[last modified: November 29, 2018]

Logical integrity: from *Maximize Presupposition!* to *Mismatching Implicatures*

Contents

1 Introduction

In this study a novel generalization, labeled Logical Integrity (hf. LI), is put forth which yields a unified account of some fairly broad class of acceptability judgments. The generalization is built with the following statement at its core, stated *pro tem* at the speech-act level.

(1) **LI's core condition.** A sentence ϕ must not be uttered in context C if it has an alternative ψ such that (i) ϕ *contextually* entails ψ in C, but (ii) ϕ does not *logically* entail ψ.

It will be argued that LI, which consists of (1) coupled with a suitable "projection recipe",[1] makes adequate predictions for a broad array of examples which has so far been chopped up by three distinct analyses that happen to capture more or less mutually incompatible generalizations: the *Maximize Presupposition!* principle originating in Heim (1991), the *Presupposed Ignorance* system in Spector & Sudo (2017) and the *Mismatching Implicature* approach of Magri (2009a). It is furthermore argued on the basis of novel evidence that the predictions made by LI are superior to each of these three piece-meal analyses individually considered (in some cases it is shown that salient modifications of the analyses would not solve the relevant problems either). As both the empirical landscape and inter-connections of the proposed analyses in the literature are somewhat complicated, in this introductory section an overview of the relevant facts is provided followed by the outline of the rest of the paper.

Heim (1991) sketched a principle of language use according to which, given the choice between two competing forms, *all else equal* the one with the stronger presupposition must be used unless

1 See the discussion in sections 2.2 and 5.

its presupposition is not known to be true. This principle was taken up in subsequent literature (Percus 2006, Sauerland 2008, Schlenker 2012, a.o.) and is standardly referred to as *Maximize Presupposition!* (hf. MP). As an example, in (2a) 'all' is blocked by 'both' because the latter triggers a stronger presupposition (that John has exactly two arms) which is satisfied in the context. In contrast in (2b) 'all' is available because this time the presupposition of 'both' (that John has exactly two fingers) contradicts background assumptions rendering the 'both'-sentence unusable.

(2) [Context: John has ten fingers and two arms.]

a. John broke {#all,✓both} of his arms.

b. John broke {✓all,#both} of his fingers.

An alternative account of the oddness of 'all' in (2a) is given by (1) above. Note that the 'all'-sentence in (2a) does not logically entail the 'both'-sentence: since it is not a logical truth that John has two arms, it is logically possible that John has more than two hands and broke all of them in which case the 'all'-sentence is true but the 'both'-sentence is not (more specifically, the 'both'-sentence is undefined due to presupposition failure). On the other hand, on the contextual assumption that John has exactly two hands the truth of the 'all'-sentence guarantees the truth of the 'both'-sentence, i.e., the argument in (3) is valid. Put differently, the 'all'-sentence contextually entails the 'both'-sentence in any context in which it is taken for granted that John has exactly two hands.

(3)
$$\begin{array}{rl} & \text{John has exactly two arms} \\ & \text{John broke all of his arms} \\ \hline \therefore & \text{John broke both of his arms} \end{array}$$

Hence on the assumption that 'all' and 'both' compete, an assumption that LI shares with MP, it is predicted by (1) that the 'all'-sentence in (2a) should be blocked by its 'both'-alternative in any context in which it is assumed that John has two hands. The "core" examples of MP are thus explained by LI as well as MP. However there is a delicate difference in predictions: MP, unlike LI, relies crucially on the condition that the presuppositions of the alternatives be common ground. As argued in section 3.4.1, MP-type effects can arise even when the presupposition of the stronger alternative is *not* known to be true, a fact that LI accounts for without further ado but is problematic for MP.

As quickly pointed out above, MP as a principle of language use is assumed to kick in to decide between a set of alternatives only when *all else is equal* between the them. What does this restiction amount to in practice? Several arguments in the literature (see in particular Percus 2006 and Schlenker 2012) point to the conclusion that in the context of MP all else is equal when (and only when) the relevant competitors are "equally informative", roughly in the sense that neither competitor can be true without the other being true as well. Hence if two sentences are not equally informative, MP cannot be called upon to decide between them. Recently, Spector & Sudo (2017) (hf. S&S) have problematized this conclusion. The crucial example discussed by S&S is (4).

(4) [Context: all students smoke.]
#John is unaware that some students smoke.
ALT = {John is unaware that all students smoke}

Assumption: 'α is unaware that ϕ' presupposes that ϕ is true and asserts that it is not the case that α believes ϕ to be true.

The unacceptability of the 'some'-sentence in (4) is reminiscent of the MP-effect in (2a) because here as well the unacceptable sentence has an alternative, the 'all'-sentence, with a stronger presupposition (that all students smoke) which is satisfied in the given context. Nevertheless MP cannot account for the oddness of the 'some'-sentence in (4) *if it is restricted by equal-informativeness*. The reason is that (in contrast to the alternatives in (2a)) the alternatives in (4) are *not* equally informative. In a situation in which all students smoke and John is sure that some students smoke but is uncertain as to whether all students do the 'some'-sentence in (4) is false but its 'all'-alternative is true. Furthermore, this particular situation is not contextually ruled out in (4) (since nothing is assumed regarding John's epistemic state) therefore the two competitors are not equally informative in that particular context. S&S take (4) at face value and propose that the equal-informativeness condition of MP must be dropped. They call the resulting, more general principle *Presupposed Ignorance*. Since PI is not controlled by equal-informativeness it can account for (4): the 'some'-sentence has an alternative with a stronger presupposition which is satisfied in the context and that is enough for PI to rule it out in that context. But by the same logic PI generates too many "false negatives", incorrectly ruling out certain sentences as infelicitous. S&S, therefore, introduce a novel implementation of the exhaustivity operator which can be inserted to "rescue" certain sentences from oddness in certain contexts.

As it happens, (1) can account for the oddness of the 'some'-sentence in (4) without any modification being necessary. Note that the 'some'-sentence does not logically entail its 'all'-alternative: in a situation in which only some students smoke, the 'some'-sentence may be true (depending on John's epistemic state) but the 'all'-sentence is certainly not, *because its presupposition is false*. On the other hand on the contextual assumption that all students smoke, the truth of the 'some'-sentence immediately guarantees the truth of its 'all'-alternative. That is to say, the following argument is valid.

(5)
All students smoke
John is unaware that some students smoke
$\therefore$ John is unaware that all students smoke

LI, therefore, predicts that any context in which it is assumed that all students smoke is a context in which the 'some'-sentence is infelicitous, which is the desired prediction. However, there is a difference in predictions between PI and LI: PI (like MP) dictates a preferences for alternatives with *stronger* presuppositions while no such preference is encoded by LI. As argued in section 3.4.2, alternatives with logically *independent* presuppositions also enter into a kind of competition that is very similar to (4), a fact that can be made to follow from LI but is problematic for PI (and MP).

Finally, a highly consequantial mechanism of exhaustification is motivated in Magri (2009a,b, 2011) on the basis of examples including (6). The generalization that Magri's system (call it *Mismatching Implicatures*, hf. MI) captures, for the simple cases, is that a sentence is odd in a context in which it is equally informative with one of its logically non-weaker alternatives. Thus in the context of (6) the 'some'-sentence and its 'all'-alternative are equally informative, since the possibility of John giving only some of his students an A is contextually ruled out, and as predicted by this generalization 'some' cannot be felicitously used. Note that since there is no difference between the presuppositions of the two alternatives in (6), MP/PI are inapplicable in this case as a matter of principle.

(6) [Context: John gave the same grade to his students.]
#John gave some of his students an A.
ALT $= \{$John gave all of his students an A$\}$

In section 4, (6) is embedded in the larger paradigm and a compact summary of Magri's proposal is provided. It is shown that the paradigm can be captured by LI without further ado, thus reinforcing the intuition that (2), (4) and (6) have the same source. Furthermore, based on some observations regarding the architecture of Magri's proposal, several obstacles for it as it stands and for possible extensions of it to MP effects (see also Singh 2009) are discussed in the same section. In section 5 a modified version of the initial version of the generalization (as put forth in section 2) is motivated on the basis of a set of problematic data points discussed in the earlier sections. Section 6 is concerned with several loose ends and open problems for the present account.

