 9.8
		1.0	109	14.7 ± 6.6	110	96.4 ± 3.5	103	11.7 ± 6.2

At face value, these results match neither the predictions made by the exceptive account nor the alternative predictions I provided in Table 2. In particular, there are a number of points in both *every* and *no* graphs, and for both *unless* and *if not* conditionals that do not reflect unanimity (or near-unanimity) from participants.

Unanimous agreement occurs for both conditionals at target proportion 0 under *every* (that is, in the fully-biconditional or "complete-split" condition where all of the non-target marbles have dots, but none of the target marbles do), but agreement drops across proportions 0.2-1 for *if not*; the rates here are statistically below 100% and above "maximal uncertainty" (50%, the expected rate if participants are guessing at random). It drops even lower across the same proportions for *unless*, plunging to near-zero at target proportion 1 (the across-the-board context for this quantifier) from maximal uncertainty on 0.2-0.8.

In the *no* conditions, both conditionals receive unanimous agreement at target proportion 1 (the complete-split context). Again, agreement drops for both conditionals on the middle range of proportions (0.2-0.8), and in this case both *if not* and *unless* have agreement rates statistically below 100% and above maximal uncertainty. *If not* stays in this range in the across-the-board context for this condition (target proportion 0), but *unless* drops to near-zero here.

The interesting points these data raise can be separated into three major puzzles, which I refer to here as (A), (B), and (C). Puzzle (A) is posed by the categorical divergence of *unless* from *if not* in across-the-board contexts. Puzzle (B) is the degraded but non-zero agreement for both types of conditionals in the middle range (and for *if not* in the across-the-board contexts as well). The third puzzle, (C), is the reliable but non-categorical divergence between *if not* and *unless*, specifically with respect to the effect size difference in positive vs negative contexts. The following analysis of these data attempt to provide an explanation for these puzzles; of course, any adequate account of the meaning of *unless* will need to do so as well.

4.1.4 Discussion

The results of Experiment 1 falsify the predictions of both exceptive accounts. Both von Fintel and Leslie claim that *unless* is semantically biconditional in positive quantificational contexts, and we thus expect that *unless* under *every* should be judged as false as soon as the proportion of dot-bearing target marbles goes above 0. Instead, although agreement drops on proportions 0.2-0.8, participants do not categorically rate the *unless* test statements as false until the target proportion reaches 1 (the across-the-board context). In particular, the sharp comparison between agreement rates on the middle range and near-null agreement at 1 suggests that *unless* ought not to be described as truth-conditionally false on the middle range: this contradicts biconditionality as semantic content. In the *no* condition, von Fintel's prediction of biconditionality is falsified for much the same reason: *unless* is degraded but not categorically false on the middle range (as compared to the near-null response at target proportion 0). Leslie's predictions for the negative cases are also incorrect, but for a different reason. Her account gives *unless* the same truth conditions as *if not* in these cases, and so the two would be expected to pattern together across all target proportions, but we find instead a categorical split in the across-the-board context (target proportion 0).

Central to this interpretation of the data is the assumption that a degraded but strongly nonzero acceptance rate is indicative of the truth of the test sentence. This is motivated by the expectation that, at least in some cases, participants will invoke pragmatic reasoning in making their true/false judgements; in the context of this experiment, I consider it to be significantly more likely that this reasoning exerts a "downward" effect on a minority of participants (that is, that is causes these participants to reject semantically true sentences) than that it causes a majority of participants to accept a sentence whose truth conditions are not met.[6] While I do not make the claim that "upwards" effects are generally nonexistent, there are no obviously relevant aspects of pragmatics in this case that could reasonably have resulted in the drastic improvements (from near-null to 50% or higher) that would be required to reconcile the experimental data with either exceptive account. On the other hand, I have already discussed the possible influence of a biconditionality implicature which might be expected to exert a "downward" pressure producing the degraded agreement in question: I discuss this in detail below.

This suggests that the results of Experiment 1 corroborate the predictions in Table 2. In particular, Table 2 only predicts null (or near-null) agreement for *unless* in the across-the-board contexts for each quantifier, and the gray cells in Table 3 (those cells whose confidence intervals failed to intersect 50% or higher agreement) pick out precisely these contexts. The truth conditions for *unless* are expected to be satisfied in all other contexts: if the degraded middle-range response is due to pragmatically-driven rejection of true statements, the data from Experiment 1 meet this expectation as well. Put another way, the results in Table 3 align very well with what is expected if (a) across-the-board contexts are the exclusive contexts in which *unless* is false or inappropriate and (b) biconditionality is associated with *unless* as a pragmatic inference rather than an entailment – this is precisely what I suggest in section 3.3. These observations offer immediate explanations for Puzzles (A) and (B), and suggest a direction for the explication of Puzzle (C).

Puzzle (A): Across-the-board contexts

The observation that *unless* is bad in across-the-board scenarios such as those described in section 3.3.1 accounts immediately for Puzzle (A). As noted above, the experimental conditions which received null or near-null agreement were precisely those conditions in which the generalization that "Q-many marbles have a dot" held across the entire domain of marbles, and not just the non-target marbles. In general, we can describe an **across-the-board** (henceforth, AtB) scenario simply by the formal expression in (36), where we retain the abstract variables used by von Fintel and Leslie in formulas (7) and (10), respectively.

(36) **AtB generalization:** $Q[C]M$

Where Q is a (positive or negative) universal quantifier, this generalization is logically equivalent to the "split" version in (37), which divides the quantification by the partition of C induced by a subset $R \subseteq C$.

[6] In further support of these claims, I note that the notion of "downwards" pragmatic pressure is by no means novel. The idea behind scalar implicatures such as the well-known inference from *some* to *not all* is that a competent language user will often interpret the weaker item *some* as indicating the negation of the stronger item *all*; that this reasoning in fact causes naive subjects to reject the weaker item in contexts where the stronger one holds has been demonstrated experimentally by Papafragou and Musolino (2003), Bott and Noveck (2004), Breheny et al. (2006), and Degen and Tanenhaus (2011), among others.

(37) **Split AtB generalization:** $Q[C \wedge \neg R]M \wedge Q[C \wedge R]M$

An *unless* conditional requires the first conjunct to be satisfied, but is bad when the second holds as well. This leads at once to the following revised proposal for *unless*.

(38) $Q[C]M$ *unless* $R := Q[C \wedge \neg R]M \wedge \neg \mathbf{Q}[\mathbf{C} \wedge \mathbf{R}]\mathbf{M}$

This preserves the first (*if not*) conjunct of both exceptive accounts, but eliminates the biconditionality requirement and replaces it with a statement encoding the requirement that the AtB generalization fails (in boldface). I refer to this new conjunct as the **negative across-the-board** (negAtB) clause. It is worth noting that in the context of universal quantifiers, it would have been sufficient to express the negAtB clause as the negation of (36); I have made the deliberate decision to encode it instead via the second conjunct of (37) in order to rule out AtB contexts for non-universal quantifiers as well.[7]

So far, I have referred to AtB contexts as "bad," "inappropriate," or "unacceptable" rather than as "false." This is due to the sense, expressed indirectly in section 3.3.1, that the negAtB clause has the flavour of a precondition or a felicity condition permitting the use of *unless* – that is, the sense that the negAtB clause may in fact represent a presupposition associated with the use of *unless*. The forced-choice true/false paradigm used in Experiment 1 lacks sufficient nuance to distinguish between an entailment or presupposition classification for this clause. Naive subjects may judge sentences carrying failed presuppositions (e.g. *The King of France is bald*) to be false (see Lasersohn 1993, Abrusán and Szendröi 2013, Schwarz 2014); thus the low acceptance rate for AtB contexts could equally well reflect the failure of a negAtB presupposition as the violation of an equivalent truth condition.

Regardless of the possible presuppositional status of the negAtB clause, however, it is evident that this condition is necessarily part of an account of *unless*. It explains the pattern of data comprising Puzzle (A) immediately, by uniquely picking out those points at which *unless* (but crucially not *if not*) received near-zero agreement rates.

Puzzle (B): Biconditionality implicatures

Puzzle (B) pertains to the reduced (but nonzero) acceptance rate for both *unless* and *if not* conditionals across the middle range of target dot proportions, as well as for *if not* in the AtB contexts. I have argued that this pattern requires a pragmatic explanation, and the discussion in section 3.3.2 provides the necessary pragmatics. In particular, I claim that the pattern in Puzzle (B) is produced by biconditionality implicatures that affect both types of conditionals: conditional perfection for *if not* and uniqueness for *unless*.

Recall example (15) from section 3.3.2:

(15) a. "Mantou is always late unless she's already out before we meet."
b. "Mantou is always late unless she's already out before we meet, but she's often just less late then."

I noted earlier that, in the absence of the continuation in (38b), (38a) tends to carry a biconditional interpretation. The middle-range responses to *unless* similarly suggest that some participants have a preference for interpreting examples like (39a) and (40a) as biconditional; when these utterances are produced in contexts that instead validate (39b) and (40b), participants who automatically generated the implicature sometimes reject them. This parallels the observed rejection of *some* in contexts verifying *all* (see Degen and Tanenhaus 2011, among others).

(39) a. Every marble has a dot unless it is blue.
b. Every marble has a dot unless it is blue, but some blue marbles have dots.

[7] The discussion in section 3.3.1 suggests that capturing the non-universal contexts is essential. For a quantifier like *most*, it is possible for (36) to be true while the second conjunct of (37) is false (for instance, if the excepted set comprises a very small fraction of the total domain C); what makes *unless* unacceptable is if Q-many of the excepted set satisfy the predicate M.

(40) a. No marble has a dot unless it is blue.

b. No marble has a dot unlss it is blue, but some blue marbles do not have dots.

Conditional perfection for *if* (and *if not*) conditionals is well documented (for an overview, see van der Auwera 1997), and would generate much the same effect as described here for *unless*. Subjects who generate perfection will sometimes reject *if not* statements in contexts where they are strictly true, but biconditionality fails. This includes AtB contexts as well as the middle range (0.2-0.8), since *if not* is not affected by the negAtB clause.

Puzzle (C): Grades of biconditionality?