2 Logical Integrity (first version)

2.1 Background assumptions

The set of propositions that are "taken for granted" by the interlocutors at a particular point in a conversation is the common ground (or, the background assumptions). The set of those possible worlds that are compatible with the background assumptions is the context set (or, global context (see below), or, simply, context). A trivalent treatment of presuppositions is assumed throughout, where the third truth-value '#' represents presupposition failure. The presupposition of a sentence is satisfied in a context iff the sentence is defined (does not denote '#') at every "point" in that context.

Within a trivalent setting, the classical, bivalent notion of entailment can be generalized in different ways (see, e.g., Chemla et al. 2017). The discussion below relies almost exclusively on the following notion of entailment, given in the generalized form as it will be applied to both proposition- and property-denoting expressions. For simplicity the definition is given for model-theoretic objects rather than object-language expressions.

(7) If $\mathcal{X}_1$ and $\mathcal{X}_2$ are two objects of a type τ type that "ends in t" and can take n arguments, $\mathcal{X}_1$ entails $\mathcal{X}_2$, $\mathcal{X}_1 \vDash \mathcal{X}_2$, iff for all type-appropriate sequences of objects $\mathcal{Y}_1, \ldots, \mathcal{Y}_n$, if $\mathcal{X}_1(\mathcal{Y}_1)\ldots(\mathcal{Y}_n) = 1$ then $\mathcal{X}_2(\mathcal{Y}_1)\ldots(\mathcal{Y}_n) = 1$.

Another, strictly weaker notion of entailment, which we will have occasion to use in its bidirectional form in section 3, is that of Strawson-entailment (von Fintel 1999).

(8) If $\mathcal{X}_1$ and $\mathcal{X}_2$ are two objects of a type τ type that "ends in t" and can take n arguments, $\mathcal{X}_1$ Strawson-entails $\mathcal{X}_2$ iff for all type-appropriate sequences of objects $\mathcal{Y}_1, \ldots, \mathcal{Y}_n$, if $\mathcal{X}_1(\mathcal{Y}_1)\ldots(\mathcal{Y}_n) = 1$ then $\mathcal{X}_2(\mathcal{Y}_1)\ldots(\mathcal{Y}_n) \in \{\#, 1\}$.

This study relies on Schlenker's (2009) theory of local contexts. Under this approach, given a sentence ϕ and a global context C, each occurrence of a property- or proposition-denoting constituent of ϕ, α, can be mapped to a model-theoretic object of the same semantic type, it's *local* context, denoted $lc(\alpha, \phi\,\mathrm{C})$. There is no need to get into the bolts and gears of Schlenker's theory here. In (9) all the relevant facts are aggregated.

(9) For any proposition-denoting expressions ϕ and ψ, property-denoting expressions α and β, individual-denoting expression υ, and generalized quantifier $\mathcal{Q}$,

a. $lc(\psi,\ [\phi \wedge \psi]\ ,\mathrm{C}) = lc(\psi,\ [\phi \rightarrow \psi]\ ,\mathrm{C}) = \lambda w.\ w \in \mathrm{C} \wedge [\![\phi]\!]^w = 1$

b. $lc(\psi,\ [\phi \vee \psi]\ ,\mathrm{C}) = \lambda w.\ w \in \mathrm{C} \wedge [\![\phi]\!]^w = 0$

c. $lc(\beta,\ [\mathcal{Q}(\alpha,\beta)]\ ,\mathrm{C}) = \lambda w.\lambda x.\ w \in \mathrm{C} \wedge [\![\alpha]\!]^w(x) = 1$

d. $lc(\phi,\ [\upsilon \text{ is (un)aware that } \phi]\ ,\mathrm{C}) = \lambda w.\ \exists w' \in \mathrm{C}:\ w = w' \vee w \in \textsc{dox}_{\boldsymbol{\upsilon}}^{w'}$

Note that for any constituent its local context has the same semantic type. Hence the different types in (9b) and (9c). Once a working notion of local contexts is available, entailment can be relativized to contexts, local and global.

(10) If $\mathcal{X}_1$, $\mathcal{X}_2$ and $\mathcal{C}$ are three objects of a type τ type that "ends in t" (the latter being the context) and can take n arguments, $\mathcal{X}_1$ contextually entails $\mathcal{X}_2$ in $\mathcal{C}$, $\mathcal{X}_1 \vDash_{\mathcal{C}} \mathcal{X}_2$, iff for all type-appropriate sequences of objects $\mathcal{Y}_1, \ldots, \mathcal{Y}_n$, if $\mathcal{X}_1(\mathcal{Y}_1)\ldots(\mathcal{Y}_n) = 1$ and $\mathcal{C}(\mathcal{Y}_1)\ldots(\mathcal{Y}_n) = 1$ then $\mathcal{X}_2(\mathcal{Y}_1)\ldots(\mathcal{Y}_n) = 1$.

Put differently, $\mathcal{X}_1 \vDash_{\mathcal{C}} \mathcal{X}_2$ iff $\mathcal{X}_1 \wedge \mathcal{C} \vDash \mathcal{X}_2$ (assuming generalized conjunction). For any expression ϕ and context C, $[\![\phi]\!] \wedge \mathrm{C}$ can be thought of as the *contextual meaning* of ϕ in C.

Perhaps some words need to be said pertaining to (9d). According to (9d), the local context of the clausal complement of '(un)aware' relative to the (global) context C is the union of C with the set of all words that are compatible with what υ possibly believes to be true. The informal justification is this. Given ϕ, the task is to identify all and only those worlds at which one needs to know the denotation of ϕ in order to compute the denotation of the full sentence. Since '(un)aware' is factive, we need to know whether ϕ holds at the worlds in C and, since '(un)aware' is doxastic, we need to whether or not ϕ is true at every world which is compatible with υ's beliefs (at the world of evaluation).[2]

A final point regarding the way with which the theory of local contexts will be used in this study is worth stressing.[3] Even though the primary motivation behind this theory is to give an explanatory account of presupposition projection, local contexts can be used for a variety of different purposes. There is technically no reason for a theory that uses local contexts for some purpose or other to use it for presupposition projection as well. As such, it is very much possible to use local contexts in a framework in which presupposition projection is handled in some other fashion. In particular, nothing in what follows hangs on the assumption that local contexts are the engine of presupposition projection. In fact, we won't have much occasion to talk about projection in any detailed way at all; we can afford to simply rely on well-established descriptive generalizations. On the other hand, in this study local contexts are worked heavily as *information sources* relative to which certain conditions can be checked.

2 The definition given in (9d) differs from Schlenker (to appear). The difference is immaterial for my purposes here.

3 See also (Spector & Sudo 2017: fn. 32), which makes the same point.

2.2 The proposal

The proposed generalization is introduced in two steps. In this subsection the first version, LI, is formulated (explicitly in (19)) and its predictions regarding the relevant examples, a subset of which was discussed in section 1, are rigorously investigated. However, certain problematic facts will motivate a modification to the first version, leading to the final version of the proposal, LI*, spelled out in section 5.

As pointed out in section 1, the core of LI is the statement in (11), repeated from (1). It dictates that for a sentence to be acceptable, a certain balance must be hit between the contextual information it conveys and the pattern of logical entailment that it enters into with its (independently characterized) alternatives.

(11) A sentence ϕ is unacceptable in context C if it has a logically non-weaker alternative ψ which it contextually entails in C.

The statement in (11), coupled with the definitions of logical and contextual entailment in the previous section (i.e., (7) and (10) respectively), yields the following generalization.

(12) a. Let ϕ and ψ be two competing forms and C some context. If it is *logically* possible for ϕ to be true and ψ to be "untrue" (i.e., either false or undefined) then this must be *contextually* possible in C as well, otherwise ϕ is unacceptable in C. A bit more formally,

b. Let ϕ and ψ be alternatives such that $\exists w \in W : [\![\phi]\!]^w = 1 \wedge [\![\psi]\!]^w \in \{0,\#\}$. Then, ϕ is acceptable in context C only if $\exists w \in \mathrm{C} : [\![\phi]\!]^w = 1 \wedge [\![\psi]\!]^w \in \{0,\#\}$.

To unpack the reasoning compressed (11), suppose ϕ and ψ are two competing sentences. According to the definition in (7), ψ is logically non-weaker than ϕ (i.e., $\phi \not\models \psi$) iff it is possible for ϕ to be true and ψ to be untrue. Assume ψ is in fact logically non-weaker than ϕ. According to (11), ϕ cannot be used in a context C if, in C, ϕ *contextually* entails ψ. By contraposition, if ϕ is acceptable in some context C, then ϕ does *not* contextually entail ψ in C. The consequent of this conditional (i.e., $\phi \not\models_{\mathrm{C}} \psi$) effectively boils down to an existential claim, in light of the definition (10): C contains at least one world in which ϕ is true and ψ is either false or undefined. If this possibility is contextually ruled out, ϕ is unacceptable. The examples worked through below illustrate the breadth of this proposal.

To begin with, consider the classic MP effect in (13).

(13) [Context: John has exactly two students.]
#John invited all his students.
ALT = {John invited both his students}

The 'all'-sentence in (13) does not logically entail the 'both'-alternative; while the truth of 'all' is sufficient to guarantee that 'both' is not false, it is not sufficient to guarantee that it is *true*: in a world in which John has seven students and invited all of them, 'all' is true but 'both' is undefined. Therefore, according to (11), the 'all'-sentence in (13) can only be used if it is contextually possible that it is true but the 'both'-alternative is either false or undefined. As just mentioned, the truth of the 'all' guarantees that the 'both' is not false, therefore the requirement boils down to that it must be contextually possible that the 'all'-alternative is true and the 'both'-alternative is undefined; in other words, it must be contextually possible that John has more than two students and he invited all of them. But the context given in (13) entails that John has exactly two students, thus ruling this possibility out. Therefore the 'all'-sentence in (13) is predicted to be blocked. In the same context, the 'both'-sentence is predicted to be vacuously acceptable to the extent that it lacks a non-weaker alternative.

Next, consider a simple Magri case.

(14) [Context: John always gives the same grade to his students.]
$^{\#}$John gave an A to some of his students.
ALT = {John gave an A to all of his students}

Obviously the 'some'-sentence in (14) does not logically entail its 'all'-alternative. Therefore, the 'some'-sentence is predicted to come with the requirement that it must be contextually possible that it is true and the 'all'-alternative is either false or undefined. But since the two sentences carry the same presupposition (that John has students), it is impossible for 'some' to be true and 'all' be undefined. Therefore, it must be contextually possible that 'some' is true and 'all' is *false*; in other words, it must be contextually possible that John has some students and gave an A to some but not all of them. This possibility is ruled out by the context specified in (14), hence the 'some'-sentence is correctly predicted to be blocked. In the context of (14), the 'all'-sentence is predicted to be vacuously acceptable to the extent that it lacks a non-weaker alternative.[4]

Consider now the example brought forth by Spector & Sudo (2017) which, as pointed out in section 1, is problematic for *Maximize Presupposition!*. The example is repeated in (15a) and the assumed lexical entry is given in (15b).

(15) a. [Context: all students smoke.]
$^{\#}$John is unaware that some students smoke.
ALT = {John is unaware that all students smoke}

b. $[\![\text{ unaware }]\!]^w = \lambda P \lambda x : P(w) = 1. \neg \mathrm{B}_x^w[P]$

c. For any world w, individual x and proposition P, $\mathrm{B}_x^w[P]$ iff x believes P to be true in w.

The 'some'-sentence in (15a) does not logically entail its 'all'-alternative: in a world in which only some students smoke, 'some' *may* be true but 'all' is certainly undefined. Therefore, 'some' is predicted to come with the requirement that it must be contextually possible that it is true and 'all' is either false or undefined. Now, it is impossible for the 'some'-sentence to be true when its 'all'-alternative is false.[5] Therefore, the requirement boils down to that it must be contextually possible that the 'some'-sentence is true and its 'all'-alternative is undefined; in other words, it must be contextually possible that some but not all students smoke and John is unsure as to whether that any student smokes (i.e., $\neg \mathrm{B}_{\mathrm{J.}}[\exists]$). This possibility is ruled out by the context of (15a), hence the 'some'-sentence is predicted to be blocked.

Importantly, for (15a) one must also check that the 'all'-sentence is not incorrectly ruled out:

4 See section 6.2 for a discussion of an immediate consequence of this prediction pertaining to homogeneity.

5 'α is unaware that ϕ' can be analyzed in as $\underline{\phi} \wedge \neg \mathrm{B}_\alpha(\phi)$, where underlining marks the presupposition. It immediately follows that if $\phi \vDash \psi$, whenever 'α is unaware that ϕ' is true, 'α is unaware that ψ' is either true or undefined: the clausal complement of 'unaware' is a Strawson-downward-entailing environment.

(16) [Context: all students smoke.]
✓John is unaware that all students smoke.
ALT = {John is unaware that some students smoke}

The novelty here is that, in contrast to (13) and (14), the two competitors in (15a) are logically independent. Having shown earlier that the presuppositionally stronger alternative (i.e., the 'all'-sentence) blocks the presuppositionally weaker one (i.e., the 'some'-sentence), in light of the acceptability of (16) we must now make sure that the opposite does not hold. It is straightforward to check that (11) predicts the 'all'-sentence to come with the requirement that it is contextually possible that all students smoke but John only believes that some students smoke. This possibility is not ruled out in (16) (since no assumption is made regarding the epistemic state of the attitude holder, John), hence 'all' is indeed predicted to be acceptable in the specified context.[6]

Let us now look at how (11) fairs with the positive counter-part of 'unaware'.

(17) [Context: all students smoke.]
✓John is aware that some students smoke.
ALT = {John is aware that all students smoke}

As pointed out by S&S, 'aware' does not show the same behavior as 'unaware'. In a context in which it is common ground that all students smoke, 'aware…some…' is fine, (17), but '*un*aware…some…' is not, (15a). This is as things should be, according to (11): the 'some'-sentence in (17) does not logically entail its 'all'-alternative (i.e., that John is aware that all students smoke), therefore it is predicted to come with the requirement that it must be contextually possible that 'some' is true while 'all' is either false or undefined. In contrast to the previous examples, here *a genuinely disjunctive requirement is generated* because the truth of 'some' in (17) is indeed compatible with both falsity *and* undefinedness of its 'all'-alternative; thus the requirement is that it must be contextually possible that the 'some'-sentence in (17) is true and *either* not all students smoke *or* all students smoke but John does not believe so. Given the background assumptions in (17), the second possibility is *not* contextually ruled out and therefore the 'some'-sentence is predicted to be acceptable.

Importantly, every effect so far discussed (and those that will be discussed in the following sections) can be reconstructed "locally".

(18) a. #No professor who has exactly two students invited <u>all</u> of them.
b. #If all students smoke, John is unaware that <u>some</u> students smoke.
c. #Either John gave his students different grades, or he gave <u>some</u> of them an A.

Take (18a) (the same point can be made with the other two examples). The problem here is that, at root, the sentence logically entails its 'both'-alternative, indeed the two are logically equivalent. The reason is that the presupposition triggered by 'both' in the sentence 'no professor who has exactly two students invited both of them' is filtered through the restrictor and boils down to the

6 The acute reader might object that even if this possibility *is* contextually ruled out (e.g., in a context in which it is common ground that (i) all students smoke and (ii) either John thinks all students smoke or he thinks that no student smokes), the 'all'-sentence (i.e., John is unaware that all students smoke) is still acceptable. Indeed this fact cannot be accounted for by (11) but it will be one of the welcomed predictions of LI* in section 5.

presupposition that every professor who has exactly two students has exactly two students, which is tautologous. Consequently, at the sentential level there is no truth-conditional difference between the sentence in (18a) and its 'both'-alternative. For our purposes this means that (11) no longer predicts any contrast to arise, contrary to fact. This is because (11) is a *global* condition that applies to sentences at root and, as such, it is blind to the internal constitution of sentences. If two sentences are globally synonymous then (11) does not "see" the difference between the two to begin with, let alone predicting one to be blocked by the other.