Puzzle (C) centers around the reliable but non-categorical divergence between *if not* and *unless* over the target dot proportions 0.2-0.8, and particularly around the fact that this effect size is significantly greater in the positive contexts than the negative ones. This is more perplexing than (A) or (B), and no immediate solution is offered by the discussion up to this point.

Using the logic I have so far employed, these effects must be explained pragmatically. Consequently, there are two possible sources for the pattern in (C). One, the lower acceptance rate for *unless* in the positive condition (as compared to the negative one) might be due to a second pragmatic pressure (separate from and in addition to the biconditionality implicature), which operates only in the presence of the positive quantifier: I do not have any speculations as to what such an inference might comprise. A second possibility is that the biconditionality implicature is simply stronger in the positive cases than it is in the weaker ones.

I favour the second option, and take the view that Puzzle (C) is most likely related to biconditionality implicatures driving the data patterns in (B). This view is supported by the fact that the empirical positive/negative split in the results of Experiment 1 to a certain extent validate the intuition (cf. Leslie) that *unless* is biconditional in positive contexts but not in negative ones. In addition, although the *unless* and *if not* agreement rates in the negative conditions do not obviously diverge in the manner of the positive conditions, *unless* agreement shows a tendency to be lower than *if not* here as well (this is supported as well by more detailed statistical tests, but I have not discussed these here as they add little to the results already reported). This supports a view on which *unless* is in general more strongly biconditional than *if not*, but the pressure exerted by this implicature is weakened in the negative cases (or, conversely, strengthened in the positive ones).

Whether or not this is the correct line to take in approaching Puzzle (C), the pattern of results here demonstrate one thing clearly: the original positive/negative split is a pragmatic matter, rather than a semantic one. Over the middle range of target dot proportions, the difference between *unless* in positive as opposed to negative contexts is non-categorical. Consequently, the split over biconditionality is not a potential problem for compositionality (*à la* Higginbotham), and should therefore be omitted from any semantic account of *unless*.

4.2 Interim summary

Based on the results of Experiment 1, I have argued for four main points. First, neither version of the exceptive account offers the correct semantics for *unless*: empirically, *unless* is not semantically biconditional under either positive or negative quantification (and it does not uniformly pattern with *if not* under *no*). Second, *unless* categorically diverges from *if not* at precisely those points where Q-many of the excepted set verify the predicate M: the negative across-the-board clause therefore plays a role in accounting for *unless*. Third, biconditionality implicatures affect both *unless* and *if not*, and account for the degraded but non-zero acceptance rates in experimental conditions that deviate from biconditionality (barring the AtB contexts for *unless*). Finally, the perception of a positive/negative biconditionality split is a reflection of pragmatic interpretation, not the semantic truth conditions of *unless*.

These points, more generally, argue for a view of *unless* as sensitive to three propositions: *if not*, the negAtB clause, and a perfection or *not if* clause.

(41) $Q[C]M$ *unless* $R :=$

if not	$Q[C \wedge \neg R]M$
negAtB	$\neg Q[C \wedge R]M$
perfection/not if	$Q[C \wedge R]\neg M$

I take it to be uncontroversial that *if not* is an entailment of *unless*, and have argued that biconditionality/perfection is an implicature. The status of the negAtB clause, however, remains undecided: although it picks out precisely those points at which a strong majority of participants judged *unless* to be false, I have argued (in section 3.3.1 and elsewhere) that it has the character of an appropriateness or felicity condition on the use of *unless* and may therefore be best classified as a presupposition.

Presuppositional content is usually identified on the basis of its projective behavior: standard projection tests include the "family of sentences" diagnostic (Chierchia and McConnell-Ginet 1990). In particular, if a proposition S carries the presupposition r, then r is also presupposed when S is negated at the sentential level, converted to an interrogative, or embedded in the antecedent of a conditional. (42) shows this for the standard "King of France" example.

(42)
a. The King of France is bald.
b. The King of France is not bald. (*alt:* It is not the case that the King of France is bald.)
c. Is the King of France bald?
d. If the King of France is bald, his crown must hurt him.

(42a) presupposes the existence of a (unique) King of France, and we see that (42b)-(42d) commit the speaker to this belief as well.

The family of sentences diagnostic, unfortunately, is not straightforward for *unless.* Like conditionals as a class, *unless*-statements resist embedding. Consider sentential negation, to start:

(43)
a. Not every marble has a dot unless it is blue.
b. It is not the case that every marble has a dot unless it is blue.

It is extremely difficult (if not impossible) to interpret the negation in (43a) as sentential, rather than as negation of the consequent clause. The less fluent negation in (43b) is an attempt to get around this, but is no less difficult to interpret.[8] The difficulty in determining even the entailed content of these sentences more or less precludes any ability to reason about their projected content.

Similar problems arise with the other members of the family. It is relatively straightforward to transform an *unless*-statement into an interrogative, as in (44), but I do not think it is clear whether or not something along the lines of the negAtB clause projects.

(44) Does every marble have a dot unless it is blue?

It seems to me that this is not a sensible question to ask if the questioner has reason to believe that the AtB condition holds (i.e., that every marble has a dot), but I have no pretheoretic intuitions about situations in which the speaker has no prior knowledge. This uncertainty seems due to the dual-clause structure of an *unless*-sentence, which is what causes the complications in examples (43a) and (43b) as well. This problem is exacerbated by embedding *unless* under a conditional, as in (45).

(45) If every marble has a dot unless it is blue, the blue marbles are worth more money.

Here, it is very difficult to separate the consequent of the *if*-conditional from some notion of the exceptionality of the set of blue marbles (that is, (45) seems to suggest that it is some dot-related difference on the blue marbles that results in their greater value). (45) gives rise to at least the negAtB inference, if

[8] Indeed, the easiest interpretation is still as negation of only the consequent clause, which is related to the issue of *conditional excluded middle* (see von Fintel and Iatridou 2002, Leslie 2008).

not biconditionality, but it does not seem possible to determine *ad hoc* whether this is due to pragmatic implicature on the outer conditional or projective content from the inner *unless*.

What other possible routes exist for testing presuppositional status? A potentially relevant difference between presuppositions and entailments has to do with accommodative behavior. Sentences carrying presuppositions can, under certain circumstances, be felicitous even though their content has not previously been explicitly entered into the stock of common knowledge: this relies on the possibility of incorporating the content of the presupposition into the set of background assumptions without causing the set to become inconsistent. For instance, the assertion in (46) is unproblematic in a discourse situation where the listener was not aware that the speaker has a sister, as long as speaker and hearer had not previously coordinated upon the fact that the speaker has no sister.

(46) My sister is taking me out for dinner tonight.

This suggests a possible avenue for testing the status of the negAtB clause. Non-universal quantifiers such as *most* and *few* are far less rigid in their interpretations than universal ones: relevant context might render any domain fraction greater than half as a suitable representation for *most*, but only proportion 1 can validate *every*. It is possible that this increased flexibility will increase the potential for accommodation of a quantified *unless*-statement in an AtB context if the quantifier in question is non-universal.

This suggests a second experiment, this time using non-universal quantifiers. Independently of the status of the negAtB clause, this offers a means of verifying the claim that *unless* can in fact occur under non-universal quantifiers (contra von Fintel).

4.3 Experiment 2: Non-universal quantifiers

4.3.1 Experiment design: materials, methods, and participants

The basic setup for Experiment 2 was the same as for Experiment 1. Experiment 2 also used the forced-choice true/false paradigm. Participants were shown displays of red and blue marbles, as before, and asked to decide whether one-sentence descriptions were true or false.

Experiment 2 had four independent variables. As in Experiment 1, both *unless* and *if not* were used, in each of three quantificational contexts: *most*, *some*, and *few*. The test statements are given in (47)-(49).

(47) a. Most marbles have a dot unless they are [target colour].
b. Most marbles have a dot if they are not [target colour].

(48) a. Some marbles have a dot unless they are [target colour].
b. Some marbles have a dot if they are not [target colour].

(49) a. Few marbles have a dot unless they are [target colour].
b. Few marbles have a dot if they are not [target colour].

Again, as before, the proportion of target marbles with dots was varied evenly from $\{0, 0.2, 0.4, 0.6, 0.8, 1\}$. The fourth variable was the proportion of *non-target* marbles that had dots: this varied evenly from *high, mid* and *low* parameters specific to each of the three quantifiers ($\{0.6, 0.8, 1\}$ for *most*, $\{0.2, 0.4, 0.6\}$ for *some*, $\{0, 0.2, 0.4\}$ for *few*). This additional variable was included to leave room for expected participant variance as to the "exemplar" proportion for each of the relevant non-universal quantifiers. In total, these four variables produced a space of $2 \times 3 \times 6 \times 3 = 108$ test items. Display variation in this experiment was limited to random selection of the target colour as either red or blue. In all cases, the ratio of red:blue marbles was 10:10.[9]

Experiment 2 also included non-test items, a number of which were designated as controls for the high, mid, and low conditions. As before, non-test statements used all three quantifiers, were either *if*-conditionals,

[9]This choice was made because each of the test sentences in (47)-49) has the plural pronoun *they* in the antecedent. The proportion 0.2 only picks out one marble from a set of 5, which might have thrown participants off due to the plural pronoun.

single-clause quantified statements, or *there*-existentials, and mentioned either red, blue, or only dots. Statements which asked about marble colour were designated as controls, and the displays for these items were generated with the target colour dot proportion varied only between the high, mid, and low conditions for the relevant quantifier. This enabled the independent collection of data regarding participants' overall acceptance of the high, mid, and low proportions as representative of the quantifiers *most, some* and *few*. The non-target dot proportion for controls was selected at random from $\{0, 0.2, 0.4, 0.6, 0.8.1\}$. The remaining non-test items (those which asked about dots) were fillers and were accompanied by displays for which both the red and blue dot proportions were selected at random from the set $\{0, 0.2, 0.4, 0.6, 0.8, 1\}$. In total, there were $3 \times 3 \times 3 = 54$ control items, each of which could appear with 6 different non-target proportions and either red or blue as the target colour. There were only $3 \times 3 \times 1 = 9$ fillers, but each of these could appear with $6 \times 6 = 36$ possible displays.