However, correct predictions are made if (11) is checked against the *local* context of the scope expression 'invited all of his students'. The local context of the scope of 'no' in (18a) is predicted to be that function which maps each world w in the context-set to the set of individuals that are professors with exactly two students in w. For simplicity, we can turn this function into the set $\mathrm{S} = \{\langle w,a\rangle : w \in \mathrm{C} \wedge [\![\text{professor who has exactly two students}]\!]^{w}(a) = 1\}$. Now relative to this context, the weaker alternative '$\lambda x.x$ invited all of x's students' contextually entails the logically stronger '$\lambda x.x$ invited both of x's students'; indeed, the two are contextually equivalent in the sense that for any $\langle w,a\rangle \in \mathrm{S}$, $[\![\lambda x.x \text{ invited all of } x\text{'s students}]\!]^{w}(a) = [\![\lambda x.x \text{ invited both of } x\text{'s students}]\!]^{w}(a)$. Therefore the 'all'-alternative is predicted to be blocked.

I take the moral of (18) to be that (11) must be supplemented with a projection recipe. For the moment, I'd like to propose (19) as a solution. This same form of "localization" in the face of the challenge raised by the data in (18) has also been proposed by Singh (2011) for *Maximize Presupposition!*, Spector & Sudo (2017) for *Presupposed Ignorance*, and Schlenker (2012) for *Mandatory Implicatures* (for the latter, see also section 3.2).[7]

(19) **Logical Integrity, LI.** (first version)

a. *Projection principle.* A sentence ϕ is <u>un</u>acceptable in context C if it contains a property- or proposition-denoting constituent β which violates CC in its local context with respect to one of its alternatives β'.

b. *Core Condition, CC.* A property- or proposition-denoting expression β violates CC in (its local) context C w.r.t. β' iff β' is logically non-weaker than β and β contextually entails β' in C (i.e., $\beta \not\models \beta'$ but $\beta \models_{\mathrm{C}} \beta'$).

In a nutshell, LI checks CC, which is effectively the level-neutralized version of (11), for every constituent (of the relevant type) of a given sentence, including the whole sentence.[8] Since a sentence is ruled out as soon as a CC-violation is detected, every sentence ruled out by (11) is ruled

7 Although the projection recipe (19a) is adopted by the mentioned authors without further ado, there is some choice involved in formulating it. For example, one could require that only *the smallest* property- or proposition-denoting constituent that contains a certain alternative-triggering item (such as 'some') not to violate CC. This formulation accounts for the basic facts, but let me briefly point out why it fails in general. Consider (15a) from above. The smallest constituent that contains 'some' in (15a) is the embedded clause 'some students smoke'. The local context of this expression is $\lambda w.\ \exists w' \in \mathrm{C} :\ w = w' \vee w \in \mathrm{DOX}^{w'}_{\mathbf{J.}}$. In this context there is no violation of CC (see fn. 8): to account for the oddness of (15a) one needs to check CC *at the root*.

8 In the case of '(un)aware' there are now two constituents to be taken into account, the whole sentence and the embedded clause. The embedded clause, however, is not ruled out by LI in either case. The reason is that the local context of the embedded clause include the worlds compatible with John's beliefs, and no contextual restriction is put on these; in particular, the requirement generated by LI (that it be contextually possible that some but not all students smoke) is not ruled out by the local contexts.

out by LI; but the converse is not true, e.g. every sentence in (18) is ruled out by LI. We will see that in some cases LI predicts false negatives. This problem is addressed by LI* formulated in section 5.

3 *Maximize Presupposition!* and related phenomena

3.1 Outline

In subsection 3.2 the bare-bones of the theory of *Maximize Presupposition!* (hf. MP) are laid out. The problem with the requirement, often associated with MP, that for it to be activated the relevant alternatives must be "equally informative" is discussed in some detail. Some relevant facts are reviewed in subsection 3.3 and it is argued, building in particular on Spector & Sudo (2017) (hf. S&S), that facts point in opposing directions regarding whether equal-informativeness is a necessary condition of MP. S&S's solution to this puzzle is briefly summarized in the same subsection, and subsection 3.4 closes by a discussion of two data points that are problematic for S&S's proposal.

3.2 Strawson-equivalent alternatives: Standard MP

Consider the following, fairly standard formulation of MP in which the "all else equal" proviso is explicitly cashed out as equal-informativeness. (Throughout this section attention is focused on sentences at root and therefore the formulation in (20) is not localized.)

(20) ***Maximize Presupposition!.*** Let ϕ and ψ be two alternatives such that the presupposition of ψ is stronger than ϕ. In any context C in which the following two conditions are met, one must use ψ.

a. The presuppositions of both ϕ and ψ are satisfied.

b. ϕ and ψ are equally informative.

There are at least two salient ways to precisify the notion of equal-informativeness. The option which has been adopted most widely in the literature is that of *contextual equivalence*: ϕ and ψ are contextually equivalent in C iff there is no world w in C in which one alternative is true and the other is not, $\forall w \in \mathrm{C} : [\![\phi]\!]^w = 1 \Leftrightarrow [\![\psi]\!]^w = 1$.[9] An alternative, which is strictly stronger and has also sometimes been utilized, is that of *contextual identity*: ϕ and ψ are contextually identical in C iff there is no world w in C in which the two alternatives deliver different truth-values, $\forall w \in \mathrm{C} : [\![\phi]\!]^w = [\![\psi]\!]^w$. As pointed out immediately below, the choice between the two is in fact moot at least for the classic examples that have been used to motivate MP.

Here are two classic examples that MP has been traditionally applied to.

(21) a. $^{\#}$A sun is shining. $\quad$ ALT = {The sun is shining}

b. $^{\#}$All of John's eyes are open. $\quad$ ALT = {Both of John's eyes are open}

As an example, the unacceptability of (21b) is accounted for as follows. The 'both'-alternative in (21b) has a stronger presupposition which is satisfied in normal contexts. Furthermore, the two alternatives are equally informative in normal contexts in the sense that they cannot convey

9 This is the bidirectional version of contextual entailment as defined in section 2.1.

differential information: it is impossible for one to be true and the other false. Consequently, it is predicted that the 'both'-alternative should be preferred and 'all'-alternative should be unacceptable, as desired.

In the previous paragraph, the fact that the 'all'-sentence was equally informative with its 'both'-alternative is not merely a contextual contingency, but rather a consequence of a logical relation that 'all' bears to 'both'. More specifically, 'all' and 'both' are Strawson-equivalent in the sense that it is impossible for one to be true and the other be *false* (rather than undefined). Within trivalent semantics, Strawson-equivalence is the closest one can get to the formalization of the intuition that two sentences "have the same assertive component". A bit more formally,

(22) a. Two sentences ϕ and ψ are Strawson-equivalent iff, $\forall w \in W : \{[\![\phi]\!]^w, [\![\psi]\!]^w\} \subseteq \{0,1\} \Rightarrow [\![\phi]\!]^w = [\![\psi]\!]^w$.[10]

b. Let β and β' be two lexical items of a type that "ends in t" which can take n arguments. β and β' are Strawson-equivalent iff for all objects $x_1, \ldots, x_n$ of appropriate types, if $[\![\beta]\!](x_1)\ldots(x_n) \neq \#$ and $[\![\beta']\!](x_1)\ldots(x_n) \neq \#$ then $[\![\beta]\!](x_1)\ldots(x_n) = [\![\beta']\!](x_1)\ldots(x_n)$.

Strawson-equivalence is a *logical* (i.e., acontextual) property that some sentence-pairs have in virtue of their semantics. One immediate *pragmatic* (i.e., context-dependent) consequence of it is that any context in which the presuppositions of two Strawson-equivalent sentences are satisfied is a context in which the sentences are contextually *identical* (and therefore also contextually *equivalent*), in the sense that they deliver the same truth-value at every world in the said context. Put differently, whenever MP is called upon to decide between a pair of Strawson-equivalent alternatives (as was the case with classic examples such as those in (21)), the condition of equal-informativeness, (20b), is redundant: in such cases, the condition of presupposition satisfaction, (20a), guarantees that (20b) holds in the relevant context.