Participants were again recruited via Mechanical Turk. After *post hoc* filtering for native English status, 373 responses were taken as data. Each participant saw 24 randomly-selected test items, interspersed with 24 items randomly selected from an undifferentiated list of fillers and controls. Test items were seen by a minimum of 61 participants (average = 82.9), and control items were seen by a minimum of 199 participants (average = 221.6).

4.3.2 Predictions

I use the working proposal in (38), restated below, to generate predictions for Experiment 2.

(38) $Q[C]M$ *unless* $R := Q[C \wedge \neg R]M \wedge \neg Q[C \wedge R]M$

Predictions are given by quantifier condition in (50)-(52) and by experimental condition (with high, mid, and low conditions collapsed) in Table 4. The predictions rely on the assumption that each of the high, mid, and low conditions is an acceptable representation of the relevant quantifier; control data will permit refinement of this picture later.

(50) Most marbles have a dot unless they are blue.
TRUE just in case *most red marbles have dots and not most blue marbles do.*

(51) Some marbles have a dot unless they are blue.
TRUE just in case *some red marbles have dots and no blue marbles do.*

(52) Few marbles have a dot unless they are blue.
TRUE just in case *few red marbles have dots and not few blue marbles do.*

Table 4: Experiment 2 predictions for *unless* by condition

	Target dot proportion					
	0	0.2	0.4	0.6	0.8	1
most	T	T	T	#/F	#/F	#/F
some	T	#/F	#/F	#/F	#/F	#/F
few	#/F	#/F	#/F	T	T	T

As before, *if not* is expected to be true in all conditions:

(53) Most/some/few marbles have a dot if they are not blue.
TRUE just in case *most/some/few red marbles have dots.*

4.3.3 Results

Control results

Figure 5 shows the agreement rates for control sentences, by construction type and target dot condition (high, mid, or low). The leftmost graph has the results for *most*, the middle is *some*, and the rightmost is *few*. The x-axis represents target dot condition, from low to high, and the y-axis (as before) shows agreement rate. The green line represents *if*-conditionals, the purple line represents single-clause quantified statements, and the orange line represents *there*-existentials. As before, error bars give 95% binomial confidence intervals.

Figure 5: Control results, by quantifier condition

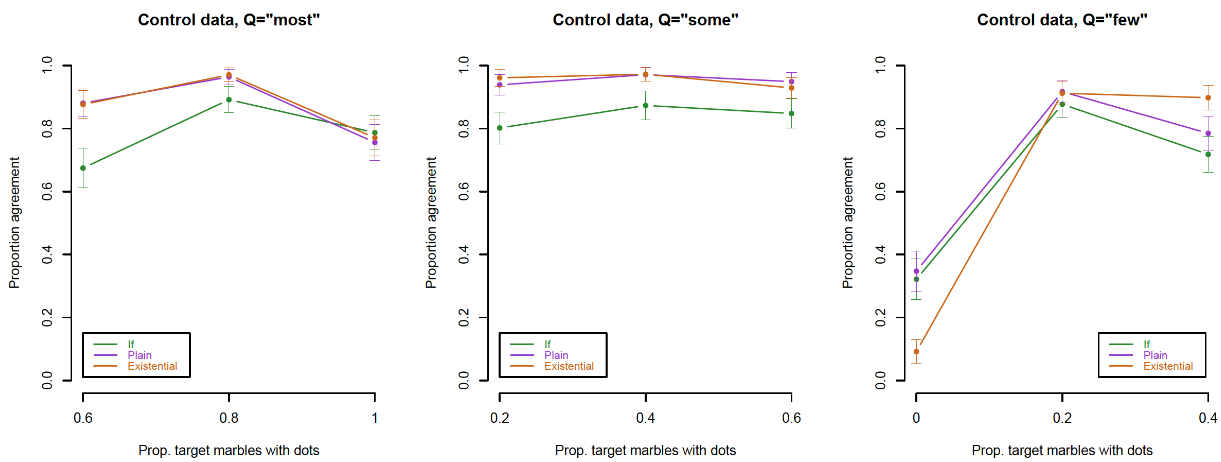


Table 5 gives the corresponding numerical data for each condition. N represents the total number of participants who responded to a given item, as before. All cells for which the upper limit of the binomial confidence interval on the agreement rate falls below 50% are marked in gray.

Table 5: Numerical data for Experiment 2 controls

			most		*some*		*few*	
			N	% agree	N	% agree	N	% agree
Target marble condition	*if*	low	215	67.4 ± 6.3	237	80.2 ± 5.1	199	32.2 ± 6.5
		mid	212	89.2 ± 4.2	205	87.3 ± 4.6	236	87.8 ± 4.2
		high	230	78.7 ± 5.3	230	84.8 ± 4.6	234	72.8 ± 5.8
	plain	low	226	88.1 ± 4.2	212	93.9 ± 3.2	216	34.7 ± 6.4
		mid	221	96.4 ± 2.5	241	97.1 ± 2.1	229	91.8 ± 3.6
		high	221	75.6 ± 5.7	215	94.8 ± 2.9	223	78.5 ± 5.4
	exist.	low	210	87.6 ± 4.5	205	96.1 ± 2.7	229	9.2 ± 3.7
		mid	237	97.1 ± 2.2	216	97.2 ± 2.2	216	91.2 ± 3.8
		high	205	77.1 ± 5.8	239	92.9 ± 3.3	225	89.8 ± 4.0

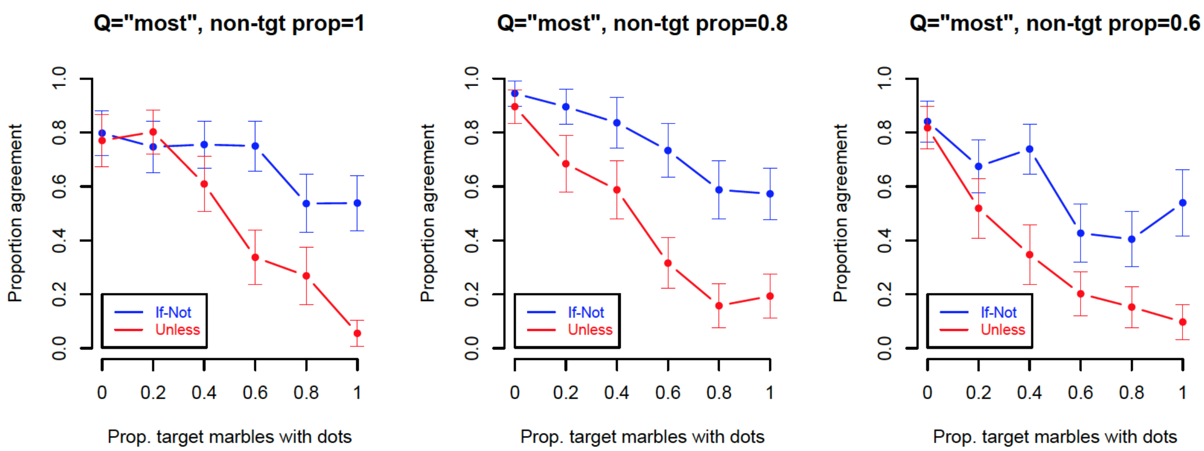
The control data provide information as to the fraction of respondents who accept a given target condition as an adequate representation of the quantifier in question. Consequently, this sets baseline acceptance against which to assess the relative strength of participant agreement on test items. For instance, the agreement rate for *if*-conditionals in *most*-low suggests that 65% or higher agreement should be treated as near-unanimous in this condition. I use the agreement rates for the *if*-conditionals as threshold markers, since these are the control items which most closely resemble the test items; these thresholds t are noted in the test data tables below, and agreement rates for which the upper limit of the confidence interval lies below half of t are marked in gray.

One point from the control data is worth noting individually, and this is the markedly low acceptance rate for *few* in the low condition, across all sentence types. This suggests that the low target proportion (0, in this case) is not really accepted as a representation of the quantifier *few*, and indicates that test results in this condition should be treated as unreliable.

Test results

Figures 6-8 give the results of Experiment 2 by non-target dot condition (left=high, middle=mid, right=low). The x-axis gives the proportion of target marbles with dots, and the y-axis is participant agreement. The red line again represents *unless*, and the blue *if not*. Binomial confidence intervals are shown.

Figure 6: *Most* results, by non-target dot condition



Tables 6-8 contain the corresponding numerical data. N is the total number of participants who responded to a given item. The threshold rates t are noted at the top of each column. Agreement rates for which the upper limit of the confidence interval falls below half of the threshold rate are marked in gray.

On the assumption that the threshold heuristic is reasonable, response rates that do not receive gray shading can be treated as reflecting participant agreement (that is, reflecting a test sentence that participants essentially find acceptable). On this view, the prediction that all *if not* items are acceptable is upheld. A comparison of the gray cells in Tables 6-8 with the gray cells in Table 4 reveals a few points of interest with respect to the predictions made for *unless*.