This observation raises a question with far-reaching consequences for the theory of *Maximize Presupposition!*: is the condition of equal-informativeness, (20b), necessary in the formulation of MP at all? In light of the discussion above, it is clear that to answer this question one must look at *non*-Strawson-equivalent alternatives. This is the topic of the next subsection. It will be argued that facts from non-Strawson-equivalent examples point in two opposing directions. Once this tension is adequately characterized, S&S's solution to it is briefly summarized.[11]

10 This is simply the bidirectional counterpart of Strawson-entailment as defined in section 2.1.

11 One possible approach is to claim that MP is a principle that is geared exclusively toward Strawson-equivalent alternatives. There are at least two ways to cash this out: to limit the scope of MP (i) to Strawson-equivalent lexical items or (ii) to Strawson-equivalent sentences/expressions. I assume that this approach is viable only to the extent that a coherent theory with a wider coverage is not feasible. As I believe such a theory is feasible, I will not pursue this option. I'd just like to point out that options (i) and (ii) are quite plausibly distinct: *it is not the case that any two sentences that differ only in that one lexical item is replaced by one of its Strawson-equivalent alternatives, are Strawson-equivalent themselves*. Consider the sentences 'a professor brought all his students' and 'a professor brought both his students'. Under the assumption that presuppositions triggered in the scope of the indefinite project existentially (which is contested in general but very plausible in this case), in a situation in which five out of ten professors have two students and the rest have more, and one professor in the latter category brought his students and no professor with two students brought his students, the 'both'-sentence is false while the 'all-'sentence is true, hence the two *sentences* are not Strawson-equivalent even though they differ only in the substitution of one lexical item with one of its Strawson-equivalent alternatives.

3.3 Non-Strawson-equivalent alternatives: Spector & Sudo (2017) proposal

As mentioned above, while the context specified in (23) is one in which the presupposition of (23b) is satisfied, it is not a context in which the two sentences are equally informative: the possibility that all students smoke while John only believes that some students smoke is one in which (23a) is false while (23b) is true and, furthermore, it is not contextually ruled out in (23). Consequently, the two alternatives in (23) are not contextually equivalent (let alone contextually identical), which is why the contrast in (23) is entirely surprising from the point of view of MP. The latter predicts (23a) to be acceptable precisely *because* it is *not* equally informative with its competitor (23b).

(23) [Context: all students smoke.]

a. #John is unaware that some students smoke.

b. ✓John is unaware that all students smoke.

On the basis of (23) one is certainly tempted to draw the conclusion that equal-informativeness is not a relevant condition for MP: it is redundant for Strawson-equivalent alternatives (as discussed in the previous subsection 3.2) and it leads to incorrect predictions for at least some non-Strawson-equivalent alternatives such as (23). The situation, however, is more nuanced. As it happens, the distribution of some non-Strawson-equivalent alternatives can *only* be accounted for by MP if it *is* restricted by equal-informativeness. To see this, consider the distribution of the positive counter-part of 'unaware'.

(24) [Context: all students smoke.]

a. ✓John is aware that some students smoke.

b. ✓John is aware that all students smoke.

The striking *absence* of an (acceptability) contrast between the alternatives in (24) comes out clearly when we notice that, like the alternatives in (23), the alternatives in (24) are neither contextually equivalent nor contextually identical in the context specified in (24). Now, the acceptability of (24a) is immediately predicted by MP *if it is restricted by equal-informativeness*: since the alternatives in (24) are not equally informative, it is predicted that the 'some'-alternative is usable even if the presupposition of the 'all'-alternative is satisfied.

Putting these two observations together, the challenge for *Maximize Presupposition!* can be summarized as in (25). Importantly, note that this problem is indeed specific to MP; as shown in section 2.2, LI predicts the relevant facts without any analogous problems.

(25) a. If MP is restricted by equal-informativeness, why is (23a) <u>un</u>acceptable?

b. If MP is <u>not</u> restricted by equal-informativeness, why is (24a) acceptable?

Before moving on to S&S's solution to this problem, it is perhaps useful to note another piece of evidence pointing in the same direction as (24), coming from the competition between 'believe' and 'know', a center-piece of *Maximize Presupposition!* literature. It is by now a mainstream assumption that there is more to 'know' than the factive presupposition and the doxastic entailment. This extra piece of information presumably has to do with rational grounds on which the subject is reported to hold the relevant belief. The precise nature of this component is not relevant here;

the crucial assumption is that this piece of information is asserted and not presupposed. If so, then 'believe' and 'know' are *not* Strawson-equivalent, which means that there are worlds in which a 'believe'-sentence is true while its 'know'-alternative is *false* (rather than undefined). The question then arises: in a context in which the factive presupposition of 'know' is satisfied, is the 'believe'-alternative *ipso facto* deviant? A positive answer would be evidence that MP must not be restricted by equal-informativeness (as the two alternative may not be equally informative in such a context) while a negative answer would be evidence for the contrary claim.

The judgments are in general rather unclear, presumably due to the fact that the notion of justification involved in the meaning of 'know' is heavily context-dependent. However, Schlenker (2012) who discusses this exact issue in the context of *Maximize Presupposition!* comes to the conclusion that "[. . .] it is only in case 'believe' and 'know' are contextually taken to have the same assertive component that *Maximize Presupposition!* applies. In other words, the competition only arises in case the context licenses the inference: if *x* believes that p and if p is true, then *x* knows that p (the converse entailment is presumably always valid)." One piece of evidence in favor of Schlenker's claim is the following.[12]

(26) [Context: We are running an experiment pertaining to color perception. We have two red-green color blind participants (for simplicity, let us assume this means people who suffer from this condition see red as green and perceive every other color accurately) and two participants with normal color perception.]
In the first trial, we showed every participant a green mug. Therefore by the time of the second trial, every participant {$^{\checkmark}$believed, $^{?\#}$knew} that he had seen a green object.

Uncontroversially, presuppositions triggered in the scope of the universal quantifier project universally. Thus the 'know'-alternative in (26) triggers the presupposition that for every participant *x*, *x* had seen a green object. This presupposition is of course satisfied in the context of the sentence. Nevertheless, the 'believe'-alternative is impeccable. This fact can be explained on the basis of the assumption that the two alternatives are not equally informative: the 'know'-alternative in (26) presumably asserts that every participant believes *on good rational grounds* that they saw a green mug, which contradicts the assumption that half of the participants were color blind (I assume this is the reason why the 'know'-alternative in (26) is degraded).

To recap, it seems that data from non-Strawson-equivalent alternatives point to two opposing directions: either MP does not rely on equal-informativeness, in which case it can derive (23) but not (24) and (26), or MP does rely on equal-informativeness, in which case it can derive (24) and (26) but not (23).

Spector & Sudo (2017), building on Sharvit & Gajewski (2008) and Gajewski & Sharvit (2012), propose a solution to this puzzle which involves breaking the theory in two halves: one mechanism specialized to deal with presupposed content and the other with assertive content. As to the former, they propose that *Maximize Presupposition!* must indeed be reformulated so as to not rely on equal-informativeness, (27). They dub the new principle *Presupposed Ignorance*.

(27) ***Presupposed Ignorance, PI.*** Let ϕ and ψ be two alternatives such that the presupposition of ψ is stronger than ϕ. In any context C in which the presuppositions of both ϕ and ψ are satisfied, one must use ψ.

12 (26) is not one of Schlenker's (2012) examples. The choice to use (26) instead is motivated by some confounds that might be involved in his original examples.

Now, since for Strawson-equivalent alternatives equal-informativeness followed from the condition that the presuppositions of the two alternatives be satisfied, (27) and MP make identical predictions for the acceptability conditions of cases like 'all' and 'both', and 'a(n)' and 'the'. The difference, as will be seen below, kicks in only for non-Strawson-equivalent alternatives.

On the assertive side, S&S rely on a formulation of the exhaustivity operator, call it *exh**, which in effect is a presupposition hole with respect to the "innocently excludable" alternatives of the prejacent. Since the exact internal working of *exh** is not directly relevant here (and in particular I will not reproduce S&S's definition of innocently excludable alternatives), I will just briefly summarize the main features of its semantics in the particular case of when the prejacent, ϕ, has only one innocently excludable alternative, ψ. The particular novelty here, of course, is the underlined portion: the expression '*exh** ϕ' is defined to be undefined if there is an innocently excludable alternative which is undefined (or if the prejacent itself is undefined).

(28) Suppose $\text{ALT}_{\text{IE}}(\phi) = \{\psi\}$. Then,

$$[\![exh*_{\text{ALT}_{\text{IE}}(\phi)}\ \ \phi]\!]^w = \begin{cases} \# & \text{if } [\![\phi]\!]^w = \# \underline{\text{ or } [\![\psi]\!]^w = \#} \\ 1 & \text{if } [\![\phi]\!]^w = 1 \text{ and } [\![\psi]\!]^w = 0 \\ 0 & \text{if } [\![\phi]\!]^w = 0 \text{ or } [\![\psi]\!]^w = 1 \end{cases}$$

To see how (27) and (28) interact to derive the relevant facts, consider the (29) (= (23) from above).