First, formula (38) predicts that *unless* under *most* will be rejected across the target proportions 0.6-1 in all three non-target dot conditions and accepted elsewhere. These expectations are met everywhere except in the high condition at target dot proportion 0.6. This point does meet with significantly lower agreement ratings than the lower target proportions in the same non-target dot condition, with the raw agreement rate at less than half of the condition threshold $t = 78.7$. However, the confidence interval here has nontrivial

Figure 7: *Some* results, by non-target dot condition

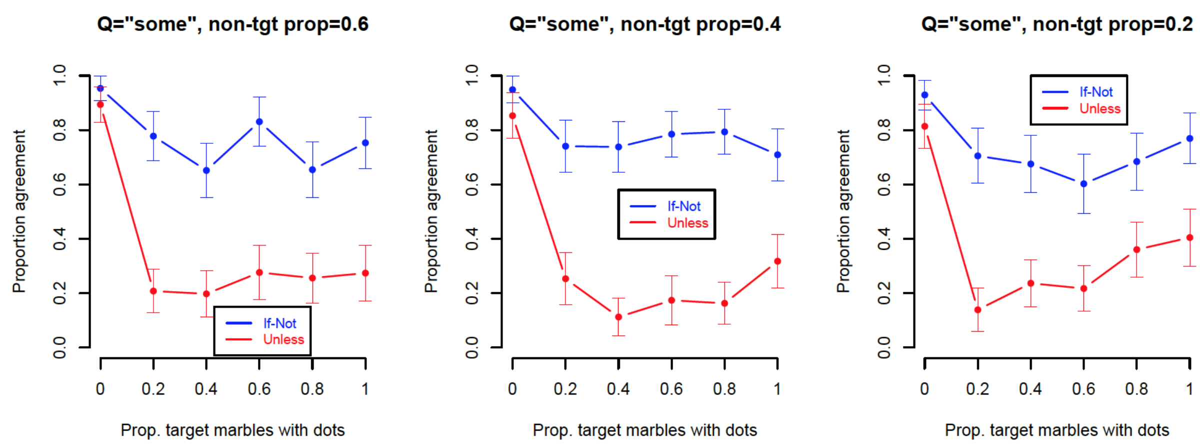


Figure 8: *Few* results, by non-target dot condition

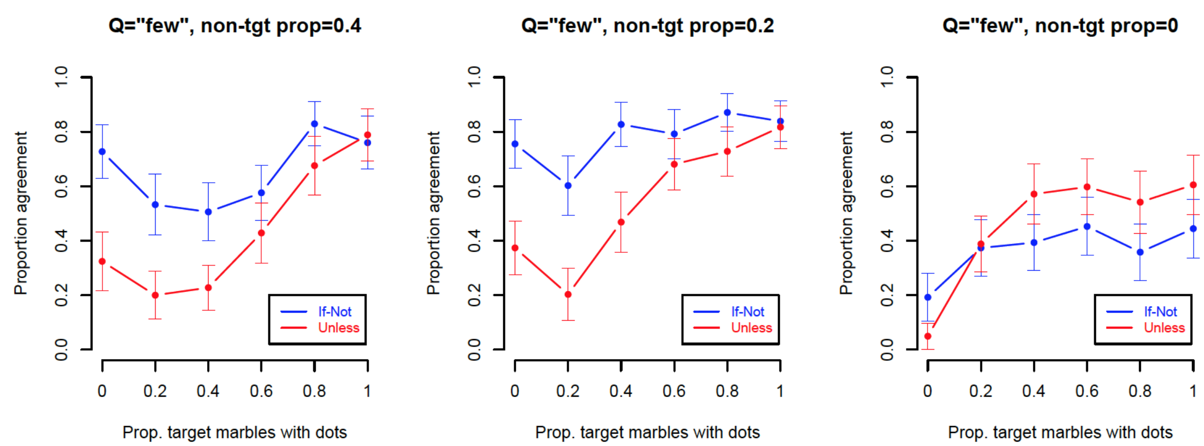


Table 6: Numerical data for *most*

			high ($t = 78.7$)		mid ($t = 89.2$)		low ($t = 67.4$)	
			N	% agree	N	% agree	N	% agree
Prop. target marbles with dots	*If not*	*0.0*	89	79.8 ± 8.3	91	95.5 ± 4.7	88	84.1 ± 7.6
		0.2	79	74.7 ± 9.6	86	89.5 ± 6.5	86	67.4 ± 9.9
		0.4	94	75.5 ± 8.7	61	83.6 ± 9.3	88	73.9 ± 9.2
		0.6	84	75.0 ± 9.3	75	73.3 ± 10.0	82	42.7 ± 10.7
		0.8	82	53.6 ± 10.8	80	58.8 ± 10.8	89	40.45 ± 10.2
		1.0	91	53.8 ± 10.2	103	57.3 ± 9.6	63	54.0 ± 12.3
	Unless	*0.0*	74	77.0 ± 9.6	96	89.6 ± 6.1	93	81.7 ± 7.9
		0.2	91	80.2 ± 8.2	76	68.4 ± 10.5	79	51.9 ± 11.0
		0.4	87	60.9 ± 10.2	80	58.8 ± 10.8	72	34.7 ± 11.0
		0.6	86	33.7 ± 10.0	95	31.6 ± 9.4	94	20.2 ± 8.1
		0.8	67	26.9 ± 10.6	76	15.8 ± 8.2	85	15.3 ± 7.7
		1.0	90	5.6 ± 4.7	93	19.4 ± 8.0	82	9.8 ± 6.4

Table 7: Numerical data for *some*

			high ($t = 84.8$)		mid ($t = 87.3$)		low ($t = 80.2$)	
			N	% agree	N	% agree	N	% agree
Prop. target marbles with dots	*If not*	*0.0*	86	95.4 ± 4.4	78	94.9 ± 4.9	85	92.9 ± 5.5
		0.2	81	77.8 ± 9.1	81	74.1 ± 9.5	78	70.5 ± 10.1
		0.4	89	65.2 ± 9.9	84	73.8 ± 9.4	77	67.5 ± 10.5
		0.6	65	83.1 ± 9.1	93	78.5 ± 8.3	78	60.3 ± 10.9
		0.8	81	65.4 ± 10.4	92	79.4 ± 8.3	76	68.4 ± 10.5
		1.0	81	75.3 ± 9.4	86	70.9 ± 9.6	78	76.9 ± 9.3
	Unless	*0.0*	85	89.4 ± 6.5	68	85.3 ± 8.4	86	81.4 ± 8.2
		0.2	101	20.8 ± 7.9	79	25.3 ± 9.6	72	13.9 ± 8.0
		0.4	86	19.8 ± 8.4	80	11.3 ± 6.9	93	23.7 ± 8.6
		0.6	76	27.6 ± 10.1	69	17.4 ± 8.9	92	21.7 ± 8.4
		0.8	86	25.6 ± 9.2	86	16.3 ± 7.8	86	36.1 ± 10.2
		1.0	73	27.4 ± 10.2	85	31.8 ± 9.9	84	40.5 ± 10.5

Table 8: Numerical data for *few*

			high ($t = 72.8$)		mid ($t = 87.8$)		low ($t = 32.2$)	
			N	% agree	N	% agree	N	% agree
Prop. target marbles with dots	*If not*	*0.0*	77	72.7 ± 10.0	90	75.6 ± 8.9	78	19.2 ± 8.7
		0.2	77	53.2 ± 11.1	78	60.3 ± 10.9	83	37.3 ± 10.4
		0.4	85	50.6 ± 10.6	81	82.7 ± 8.2	89	39.3 ± 10.2
		0.6	92	57.6 ± 10.1	77	79.2 ± 9.1	84	45.2 ± 10.6
		0.8	82	82.9 ± 8.1	93	87.1 ± 6.8	81	35.8 ± 10.4
		1.0	75	76.0 ± 9.7	93	83.9 ± 7.5	81	44.4 ± 10.8
	Unless	*0.0*	74	32.4 ± 10.7	91	37.4 ± 9.9	81	4.9 ± 4.7
		0.2	80	20.0 ± 8.8	69	20.3 ± 9.5	85	38.8 ± 10.4
		0.4	101	22.8 ± 8.2	79	46.8 ± 11.0	77	57.1 ± 11.0
		0.6	77	42.9 ± 11.0	94	68.1 ± 9.4	87	59.8 ± 10.3
		0.8	74	67.6 ± 10.7	92	72.8 ± 9.1	72	54.2 ± 11.5
		1.0	71	78.9 ± 9.5	93	81.72 ± 7.9	76	60.5 ± 11.0

overlap with the half-threshold mark, where the higher target proportions (marked in gray) do not: this point can be said to reflect maximal uncertainty in this condition, then, rather than majority agreement or majority rejection.

Under *some*, formula (38) predicted agreement only at target dot proportion 0 across all three non-target dot conditions (since the negAtB clause requires that no target marbles have dots). These expectations are matched at all points in the high and mid conditions, but fail at target proportions 0.8 and 1 in the low condition. The agreement rate for the aberrant target proportions under *some* also reflect maximal uncertainty: the raw rate are at or below 50% of the condition threshold ($t = 80.2$), but well overlap this mark when confidence intervals are taken into account.

The data under *few* are the most complex. Formula (38) predicts agreement on target proportions 0.6-1 across all three non-target conditions, and rejection on target proportions 0-0.4. These expectations are not met in any of the non-target conditions. In the high condition, participants seem to accept *unless* on the predicted 0.6-1 range, and the majority reject it on 0.2 and 0.4, but target proportion 0 reflects maximal uncertainty and does not lie below the half-threshold ($t = 72.8$) mark for this condition. In the mid condition, we have majority acceptance again on 0.6-1, but rejection only at 0.2; target proportion 0 reflects maximal uncertainty, and the agreement rate at 0.4 does as well, with raw agreement rate lying above the half-threshold ($t = 87.8$) mark for this condition, but significant overlap in confidence intervals. In the low condition, we get rejection only at target proportion 0; all other agreement rates are in fact at or higher than the threshold mark $t = 32.2$. This may be an effect of the majority rejection of 0 as an acceptable representation of the quantifier *few* – as noted above, the significantly degraded control response in this condition renders the test data in *few*-low unreliable.

4.3.4 Discussion

The motivation for this experiment was an investigation of the presuppositional status of the negAtB clause (the second conjunct of (38)). I argued that, if the negAtB clause in fact reflects a presupposition or admittance condition for *unless* (and not an entailment), we might expect to find evidence of accommodation.

How might this potential accommodation be reflected in the data? One view of accommodation (as per example (46)) involves the incrementation of the common ground (or set of background beliefs and assumptions, Stalnaker 2002) with the presupposition(s) of a proffered proposition. More broadly, this sort of accommodation can be seen as a reflection of co-operative communicative behavior in general, and in particular of a general communicative tendency to try to make sense of speaker's contributions.

In the context of Experiment 2, then, attempts to accommodate a presuppositional negAtB clause might lead a fraction of participants to accept *unless*-statements in AtB scenarios where they perceive a salient reasonable difference between the target and non-target dot proportions. For instance, we would not expect an effect where the dot proportion was identical across target and non-target marbles (since this precisely verifies the AtB condition). Consider the following *most* scenario, however: of the twenty display marbles, 6 of 10 blue marbles have dots and 8 of 10 red ones do as well. Participants might then conjecture that the (hypothetical) speaker regards the proportion 0.6 as unrepresentative of *most* by comparison to the higher proportion 0.8 on the red marbles, and these participants might therefore accept the description "Most marbles have a dot unless they are blue." This reflects the principle of cooperative communication that underlies many theories of accommodation: practically, accommodation of this sort would be expected to exert an *upwards* pressure on otherwise bad sentences, which could be sufficient to move them out of the strongly rejected "gray zone."

The results of Experiment 2 are compatible with a presuppositional view of the negAtB clause. I consider the *few* data first. A very literal interpretation of the negAtB clause would predict gray-zone agreement rates for *unless* at target proportion 0 in the high and mid conditions, but the data instead shows maximal participant uncertainty at these points. Looking back at the control data, it is clear that participants largely reject the proportion 0 as a representation of *few*: thus a picture in which none of the target marbles have dots might easily be interpreted as satisfying the negAtB requirement that *not few target marbles have dots*. Example gives the conditions for "Few marbles have a dot unless they are blue" according to (38).

(54) Few marbles have a dot unless they are blue.
Few red marbles have dots and not few blue marbles have dots.

In the absence of a display context, example (54) seems to suggest that more blue marbles have dots than red ones, but a participant trying to make sense of the given statement in for a display where, for instance, 4 of 10 red marbles have dots and no blue ones do might accommodate by interpreting 0 as a reasonable representation of *not few*. Crucially, both the high and mid conditions for the statement in (49a) reflect situations in which the proportions of target and non-target marbles with dots are noticeably different.

The other unexpected point of maximal uncertainty in the *few* data (target proportion 0.4 in the mid condition) also involves a display context which reflects this sort of salient difference between the target and non-target dot proportions. In a group of ten marbles, the proportion 0.4 is visually close to *half*, which seems to be a poor representation of *few*: this proportion was only accepted at a rate of 72.8% in control data, as opposed to a rate of 87.8% for proportion 0.2. The observable difference between the target and non-target proportions, then, leads to increased acceptability (via accommodation) for this test condition.

A similar effect holds for the *most* data. The only unexpected point of maximal uncertainty here is in the high condition, at target proportion 0.6. As for the *few* point described immediately above, this target proportion is visually similar to *half*, especially when compared to the non-target proportion 1 in the high condition. Thus the salient difference leads to increased acceptability for this test condition.

Finally, the two important points of maximal uncertainty under *some* occur in the low non-target condition, at target proportions 0.8 and 1. I have already noted the tendency of naive respondents to reject *some* as a description of situations in which *all* is verified; this implicature would be particularly salient in the low non-target condition, where *some* is represented by the proportion 0.2 (less than half). A similar argument holds for target proportion 0.8 in this condition (since the number of target marbles with dots is still more than twice as many as the number of non-target marbles with dots), and might be guided by lexical competition with *some*. Although it is not enough to extract it from the gray zone, we can also observe an increase in acceptance at target proportion 0 in the *some*-mid condition, presumably driven by the sharp difference in proportion between the target and non-target marbles.

To a certain extent, these interpretations the Experiment 2 data remains preliminary. I do not claim that these data uniquely support a presuppositional account of the negAtB clause, but rather argue that the data seem highly compatible with this view. More broadly, the notion that salient differences between the target and non-target proportions in fact drive the differentials in acceptance rates (and, specifically, account for the difference between gray zone acceptance and maximal uncertainty) suggests that the negAtB clause should not be interpreted with the same strictness as an entailment. I discuss this in greater detail in section 5.2.

One additional interesting pattern emerges from the Experiment 2 results. As in Experiment 1, *if not* and *unless* diverge in the conditions with non-unanimous acceptance rates (with *unless* rating lower than *if not* in all conditions except the unreliable *few*-low). This divergence is noticeable but non-categorical across the predicted white-zone proportions for both *most* and *few*. Linear mixed-effects regression reveals this effect to be significant for both quantifiers; what is of particular interest is that the effect size is greater under *most* than under *few*. This parallels the effect size difference between *every* and *no* from Experiment 1. Since *no* and *few* are downward-entailing where *every* and *most* are not, this suggests that the responsible pragmatic pressures (e.g. biconditionality implicatures) responsible for the difference between *unless* and *if not* are not only weaker under negative quantifiers, but in fact in downward-entailing contexts in general.

4.4 Summary of experimental investigation

Before moving forward with a revised theoretical proposal for *unless* in section 5, I summarize the desiderata for such an account that have emerged from the results of the experimental investigations detailed here.

First, an account of *unless* should not stipulate biconditionality as a semantic constraint (i.e. an assertion), although it should be compatible with biconditionality as a strong implicature. Secondly, the significance of the "across-the-board" condition (that *unless*-statements are bad in contexts where their consequents hold for the excepted set as well as the remainder of the domain) is demonstrated clearly by

Experiment 1. Experiment 2 supports a view of the negAtB clause as presuppositional, but also suggests that this condition is better framed as relating to salient differences on the excepted set vs its complement, rather than strictly as given in formula (38). Finally, any account of *unless* ought to offer an explication of the differing perception of biconditionality in negative vs positive contexts – or, as Experiment 2 suggests, in downward entailing (DE) contexts vs non-DE contexts.

5 A revised theoretical account of *unless*

Consider the schemata for an account of *unless* given in (41)

(41) $Q[C]M$ *unless* $R :=$
if not $\quad Q[C \wedge \neg R]M$
negAtB $\quad \neg Q[C \wedge R]M$
perfection/not if $\quad Q[C \wedge R]\neg M$

According to the desiderata given in section 4.4, both the negAtB clause and the perfection/*not if* inference have to do with pragmatic aspects of interpretation. The negAtB clause is either itself a presupposition or the consequence of some more general felicity or appropriateness condition for *unless*, and the inference to biconditionality ought to be accounted for as an implicature. The *if not* proposition is uncontroversially entailed by *unless*.

Oustide of the three propositions/inferences in (41), this investigation has not uncovered any additional considerations or conditions that seem relevant to the meaning of *unless*. Consequently, we are left with the conclusion that the asserted content of *unless* is precisely (no more and no less than) the asserted content of *if not*. The empirical differences between the two are attributable to the joint effects of the negAtB condition and perfection/*not if*. In this section, I set out a revised account for *unless* that attempts to deal with these interactions at a more abstract and theoretically-grounded level than I have so far done.

I first describe a semantics for *unless* based on Leslie's 2008 proposal for the semantics of *if*-conditionals. This captures the equivalence between *unless* and *if not* while circumventing Higginbotham-style embedding issues. I then situate the negAtB condition as a special case of a general appropriateness condition (*conditional strengthening*) on the use of conditional statements, and argue that this interacts differently with *unless* and *if not* to produce the observed across-the-board effects. Finally, I argue that the uniqueness implicature is in fact driven by the same pressures and mechanisms as conditional perfection on regular *if*- and *if not*-conditionals, and suggest that this explains the relative strength of biconditionality on *unless* as compared to *if not*.

5.1 The semantics of *unless* and *if not*

Leslie's (2008) "modalized restrictive account" of the semantics of *if* provides the necessary tools for circumventing the embedding problem encountered by Higginbotham in equating *unless* with *if not*. I spell out only the essential pieces of this treatment here.

The first essential piece of Leslie's account is the adoption of a Kratzer-style restrictive conditional (Lewis 1975, Kratzer 1986). On this account, a "bare" conditional, e.g. *q if not p*, is presumed to contain a covert universal modal quantifier, as shown in (55). I formulate this here using *if not*, but the observations apply to *if*-conditionals as well (where the negation inside the restriction is omitted).

(55) $q \text{ if not } p := \forall w[\neg p(w)]q(w) \qquad (\leftrightarrow \textit{must } q \textit{ if not } p)$

Leslie extends this basic proposal in two ways. First, she argues that conditional operators must be treated as capable of restricting nominal quantifiers as well as modal quantifiers.[10] This is motivated by an updated version of Higginbotham's problem, which arises from simply embedding the interpretation in (55) under wide-scope nominal quantifiers:

[10]This is an extension with respect to Kratzer (1986), but may in fact maintain the spirit of Lewis (1975).

(56) a. Everyone will succeed if he does not goof off.
$= \forall x(\forall w[\neg \text{goof}(x, w)]\text{succ}(x, w))$

b. No one will succeed if he does not work hard.
$= \neg\exists x(\forall w[\neg \text{work}(x, w)]\text{succ}(x, w))$
$= \forall x \neg(\forall w[\neg \text{work}(x, w)]\text{succ}(x, w))$
$= \forall x(\exists w[\neg \text{work}(x, w)]\neg \text{succ}(x, w))$

(56a) stipulates that all individuals are such that all worlds where they work hard are worlds where they succeed. This seems reasonable, but things go wrong again under the negative quantifier. (56b) stipulates that there is no one such that all worlds where he goofs off are ones where he succeeds – or, equivalently, that everyone is such that there is at least one world where he goofs off that is not a success world. This is compatible with the truth conditions we want for (56b), but is evidently not strong enough. Leslie points out that allowing the conditional to restrict the domain of the nominal quantifier directly avoids this problem.

(57) a. Everyone will succeed if he does not goof off.
$= \forall x[\neg \text{goof}(x)]\text{succ}(x)$

b. No one will succeed if he does not work hard.
$= \neg\exists x[\neg \text{work}(x)]\text{succ}(x)$
$= \forall x[\neg \text{work}(x)]\neg \text{succ}(x)$

This improvement comes at a cost. In particular, we lose the sense that conditional statements like those in (57) can refer to a space of possible situations, rather than just some one particular instance. To illustrate the need for quantification over situations, Leslie provides the example of a lucky student named Meadow, who is exceptionally bright and will succeed regardless of the amount of work she puts in. Meadow's inclusion in the domain of quantification ought to be sufficient to falsify (58b). That is, (58b) ought to be false even in a situation where Meadow chooses to work hard. Leslie argues that this is precisely because, in any relevant situation, Meadow would also succeed without working hard.

This motivates Leslie's second modification of the basic restrictive conditional: she postulates that any conditional lacking an overt modal quantifier contains a covert wide-scope modal. (58) shows how this improves matters.

(58) a. Everyone will succeed if he does not goof off.
$= \forall w(\forall x[\neg \text{goof}(x, w)]\text{succ}(x, w))$
All situations are such that everyone who does not goof off succeeds

b. No one will succeed if he does not work hard.
$= \forall w(\neg\exists x[\neg \text{work}(x, w)]\text{succ}(x, w))$
$= \forall w(\forall x[\neg \text{work}(x, w)]\neg \text{succ}(x, w))$
All situations are such that everyone who does not work hard does not succeed.