(29) [Context: all students smoke.]
#John is unaware that some students smoke.
ALT = {John is unaware that all students smoke}

In S&S's framework, the sentence ϕ = 'John is unaware that some students smoke' can have two different parses: with and without matrix exhaustification.[13] However, since 'unaware' is Strawson-downward-entailing in its clausal complement (see fn. 5), '*exh** ϕ' is predicted to be vacuously equivalent to ϕ, because the 'all'-alternative is not innocently excludable (and there is no other alternatives by assumption). Consequently the only usable parse is the one without exhaustification. But PI, (27), predicts this parse to be unacceptable if the presupposition of the 'all'-alternative is satisfied, thereby deriving the unacceptability of (29).

Next, consider (30) (= (24) from above).

(30) [Context: all students smoke.]
✓John is aware that some students smoke.
ALT = {John is aware that all students smoke}

Here again there are two available parses, with and without exhaustification. The difference is that this time exhaustification is *not* vacuous: the parse '*exh** John is aware that some students smoke' is predicted to presuppose that all students smoke and assert that John only believes that some students smoke. Since the parse without *exh** is ruled out by PI for the same reason as before, the

13 I follow S&S in ignoring the possibility of embedded exhaustification, a potentially significant omission.

parse with *exh** can be used to "rescue" the sentence in the context of (30). This, in a nutshell, is S&S's solution to the problem raised in (25).[14]

3.4 Two challenges for S&S's proposal

3.4.1 Weak predictions

Let ϕ and ψ be two Strawson-equivalent alternatives such that the presupposition of ψ is stronger than that of ϕ. As already discussed, PI predicts that ϕ can be felicitously used in context C only if the presupposition of ψ is not satisfied in C (i.e., there is a world in C in which ψ is not defined). LI, on the other hand, makes a stronger prediction, namely that ϕ can be felicitously used in C only if it is contextually possible that ϕ is true and ψ is undefined (i.e., there is a world in C in which ϕ is true and ψ is undefined). This means that LI is more strict that PI: more sentence-context pairs are ruled out by LI than PI. In this section I'd like to argue that the relative strictness of LI makes it empirically more adequate than PI (and a modification thereof).

To begin with, consider the contrast in (31).[15]

(31) [Context: we have not established whether Mary has any students this semester, but it is common knowledge that as a rule she takes two students on at a time.]
Mary will bring {#all, ✓both} her students.

The crucial feature of (31) is that while the two alternatives are contextually equivalent, which means in particular that the 'all'-alternative contextually entails the 'both'-alternative, the presupposition of the latter (that Mary has exactly two students) is *not* satisfied in the given context; the context merely entails that *if* Mary has any students, she has exactly two. But, then, the unacceptability of the 'all'-alternative is unexpected for PI, as it predicts the 'all'-alternative to be infelicitous only if the presupposition of the 'both'-alternative is satisfied, which is not the case in (31).

The pair of examples in (32) exhibit the same feature. In (32a), the presupposition of the 'the'-alternative (that there was exactly one winner) is not satisfied; the context merely entails that *if* there was a winner, there was exactly one. In (32b), the information contained in the context is even weaker; here, the context entails that if there was a winner *who got a gold medal*, then there was exactly one winner (although in general there could have been any number of winners).

(32) a. [Context: in chess at most one player wins, if that.]
I just saw two people playing chess...
{#A, ✓the} winner was very smart.

b. [Context: there is a regional contest in which any number of contestants may win. In case only one contestant wins, the judges may decide to give him/her a gold medal.]
{#A, ✓the} winner was given a gold medal.

14 S&S assume that PI, much like LI, must be checked against every constituent (of the relevant type) to account for, e.g., (18b) in section 2.2. On the other hand, it is crucial to their account that PI not be checked below the exhaustivity operator, for otherwise they would predict (30) to be unacceptable in the given context. Consequently, they need to stipulate that PI is checked at every constituent except for those constituents that are not in the immediate scope of *exh*. As far as I can see, LI need not be modified in the same way.

15 The data points in (31) and (32) which essentially make the same point were noticed independently by, respectively, Percus (2010) and Anvari (2015).

The challenge raised by these examples is that in each case the presuppositionally weaker alternative (the 'all'-alternative in (31), the indefinite alternatives in (32)) is blocked even though the presupposition of the stronger alternative is in fact *not* satisfied. Since PI relies crucially on the presupposition of the stronger alternative being satisfied in the context, it does not predict any contrast to arise in (31) and (32), contrary to fact.[16] LI, on the other hand, correctly predicts the weaker alternatives in (31) and (32) to be blocked. For example, in (31), the 'all'-sentence is predicted to be felicitous in the given context only if it is contextually possible that it is true while its 'both'-alternative is not; i.e., only if there is a world in the context in which Mary has more than two students and invited all of them. Since this possibility is ruled out in the context of (31), the sentence is predicted to be unacceptable (the exact same reasoning applies to the cases in (32)).

Now it might be argued that PI can be reformulated to accommodate these facts quite easily. Note that in all three examples above the presuppositionally stronger alternative is felicitous. Since the relevant presuppositions are not satisfied in the given contexts, the felicity of these sentence would require some form of accommodation. But this may not be an obstacle for these items to block their presuppositionally weaker alternatives. In other words, perhaps all one needs to acknowledge is that when two alternatives compete the one with the weaker presupposition is blocked *as soon as the one with the stronger presupposition can be felicitously uttered.*[17]

(33) Let ϕ and ψ be two alternatives such that the presupposition of ψ is stronger than ϕ. For any context C, ϕ can be used in C only if ψ <u>cannot be felicitously used in C</u>.

The point of (33), of course, is that the requirement of PI that the presupposition of the stronger alternative be satisfied is unnecessarily restrictive: the sheer fact that the stronger alternative can be felicitously used is enough to block the weaker alternative, even if a felicitous use of the stronger alternative would necessitate such a mechanism as presupposition accommodation.

It is not difficult to argue against the counter-proposal in (33). In particular, since according to (33) the felicity of the stronger alternative entails the infelicity of the weaker alternative it predicts that,

(34) There is no context in which both elements of a pair of competing forms, with one presuppositionally stronger than the other, can be used felicitously.

This prediction, however, cannot be true. Consider the example (35) from Heim (1991).

(35) a. A pathologically nosy neighbor of mine broke into the attic.
 b. {The pathologically nosy neighbor of mine, my pathologically noisy neighbor} broke into the attic.

If no information is taken for granted regarding the sanity of the speaker's neighbors, both sentences in (35) can be uttered felicitously. If (33) was on the right track, however, the felicity of (35b)

16 Note that these data points are also problematic for standard *Maximize Presupposition!* as formulated in section 3.2. The reason is that regardless of how equal-informativeness is cashed out, since the presuppositions of the stronger alternatives are not satisfied, no contrast is predicted to arise by standard MP.

17 Thanks to Y. Sudo and P. Schlenker for bringing this possibility to my attention.

would have bocked (35a) as it has a stronger presupposition and it can be felicitouly asserted.[18] According to LI, (35a) is predicted to be felicitous as long as the possibility is allowed that the speaker has more than one pathologically noisy neighbors one of which broke into the attic. Since this possibility is not ruled out, the sentence is correctly predicted to be felicitous.

(35) is a fairly strong piece of evidence against (33), but there is another prediction it makes which is interesting in its own right and merits investigation.

(36) Given a pair of competing forms, if the presuppositionally stronger one is ruled out on independent grounds (and thus cannot be used felicitously), then the presuppositionally weaker alternative can in principle be used felicitously.