Leslie (2008) only deals with cases where a conditional statement contains at most one overt quantifier, which justifies the assumption that, absent an overt modal quantifier, a universal modal can be given wide scope. One of the benefits of the modalized restrictor account, however, is that it can easily be extended to account for examples which contain more than one overt quantifier and where, in some sense, we would predict the possibility of scopal ambiguities. (59) provides an example for which both scope possibilities are at least plausible (although, of course, one may be privileged over the other).

(59) a. No one usually succeeds if he doesn't work hard.

b. $\neg\exists x(\text{MOST}\ w[\neg \text{work}(x, w)]\text{succ}(x, w))$
No one is such that most situations where he works hard are ones where he succeeds.

c. $\text{MOST}\ w(\neg\exists x[\neg \text{work}(x, w)]\text{succ}(x, w))$
Most situations are such that no one who does not work hard succeeds.

Arguably, these ambiguities can be generated even when the only overt quantifier is the nominal one, as in example (60).[11]

(60) a. Most people leave a party if the host does not talk to them.

b. $\forall w(\text{MOST } x[\neg\text{talk-to}(h, x, w)]\text{leave}(x, w))$
All situations are such that most people who the host does not speak to leave.

c. $\text{MOST } x(\forall w[\neg\text{talk-to}(h, x, w)]\text{leave}(x, w))$
Most people are such that in all situations where the host does not talk to them, they leave.

Finally, this view of Leslie's modalized restrictor account permits a consistent interpretation for examples like (61), where the nominal quantifier does not appear to scope over the antecedent clause at all. In these cases, we simply interpret the conditional operator as restricting the covert modal.

(61) Most people swim outside if it is not raining.
$\forall w[\neg\text{raining}(w)](\text{MOST } x \text{ swim-outside}(x, w))$
All worlds in which it is not raining are such that most people swim outside.

This account has been explicated using *if not*; as suggested above, I adopt the identical formulation for the asserted content of *unless*. That is, *unless*, like *if* and *if not*, is a restrictive operator on quantifier domains. It imposes the same restriction as *if not*: the negation of its complement. In a bare *unless*-conditional, this restriction is imposed on the domain of a covert universal modal. When the conditional includes only an overt nominal quantifier, the restriction is imposed on the domain of this quantifier, and a covert universal modal takes wide scope. Finally, if more than one quantifier which can be interpreted as scoping over both antecedent and consequent clauses is present, *unless* restricts the quantifier with narrower scope, and scopes under the other. Leslie's modalized restrictor (as set out here), then, provides a working semantics for both *unless* and *if not*: crucially, it makes the same contribution in any quantificational context, and thus avoids the embedding/compositionality problems first noted by Higginbotham (1986).

5.2 Conditional strengthening, presupposition and across-the-board contexts

Taken alone, of course, the semantics offered above fail to provide an explanation for the empirical differences between *unless* and *if not*. These differences are due to pragmatic factors. I consider first the origin of the across-the-board condition.

Conditional statements in general invoke reasoning about the circumstances in which they are appropriately asserted. Much of this reasoning can be encapsulated by the inference below, which I call *conditional strengthening*, after von Fintel. Von Fintel (2001) discusses an implicature on *if*-conditionals which has similar force to the one described in (62), although his theory of its source and derivation differ in significant ways from the discussion here.

(62) **Conditional strengthening:** Given a conditional operator COND and two propositions p and q, the statement q COND p carries with it the suggestion that the speaker is unwilling or unable to commit to the unconditional proposition q.

Conditional strengthening reflects the intuition that it is pragmatically peculiar to use a conditional when you are able to simply assert its consequent. We typically infer that someone using a conditional statement instead of the simpler alternative is either unable or for some other reason unwilling to do so. This might be due, for instance, to explicit knowledge contradicting the unconditional assertion q, or it might, alternatively, be due to a lack of knowledge about the across-the-board status of q. This is illustrated by the pragmatic weirdness of (63): since the speaker is apparently both willing and able to commit to the consequent proposition, the use of a conditional seems unjustified and highly marked.

(63) ?John will leave if Bill does not call him. Actually, he's going to leave no matter what.

[11] For further discussion of scope ambiguities in quantified conditionals, see Geurts (2004).

The same markedness (perhaps even more strongly) applies to *unless*-conditionals:

(64) ??John will leave unless Bill calls him. Actually, he's going to leave no matter what.

Von Fintel (2001) treats conditional strengthening as an implicature directly to *not unconditionally q*, but this bears some examination. Observe, in particular, that while example (63) is marked due to the assertion of *unconditionally q*, we can recover a justification for it by constructing a context in which the listener had previously expressed the belief that John's leaving was tied to Bill's calling. What this shows is that the crucial force of conditional strengthening is the notion that the speaker had a communicative/strategic/epistemic *reason* for invoking the more complex conditional statement: the inference to the existence of such a reason appears to be non-defeasible.

As it affects *if*-conditionals, then, conditional strengthening appears to be both non-optional and non-defeasible. Based on these observations, it belongs to the class of non-optional, non-defeasible "Need a Reason" (NaR) implicatures proposed by Lauer (2013). NaR implicatures are in all other respects regular (Gricean) pragmatic implicatures. Lauer's motivating case provides the template for the "justification" logic explicated here, and involves the **ignorance implicature** associated with disjunctive assertions such as (65).

(65) John is in Paris or he is in London.
$\rightsquigarrow$ The speaker is unable/unwilling to say which.

Paraphrasing from Lauer (2013), a general communicative preference for less complex utterances (i.e. either of the disjuncts) can only be overridden if there is a communicative need for the speaker to do so. In the case of (65), and with the conditionals above, the reason cannot be informativity, since the shorter alternative is actually the more informative. The reason must therefore be something else; but the need for a reason is, crucially, what cannot be canceled.

This reasoning carries over to *if*-conditionals: we can "rescue" example (63) to a certain degree by coming up with a context that supplies the speaker with a reason for choosing to use the conditional statement. However, it seems almost impossible to rescue the parallel *unless* example in (64): the *unless* case is worse than the *if* case. In (64), the speaker seems to have somehow contradicted information that she previously conveyed. In particular, the use of *unless* seems to indicate that she believes herself to be in a context which supplies a reason for using the conditional: she then contradicts this by asserting the unconditional alternative. To put it another way: where conditional strengthening on *if*-conditionals seems to be about *post hoc* rationalization, the conclusion it mandates appears to take on the role of an admittance condition or presupposition for *unless*. That is, the speaker must have a contextually-recoverable reason for using a conditional form in order for *unless* to be felicitous; it is not enough to simply infer the existence of an unverifiable reason as a consequence of the utterance.

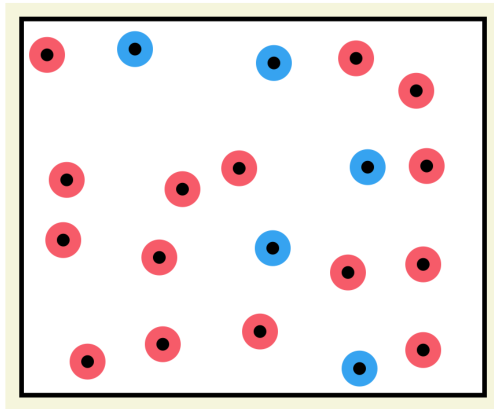
Conditional strengthening as a presupposition on *unless* is the source of the experimentally-demonstrated AtB condition discussed in section 4. Consider the AtB context for *every* shown in Figure 9. Empirically, this context yielded an agreement rate of 66.7% for the *if not* conditional (66a) but of only 14.7% for the *unless* conditional (66b). This was attributed to the significance of the negAtB clause for *unless*, which here would require that not every blue marble has a dot.

(66) a. Every marble has a dot if it is not blue.
b. Every marble has a dot unless it is blue.

Conditional strengthening as described above also predicts this outcome. Under experimental conditions, the only context provided for the assessment of the statements in (66) is the display in Figure 9 itself. Since this display in fact validates the simpler alternative "Every marble has a dot," and no additional conversational or metalinguistic context is provided, conditional strengthening marks both examples in (66) as odd, but – crucially – in different ways. Since there is no obvious reason to avoid the simpler statement, (66a) is a marked or unusual way to describe the situation.[12] The use of *unless*, on the other hand, carries

[12]It may be recovered by interpreting it as a pedantic description, in the context of an online experiment, but the point being made here is that this sort of recovery effort does not succeed with *unless*.

Figure 9: Positive across-the-board context



with it the not-at-issue information that there *is* a clear contextual reason for using a conditional statement: the absence of such a reason makes (66b) unacceptable. The theoretical account from strengthening, then, predicts majority rejection of (66b) in this context. Crucially, it predicts rejection for the reason given in section 4.1.4: the display falsifies the negAtB clause *not every blue marble has a dot.* The negAtB clause is the particular manifestation that the conditional strengthening presupposition takes in the context of Experiment 1.

The account from strengthening also explains the Experiment 2 results suggesting that the precise formulation of the negAtB clause was less significant than the existence of a salient (proportional) difference between the number of dots on target vs those on non-target marbles. In less circumscribed situations, conditional strengthening predicts that other (perhaps less direct) contextual or even metalinguistic reasons could provide the relevant justification for use of *unless.* I do not investigate this possibility here, but it is an important line of investigation for future work.[13]

To restate the proposal outlined in this section, treating conditional strengthening as an implicature for *if not* but an admittance condition or presupposition for *unless* draws the following distinction between the two. Where *if*/*if not* lead us to the inference that the speaker has a reason for avoiding the unconditional assertion of q, *unless* communicates directly the speaker's belief that she is in a context which (observably) supplies this reason. Ultimately, this presupposition captures the original idea behind the exceptive account(cf. Dancygier 1985, etc.): *unless*-statements, unlike their *if not* counterparts, draw attention to the relationship between their main generalization (the simpler alternative) and the excepted set.