In other words, the prediction is that no example can be constructed with the following features: (i) the presuppositionally weaker alternative is infelicitous due to competition with its stronger alternative, and (ii) the presuppositionally stronger alternative itself is infelicitous due to some independent reason. Corroborating this prediction is a rather complex matter due to potential confounds. Before discussing one possible attempt, let us note that while LI, as it stands, can allow two expressions to compete even if one of them is rendered infelicitous for independent reasons, nothing prevents the stipulation of a functional constraint on LI to avoid this;[19] after all, if one alternative is ruled out independently, there is in principle no reason to allow it to enter into the competition. The possibilities, then, are as follows: (i) a counter-example to (36) can be constructed where the weaker alternative is blocked even though the stronger one is independently ruled out, in which case LI as it stands is sufficient and (33) is made even less plausible, (ii) no such counter-example can be constructed, in which case (33)'s prediction (in this case) is corroborated and LI needs to be coupled with a suitable constraint. Here is one attempt at constructing the relevant counter-example.

(37) [Context: an excursion was planned by the department. Of the 70 professors who were invited, 30 have exactly two students and the rest have more. Due to lack of space, only professors who have exactly two students were allowed to invite zero, one, or both of their students, the choice being up to them.]
The week after the excursion the following dialogue takes place.
Q: How many guests attended the excursion?
A: In total, 60 professors showed up. However, not all of them invited {$^{?}$all,$^{\#}$both} their students. (Therefore, the number of attendees was rather less than we expected.)

The infelicity of the 'both' in (37) is entirely expected. As seen in (38), presuppositions triggered in the scope of 'not all' project universally to root, meaning that in (37) the 'both'-alternative presupposes that every professor (who showed up) has exactly two students. This presupposition directly contradicts the preceding discourse (if 60 professors showed up, then at least half of the

18 This argument applies to the relaxed version of PI, i.e. (33), specifically. In principle, if one makes the same move for MP, given the restriction to MP to equally informative alternatives, the result might yield an account of the puzzle raised in this subsection. Importantly, however, the original problem raised for MP, i.e. (25), will not be addressed by such a maneuver. Therefore, LI aside, the choice points are these: (i) PI solves (25) but cannot be made to solve (31) and (32), and (ii) MP *might* be able to solve (31) and (32) but cannot solve (25).

19 In fact, LI*, the final version of LI, involves adding a (different) functional constraint of this sort on top of LI.

professors who showed up have more than two students, under the assumption that *in toto* only 30 professors have exactly two students), hence the infelicity of the 'both'-alternative.[20]

(38) Not all students stopped smoking.
⇝ Every student used to smoke

Now, what about 'all'? The judgments are unclear. Those I have consulted report that 'all' in the context of (37) is not as infelicitous as 'both' but not entirely felicitous either. My feeling, for what is worth, is that the perceived oddness of 'all' negatively correlates with the salience of the oddness of 'both'. As the choice points are made clear above, and as strong evidence against (33) is already available (i.e., (35)), I take this question to be somewhat orthogonal and leave the further investigation of examples like (37) to future research.[21]

3.4.2 Strawson-equivalent alternatives with non-monotonic presuppositions

The data pointed out here is problematic for MP, PI *and* for LI as currently stated. However, once LI*, the final version of LI, is formulated in section 5 in reaction to some (loosely related but) independent data, it will be shown that a correct prediction regarding this data point will naturally fall out without further ado. Consider the contrast in (39).

(39) [Context: two students solved all of the math problems and the rest solved none.]

a. #Both students who solved some of the math problems passed.
⇝ Exactly two students solved some of the math problems (presupposition)

b. ✓Both students who solved all of the math problems passed.
⇝ Exactly two students solved all of the math problems (presupposition)

Three features of the two sentences in (39) are crucial. First, note that the presuppositions triggered by these two sentences are logically independent.[22] In a world in which two students solved all of the math problems, one solved only some, and the rest solved none, the presupposition of (39a) is false while that of (39b) is true. In a world in which one student solved all of the math problems, one solved only some, and the rest solved none, the presupposition of (39a) is true while that of (39b) is false. Second, note that both presuppositions are satisfied in the context specified in (39).

20 One might wonder why quantifier domain restriction does not kick in to rescue the 'both'-sentence. That is, why the domain of the quantifier 'not all' is not covertly restricted to professors who showed up *and* have exactly two students? This would certainly rescue the 'both'-alternative from infelicity. As pointed out by C. Ebert (pc), the absence of covert restriction in this case is probably due the partitive structure, in particular the plural pronoun in the restrictor of the quantifier.

21 Here is a quick sketch of why LI, unconstrained, predicts 'all' to be unacceptable in (37). The predicted requirement is that the local context of the predicate '$\lambda x.\ x$ invited all x's students' must allow for the possibility of there being at least one professor who has more than two students and invited all of them. Since the local context (viewed extensionally) contains all and only those professors who showed up, the requirement boils down to there being a world in the context in which there is at least one professor who showed up, who has more than two students and who invited all of his students; this possibility being ruled out by the context, the 'all'-alternative is predicted to be unacceptable.

22 The intended reading of (39) is that both students who solved *at least* some of the math problems passed. This data point can be viewed (and accounted for) from the point of view of Magri's (2009a) account generalized to presuppositions. Magri's account will be discussed in detail in section 4. The point of (39) is the problem it poses for MP and PI.

Third, and finally, note that (39a) and (39b) are Strawson-equivalent: it is logically impossible for one to be true and the other be *false* (rather than undefined).

Given these three observations, how do MP/PI fair with respect to the contrast in (39)? Since MP/PI encode a preference for the presuppositionally *stronger* alternatives, they, in fact, do not dictate any preference between (39a) and (39b) at all, precisely because neither presupposition is stronger than the other one (since the two are logically independent). What if MP/PI are modified to be sensitive to *non-weaker* presuppositions, rather than stronger ones, along the lines of (40)?

(40) Let ϕ and ψ be two alternatives such that (thy encode the same assertive content but) the presupposition of ψ is <u>not weaker</u> than ϕ. For any context C, ϕ can be used in C only if the presupposition of ψ is not satisfied in C.

This move will certainly capture the oddness of the 'some'-alternative, (39a): the presupposition of the 'all'-alternative is not weaker than the presupposition of the 'some'-alternative (point one) and the two encode the same assertive content (point three), therefore since the presupposition of the 'all'-alternative is satisfied (point two), the 'some'-alternative is predicted to be unacceptable by (40). The problem is that the exact same reasoning now applies to the 'all'-alternative, (39b), predicting that it too should be unacceptable, contrary to fact. Thus under a principle such as (40), both sentences (39a) and (39b) are predicted to come with the requirement that the context must allow there being students who solved only some of the math problems, a prediction which is only true for (39a).

In a nutshell, the challenge posed by (39) for MP/PI is this: either the principle dictates a preference for presuppositionally stronger alternatives, in which case it predicts both sentences in (39) to be acceptable, or it dictates a preference for presuppositionally non-weaker alternatives, in which case it predicts the two alternatives in (39) should "cancel each other out". At this point, it is not difficult to see that LI has the second problem: it incorrectly predicts both sentences in (39) to be unacceptable. Nevertheless, the seeds of a possible solution from the vantage point of LI are already present. Note that the 'all' alternative is "less bad" than the some alternative because the former violates CC only at root while the latter both at root *and* in the restrictor (because the predicate 'students who solved some of the problems' contextually entails its 'all'-alternative). To jump ahead, for LI* this very observation is operationalized to break the tie against 'some' thus capturing the contrast in (39).

4 Mandatory implicatures

4.1 Basic cases and Magri's (2009a) proposal

The most straightforward illustration of the basic puzzle with which this section is concerned is offered by the pair of contrasts in (41).

(41) a. [Context: John always gives the same grade to all his students.]
This semester, he gave {#some, ✓all} of his students an A.

b. [Context: John and Mary traveled from Vienna to Paris together.]
John {#or, ✓and} Mary traveled by train.

The adherent of a principle such as *Maximize Presupposition!* (or, its cousin, *Presupposed Ignorance*) must invoke a different principle to account for the contrasts in (41) because the relevant alternatives in these examples certainly do not carry different presuppositions. However, since LI is not specifically keyed to presuppositions *per se*, it makes predictions here as well. The case of (41a) was discussed in section 2.2. As for (41b), the 'or'-sentence is predicted to be unacceptable in the given context because (i) it does not logically entail its 'and'-alternative, therefore (ii) LI predicts 'or' to be usable only contexts in which the possibility that 'or' is true and 'and' is not is allowed (that is, the possibility is allowed that John or Mary <u>but not both</u> traveled by train), but (iii) since this possibility is explicitly ruled out in the context of (41b) the 'or'-sentence is predicted to be unacceptable.