5.3 Conditional perfection, implicatures, and scalar reasoning

One more piece remains: an explanation of the apparent biconditionality of *unless*-statements. I have argued that both *if* and *unless* conditionals are subject to biconditionality implicatures, and propose that in fact these implicatures have the same source (in terms of motivation and mechanism) for both types of conditional. I will refer to this implicature as *perfection* (cf. Geis and Zwicky 1971). In addition, I suggest here that the interaction of presuppositional conditional strengthening with conditional perfection (a) results in a stronger inference to biconditionality with *unless* than *if not* and (b) may also explain the positive/negative (DE/non-DE) contextual difference in this regard.

Since *unless* and *if not* have equivalent assertive content, there is no *a priori* reason to suppose that the biconditionality implicatures affecting them are derived by different means. As it affects *if*-conditionals, conditional perfection has generally been argued to accompany specific illocutionary uses: examples include threats, promises, and warnings (see Fillenbaum 1986, van Canegem-Ardijns and van Belle 2008, among

[13]The notion of a salient difference between the excepted set and its complement also features in Garcia-Alvarez (2008)'s work on other exceptive constructions.

others). These types of speech acts are cases where a strong contextual motive is given for the hearer to be interested in the truth of the consequent proposition on the target set as well as its complement.

(67) If you don't give me your money, I'll kill you!

In (67), for instance, the hearer has obvious motivation for considering whether or not he will be killed if he *does* hand over his money. Crucially, the threat only works from the speaker's perspective if the hearer believes that he can avoid certain death by paying up. Generally speaking, the inference to conditional perfection is drawn in cases where both speaker and hearer recognize the hearer's interest in the truth-value of the consequent proposition at all points in the domain (this is argued in greater detail in Nadathur 2013).

By means of the conditional strengthening presupposition, *unless* conditionals *invariably* draw attention to the truth value of the consequent on the excepted set, as well as on its complement. The conditions governing the availability of the perfecting implicature are necessarily met in any situation where *unless* is felicitous, and it is therefore to be expected that *unless* is perfected with a higher frequency than *if not* (where the implicature availability is more contextually limited). Since *unless* and *if not* share asserted content, speakers in a sense have a choice between the two items. Higher level reasoning, then, suggests that a speaker who selects the option more amenable to perfection is more likely to be gesturing towards the unasserted content of this implicature, and *unless* consequently invites perfection more strongly than *if not*.

The joint effects of conditional strengthening and conditional perfection thus conspire to produce a stronger sense of biconditionality on *unless*-statements than on *if not* statements. This explains the empirical results in the *every* condition of Experiment 1. What remains to be explained is the pattern of results described as Puzzle (C) in section 4: why this "stronger" biconditionality is not in effect in downward-entailing contexts. In what follows, I suggest one possible means of accounting for this difference.

Often, the most immediate inference from a conditional is that the speaker's reason for not asserting the simpler alternative (q) is that she knows q not to be the case across the board. This is the practical upshot of von Fintel (2001)'s discussion of conditional strengthening. From his perspective, conditional strengthening is essentially a scalar implicature; the inference that *not unconditionally* q arises from the negation of the stronger element on the Horn scale $\langle q \text{ COND } p, q \textit{ no matter what}\rangle$ (see also Horn 2000). While I have argued for a broader view of conditional strengthening, it seems extremely plausible that this type of scalar reasoning is invoked in cases where hearers are given neither (a) an obvious alternative (non-scale-based) reason for the speaker's choice of a conditional (and *unless* in particular) nor (b) a reason to suspect that *not unconditionally* q fails in the context of utterance. I suggest in particular that, despite the presuppositional nature of conditional strengthening on *unless*, hearers use scalar logic to arrive at a more specific (negAtB) conclusion. This notion is gains plausibility from the fact that, in order for conditional strengthening to point to conditional perfection, it must at least give rise to the inference that q fails on the excepted set – that is, it must give rise to a negAtB inference (which is a logical consequence of perfection).

If this conjecture is correct, then we may reasonably expect the negAtB realization of conditional strengthening to exhibit some behaviors in common with scalar inferences, and this provides a route to explicating Puzzle (C). Scalar implicatures have been argued to either weaken or reverse in downward-entailing contexts. For instance, while *some* in (68a) suggests the inference to *not all*, it is difficult to generate the parallel inference in (68b).

(68) a. Some of the students failed the exam.
$\rightsquigarrow$ *Not all of the students failed the exam.*

b. It's not the case that some of the students failed the exam.
$\not\rightsquigarrow$ *All of the students failed the exam.*

This result is discussed as implicature weakening by Horn (1989) and has been investigated by a number of others, including Chierchia (2004) (who suggests that implicatures are "recalibrated" in DE contexts), Sauerland (2004), Russell (2006), and Geurts (2009). It has also been experimentally demonstrated by Schwarz et al. (2008) for implicatures over disjunctions, and by Panizza et al. (2009) for implicatures over numerals. Levinson (2000) and others argue for a stronger view, which is that downward-entailing contexts in fact *reverse* the direction of scalar implicatures: thus the reason that *some* in (68b) fails to give rise to a scalar implicature is that, in the context of negation, *some* is in fact the stronger scale element. It therefore lacks the potential to give rise to a scalar implicature. Note, in support of this view, that the use of *all* in the same context does in fact produce an implicature to *some*.

(69) It's not the case that all of the students failed the exam.
$\rightsquigarrow$ *Some of the students failed the exam.*

Similarly, the inference to the negAtB realization of conditional strengthening may be weakened due to scalar reversal in DE contexts. In particular, in a DE context, the conditional assertion in fact becomes the stronger element of the scale $\langle q$ COND p, q no matter what$\rangle$, and thus does not necessarily give rise to the negAtB inference (*not unconditionally q*). This inference can, of course, still be drawn based on contextual and empirical evidence supporting *not unconditionally q* (and indeed, often is), which accounts for the fact that biconditionality effects do apparently arise to some degree in the negative condition. Crucially, however, the size of the biconditionality effect seems roughly the same as with *if not* in negative contexts, which is precisely the result we would expect if the boost that conditional strengthening gives to conditional perfection does not (typically) apply in these contexts.

This conjecture is, of course, speculative, and is subject to additional scrutiny and further investigation. A number of points tend to favour it, however. First, the proposal in this section permits unification of the biconditionality effects on both *if* and *unless* conditionals, while supporting the arguments given earlier that *unless* is not semantically biconditional. Secondly, since presuppositional conditional strengthening draws attention to the truth-value of the main generalization of an *unless*-statement over its excepted set, we have a natural explanation for existence of a perfecting inference on *unless*: conditional strengthening "invites" perfection by inducing the same sort of salience conditions for the implicature that are provided by the illocutionary force of threats and promises for *if*-conditionals. In addition, the presence of a scalar aspect to presuppositional conditional strengthening supports the intuitive conclusion that this precondition is related to the NaR implicature that arises on *if* conditionals, and appeals to independently-motivated results in scalar implicatures. Finally, this enables capture of many of the original intuitions about the differences between *unless* and *if not* without pushing these differences into asserted content.

6 Conclusions and future directions

I have examined in some detail a set of challenges to the exceptive solution to the compositionality puzzle apparently posed by *unless* (von Fintel 1992, Leslie 2008). I first discussed some reasons to doubt the claim that *unless* is semantically biconditional (in any context), as well as to consider the significance of an across-the-board condition. Subsequently, I presented experimental evidence arguing both against biconditionality as well as other predictions made by the exceptive account. The experimental data presented in this paper necessitates a reframing of the "compositionality" puzzle surrounding *unless* as a pragmatic puzzle, rather than a semantic one, and suggests a strong similarity between *unless* and *if*-conditionals (in particular, *if not*). I have, therefore, motivated the need for an account of *unless* that captures this similarity, is compatible with the empirically-demonstrated points of categorical divergence between *unless* and *if not*, and also incorporates a (pragmatic) explanation of the non-categorical divergences.

In moving towards such an account, I first motivated replacing the biconditionality clause of the exceptive account with a negative "across-the-board" clause stating that the main generalization of an *unless*-statement does not hold over the excepted set. I showed that this makes the correct predictions for *unless*-statements under universal quantifiers, and subsequently put forward some preliminary arguments for treating the negAtB condition as preconditional or presuppositional. In addition, I reported on the results from a second experiment using non-universal quantifiers. These results are first of all compatible with a presuppositional account and secondly suggest that the negAtB clause is a reflex of a broader condition enforcing a "salient" difference between the expected set of an *unless*-statement and its complement in the quantifier domain.

Taken together, these points suggest a number of revisions to the theoretical account of *unless*. The differences between *unless* and *if not* appear to be located in specific aspects of pragmatics: the NaR implicature (Lauer 2013) to conditional strengthening is realized as a precondition for *unless* (which produces the negAtB clause in the experimental context described here). This in turn creates the necessary contextual conditions for inviting the implicature to conditional perfection (cf. Geis and Zwicky 1971) and thus accounts for (a) the prominence of the biconditional interpretation, (b) its relative strength on *unless* as opposed to *if not*, and (c) the weakening of biconditionality in downward-entailing contexts (as a consequence of the involvement of scalar reasoning).

There is, of course, more work to be done either to properly establish or to refine these claims. First, the suggestion that the original compositionality problem is the reflection of reversed scalar reasoning in DE contexts needs to be more thoroughly investigated: it stands and falls both with the claim that biconditionality on *unless* is the same phenomenon as conditional perfection, and secondly, that scalar reasoning to the negAtB condition comes into play on the path from presuppositional conditional strengthening to full conditional perfection. An interesting expected consequence of the proposal in section 5.3, and one which certainly merits investigation, would be reduced biconditionality effects in situations supplying an obvious reason for the use of *unless* which is, crucially, not the negAtB condition. Secondly, results from the follow-up study on non-universal quantifiers remain preliminary. More detailed analysis of the results of this study, as well as further investigation which, for instance, more thoroughly controls both for the effects of variance in quantifier interpretation as well as the effects of secondary implicatures (e.g. *some* $\rightsquigarrow$ *not all*) would be helpful in establishing either more conclusive support for the proposal set out here, or the need to consider additional/alternative factors.

The proposals made here bear on a number of broader questions about the practical differences between presuppositions and implicatures. In arguing for a presuppositional account of conditional strengthening on *unless*, I have suggested that it simply establishes that the speaker believes herself to be in a context which satisfies the inference, whereas *if not* indicates a speaker's desire to convey the relevant information about the context. I have not, however, investigated in any detail what this difference means as part of a broader set of rules governing conversational strategy and/or cooperative behavior. Relatedly, it seems likely that the presuppositional force of conditional strengthening on *unless* can be better explicated by considering what sorts of communicative needs might drive this sort of attachment. As one illustration of the issues here, consider the following data point:

(70) I won't put your head under the water unless you ask me to.[14]

Offered in a context where the speaker is teaching his daughter how to swim, and she has expressed concern over having her head fully immersed, the force of *unless* as opposed to *if not* here seems to be to offer her some reassurance as to her control in the matter of immersion: it is in fact irrelevant whether a request to be put under the water would be met with compliance or not. This suggests that the presuppositional nature of conditional strengthening might be linked, for instance, to a condition requiring some sort of epistemic relationship between the p and q propositions in the statement q *unless* p. An exploration of the data on *unless* which makes use of the diagnostic toolkit for projective content developed in Tonhauser et al. (2013) is an obvious direction to take here, and seems extremely likely to shed light on these questions.

There is, evidently, plenty of room for future work on *unless* and its relationship to other conditional constructions alone. In addition to this, however, work on other exceptive constructions (e.g. Garcia-Alvarez 2008) suggests a role for salience conditions of the sort discussed here. The literature on exceptive constructions, deals, more generally, with a number of presupposition-like conditions (see Hoeksema 1990, Moltmann 1995). An investigation into these conditions, and in particular, how (and whether) they relate to the pragmatic considerations discussed here would provide an excellent starting point for investigating the broader questions about pragmatic theory that this paper has only touched on.

[14] This example was given to me by Itamar Francez (p.c., University of Chicago, 2013).

Acknowledgements

My thoughts and ideas about the issues in this paper have been shaped and refined by quite a number of people. I wish to thank my qualifying paper committee at Stanford for much helpful discussion: Cleo Condoravdi, Chris Potts, and in particular Dan Lassiter, with whose guidance and assistance the experiments reported here were conducted. I have also benefitted greatly from feedback from audiences at the ESSLLI 2014 Student Session in Tübingen and Sinn und Bedeutung 19 in Göttingen. Lastly, many thanks are due to Ash Asudeh, who told me in our first meeting at Oxford that the only part of my eleven-page essay on compositionality that he wanted to talk about was the short footnote objecting to the *unless* "counterexample." That conversation is by no means the only contribution I have to thank him for, but it was what got this project started in the first place.

References

Abrusán, Márta and Szendröi, Kriszta. 2013. Experimenting with the King of France: Topics, verifiability, and definite descriptions. *Semantics and Pragmatics* 6, 1–43.

van der Auwera, Johan. 1997. Conditional perfection. In Angeliki Athanasiadou and René Dirven (eds.), *On conditionals again*, pages 169–190, Philadelphia: John Benjamins Publishing Company.

Bott, Lewis and Noveck, Ira. 2004. Some utterances are underinformative: the onset and time course of scalar inferences. *Journal of Memory and Language* 51, 437–457.

Breheny, Richard, Katsos, Napoleon and Williams, John. 2006. Are generalised scalar implicatures generated by default? An on-line investigation into the role of context in generating pragmatic inferences. *Cognition* 100, 434–464.

van Canegem-Ardijns, Ingrid and van Belle, William. 2008. Conditionals and types of conditional perfection. *Journal of Pragmatics* 40, 349–376.

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

Chierchia, Gennaro and McConnell-Ginet, Sally. 1990. *Meaning and Grammar*. Cambridge, MA: MIT Press.

Clark, Herbert and Clark, Eve. 1977. *Psychology and language: an introduction to psycholinguistics*. London: Thomson Learning.

Comrie, Bernard. 1986. Conditionals: a typology. In Elizabeth Traugott, Alice ter Meulen, Judy Reilly and Charles Ferguson (eds.), *On conditionals*, pages 77–99, Philadelphia: John Benjamins Publishing Company.

Dancygier, Barbara. 1985. If, unless, and their Polish equivalents. *Papers and Studies in Contrastive Linguistics* 20, 64–72.

Degen, Judith and Tanenhaus, Michael. 2011. Making inferences: the case of scalar implicature processing. In L. Carlson, C. Hölscher and T. Shipley (eds.), *Proceedings of the 33rd Annual Conference of the Cognitive Science Society*, Austin, TX.

Égré, Paul and Cozic, Mikael. 2012. Conditionals. In Maria Aloni and Paul Dekker (eds.), *Handbook of semantics*, Cambridge: Cambridge University Press.

Fillenbaum, Samuel. 1986. The use of conditionals in inducements and deterrents. In Elizabeth Traugott, Alice ter Meulen, Judy Reilly and Charles Ferguson (eds.), *On Conditionals*, pages 179–195, Cambridge: Cambridge University Press.

von Fintel, Kai. 1992. Exceptive conditionals: the meaning of "unless". In *Proceedings of the North East Linguistics Society*, volume 22, pages 135–151.

von Fintel, Kai. 1994. *Restrictions on quantifier domains*. Ph. D.thesis, University of Massachusetts–Amherst.

von Fintel, Kai. 2001. Conditional strengthening: a case study in implicature, ms, MIT.

von Fintel, Kai and Iatridou, Sabine. 2002. If and when if-clauses can restrict quantifiers, ms, Workshop in Philosophy and Linguistics at the University of Michigan.

Garcia-Alvarez, Ivan. 2008. *Generality and exception: a study in the semantics of exceptives*. Ph. D.thesis, Stanford University.

Geis, Michael. 1973. "If" and "unless". In Braj B. Kachru, Robert B. Lees, Yakov Malkiel, Angelina Pietrangeli and Sol Saporta (eds.), *Issues in linguistics: papers in honor of Henry and Renee Kahane*, pages 231–253, Urbana, IL: University of Illinois Press.

Geis, Michael and Zwicky, Arnold. 1971. On invited inferences. *Linguistic Inquiry* 2, 561–566.

Geurts, Bart. 2004. On an ambiguity in quantified conditionals, ms, University of Nijmegen.

Geurts, Bart. 2009. Scalar implicatures and local pragmatics. *Mind and Language* 24, 51–79.

Grice, Paul. 1975. *Studies in the way of words*. Cambridge MA: Harvard University Press.

Higginbotham, James. 1986. Linguistic theory and Davidson's program in semantics. In Ernie Lepore (ed.), *Truth and interpretation: perspectives on the philosophy of Donald Davidson*, pages 29–48, Oxford: Blackwell.

Hoeksema, Jack. 1990. Exploring exception phrases. In Martin Stokhof and Leen Torenvliet (eds.), *Proceedings of the Seventh Amsterdam Colloquium*, pages 165–190.

Horn, Laurence. 1989. *A Natural History of Negation*. Chicago: University of Chicago Press.

Horn, Laurence. 2000. From if to iff: conditional perfection as pragmatic strengthening. *Journal of Pragmatics* pages 289–326.

Janssen, Theo. 1997. Compositionality. In Johan van Benthem and Alice ter Meulen (eds.), *Handbook of logic and linguistics*, pages 417–475, Amsterdam: Elsevier.

Kratzer, Angelika. 1986. Conditionals. In *Proceedings of Chicago Linguistic Society*, volume 22, pages 1–15.

Lasersohn, Peter. 1993. Existence presuppositions and background knowledge. *Journal of Semantics* 10, 113–122.

Lauer, Sven. 2013. *Towards a dynamic pragmatics*. Ph. D.thesis, Stanford University.

Leslie, Sarah-Jane. 2008. "If," "unless," and quantification. In Robert Stainton and Christopher Viger (eds.), *Compositionality, context, and semantic value: essays in honor of Ernie Lepore*, pages 3–30, Amsterdam: Springer.

Levinson, Stephen. 2000. *Presumptive meanings: the theory of generalized conversational implicature*. Cambridge MA: MIT Press.

Lewis, David. 1975. Adverbs of quantification. In Paul Portner and Barbara Partee (eds.), *Formal semantics: the essential readings*, pages 178–188, Oxford: Blackwell.

Moltmann, Friederieke. 1995. Exception sentences and polyadic quantification. *Linguistics and Philosophy* 18, 223–280.

Nadathur, Prerna. 2013. *If...(and only if): conditional perfection and completeness*. Masters Thesis, University of Oxford.

Nadathur, Prerna and Lassiter, Daniel. 2014. Unless: an experimental approach. In *Proceedings of Sinn und Bedeutung*, to appear.

Panizza, Daniele, Chiercia, Gennaro and Clifton Jr., Charles. 2009. On the role of entailment patterns and scalar implicatures in the processing of numerals. *Journal of Memory and Language* 61, 503–518.

Papafragou, Anna and Musolino, Julien. 2003. Scalar implicatures: experiments at the semantics-pragmatics interface. *Cognition* 86, 253–282.

Potts, Christopher. 2005. *The logic of conventional implicatures*. Oxford: Oxford University Press.

Quine, Willard. 1959. *Methods of logic*. New York: Holt, Rinehart & Winston.

Reichenbach, Hans. 1947. *Elements of symbolic logic*. New York: The Free Press.

Russell, Benjamin. 2006. Against grammatical computation of scalar implicatures. *Journal of Semantics* 23, 361–382.

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

Schwarz, Florian. 2014. False but slow: Evaluating statements with non-referring definites. *Journal of Semantics* to appear.

Schwarz, Florian, Clifton, Chuck and Frazier, Lyn. 2008. Strengthening 'or': Effects of focus and downward entailing contexts on scalar implicatures, ms, UMass Amherst.

Stalnaker, Robert. 2002. Common ground. *Linguistics and Philosophy* 25, 701–721.

Szabó, Zoltan. 2008. Compositionality. In *The Stanford Encyclopedia of Philosophy*, winter 2008 edition.

Tonhauser, Judith, Beaver, David, Roberts, Craige and Simons, Mandy. 2013. Toward a taxonomy of projective content. *Language* 89, 66–109.

Zuber, Richard. 1999. Exclusion phrases and criticisms of semantic compositionality. In Akira Ikeya and M. Kawamory (eds.), *Proceedings of the 14th Pacific Asia Conference on Language, Information, and Computation*, pages 401–412.