In Magri (2009a,b) a novel proposal is advanced on the basis of a set of constructions including (but far from limited to) (41). Intuitively, according to this proposal the 'some'-sentence in (41a) and the 'or'-sentence in (41b) are odd because in their contexts they are *obligatorily* interpreted, respectively, as John gave an A to some *but not all* of his students and John or Mary *but not both* traveled by train, thus contradicting the background assumptions that John always gives the same grade to his students and John and Mary traveled together. The axioms, as it were, of this proposal are aggregated in (42).

(42) *Mandatory Implicatures* (hf.MI)

a. There is a covert exhaustivity operator, *exh*, that is *obligatorily* attached to every possible scope site.[23]

b. *exh* operates solely on the basis of *logical* (as opposed to contextual) entailment.

c. To any expression ϕ, a set of alternatives is associated, $\textsc{alt}(\phi)$.[24] In general, the "domain" of *exh* is restricted to a subset of $\textsc{alt}(\phi)$, i.e. the "innocently excludable" alternatives (Fox 2007), $\textsc{alt}_{\text{IE}}(\phi)$. Intuitively, this is the collection of those alternatives that, in conjunction with the "prejacent" ϕ, can be consistently negated.

d. Let ϕ be some expression. *exh* ϕ always entails that ϕ is true. For any $\psi \in \textsc{alt}_{\text{IE}}(\phi)$, *exh* ϕ entails that ψ is false if and only if ψ is relevant.

e. Any sentence used is relevant.[25] In a context in which two propositions are equally informative, one is relevant and only if the other is relevant as well.

For the present purposes, it suffices to postulate that ψ is "innocently excludable" with respect to ϕ iff ψ is logically non-weaker than ϕ. Therefore,

(43) a. $[\![exh\ \phi]\!]^w = 1$ iff $[\![\phi]\!]^w = 1 \wedge \forall \psi \in \textsc{alt}_{\text{nw}} : ([\![\psi]\!]^w = 0 \vee \neg \text{Rel}(\psi))$

b. $\psi \in \textsc{alt}_{\text{nw}}(\phi) \Leftrightarrow \psi \in \textsc{alt}(\phi) \wedge \phi \not\models \psi$ (for '$\models$', see (7) from section 2.1)

23 Any insertion of an exhaustivity operator introduces a new possible scope site for the insertion of yet another exhaustivity operator; assuming '•' stands for a possible scope site, $\bullet[\phi] \longrightarrow \bullet[exh[\phi]] \longrightarrow \bullet[exh[exh[\phi]]] \longrightarrow \ldots$. Magri's claim makes sense only if it is interpreted as dictating that only the first layer of exhaustification is obligatory, "recursive" exhaustification being optional.

24 For the present purposes a simple replacement mechanism suffices to generate alternatives.

25 Assuming that the speaker is being cooperative, he wouldn't use the sentence if it weren't relevant.

Putting all this together, here is how this system accounts for the contrast in (41a). First, the 'some'-sentence is obligatorily parsed as (44a). Second, its only non-weaker alternative is the 'all'-sentence, (44b). Third, since that alternative is *logically* stronger than the 'some'-sentence, it will give rise to a scalar implicature iff it is relevant. Fifth, the 'some'-sentence is relevant by assumption. Since the 'some'-sentence and its 'all'-alternative are contextually equivalent,[26] the 'all'-alternative is relevant as well. Conclusion: (44a) is obligatorily interpreted as (44c), thereby contradicting the background assumptions in (41a) and incurring oddness.

(44) a. $[exh\ \underbrace{\text{John gave some of his students an A}}_{\exists}]$

b. $\text{ALT}_{\text{nw}}(\exists) = \{\ \underbrace{\text{John gave all of his students an A}}_{\forall}\ \}$

c. $[\![\ (44a)\ ]\!]^w = 1$ iff $\exists \wedge \neg\forall$. ($\lightning$ in the context of (41a))

There is a sense in which, according to this proposal, the 'some'-sentence in the context of (41) is odd because in such a context the exhaustification system short-circuits. It is an elegant analysis to the extent that it is designed by a particular arrangement of more or less independently motivated ideas enumerated in (42). It will be helpful to flag a feature of Magri's system which was first pointed out by Schlenker (2012). For Magri's system to work in full generality, it has to be "localized" in two ways; not only does it rely on a covert exhaustivity operator which can be inserted in embedded positions, it also *must be* relativized to local contexts. To see the latter, consider (45), a variation of the example used by Schlenker (2012).

(45) No teacher who will assign the same grade to each of his students wants to give {#some, ✓all} of them an A.

Note that at root, the 'some'-sentence is at least as strong as its 'all'-alternative, if not stronger, given any reasonable rendition of the notion of entailment.[27] Therefore no exhaustification-related inference is expected to arise at root. The only way to capture the oddness of the 'some'-sentence within Magri's system is to make sure that the embedded *exh* triggers a local implicature. Now this inference *may* be triggered, but it is *not* obligatorily: we have no means to guarantee that the alternative to the scope expression, namely the predicate '$\lambda x.\ x$ has decided to give all of x's students an A', is relevant.[28,29] The reason is that the *global* context does not guarantee that this predicate is

26 While the 'all'-sentence contextually entails the 'some'-sentence in *any* context, in the particular context spelled out in (41) the reverse holds as well since the possibility of John giving some but not all of his students an A is ruled out.

27 Schlenker's (2012) original example involves the universal quantifier 'every' instead of the negative existential 'no' in (45). As he points out, 'every teacher who assigned the same grade to each of his students gave some of them an A' is logically equivalent with its alternative 'every teacher who assigned the same grade to each of his students gave all of them an A', and therefore no implicature is predicted to arise in this case as well. However, this example allows for the possibility of using a more abstract notion of logical entailment *à la* Gajewski (2004), which would ignore the identity of predicates involved. If Magri makes this move, then Schlenker's 'some'-sentence would *not* logically entail its 'all'-alternative, opening the possibility of deriving the target implicature. The point made with the example in (45) will go through even if Gajewski-type LFs are used because the scope of 'no' is a downward-entailing environment, a fact to which even Gajewski-type entailment is sensitive.

28 For the sake of discussion I am assuming it makes sense to talk about the relevance of properties, as well as propositions, although this claim can be plausibly challenged. Magri is certainly commited to this assumption.

29 In principle, there may be more than one reason for a proposition/property to be relevant, but Magri does not make any proposals which can be applied to (45) other than contextual equivalence.

contextually equivalent with its alternative 'λx. x has decided to give some of x's students an A'. On the other hand, if relevance is assumed to be closed under contextual equivalence relative to, not just *global* contexts, but *local* contexts as well, then the system is predicted to obligatorily trigger the mismatching implicature.[30] Indeed relative to the local context of the scope (which, viewed extensionally, is the set of professors who will give the same grade to all their students) the scope predicate 'λx. x has decided to give some of x's students an A' is equivalent with its alternative 'λx. x has decided to give all of x's students an A'.

I therefore conclude with Schlenker (2012) that, in general, *local* contexts are the information sources relative to which expressions and their alternatives are checked with respect to contextual equivalence. The following point is worth mentioning. Magri makes the stipulation that relevance is closed under contextual equivalence, a plausible enough assumption as far as *sentences* are concerned. But once subsentential constituents are taken into consideration a conceptual glitch is introduced into the system: on the one hand relevance is typically understood as a speech-act level phenomenon, blind to the internal constitution of sentences; on the other hand, to the extent that the theoretician is willing to live with local contexts (either *à la* Heim 1983 or *à la* Schlenker 2009), contextual equivalence can be (and, indeed, in light of (45) must be) localized. It therefore seems that Magri is committed to the claim that relevance can be manipulated both inter- *and* intra-sententially (see also Katzir & Singh 2015 for some pertinent discussion), see also fn. 30.

4.2 Problems with *Mandatory Implicatures*

4.2.1 The problem of absence of primary implicatures

The set of assumptions that Magri subscribes to, (42), yield the following lemma.

(46) If ψ is an innocently excludable alternative of ϕ, then any utterance of ϕ will trigger the scalar implicature that $\neg\psi$ if and only if ψ is relevant.

Another assumption, which is widely held in the literature and which Magri shows no inclination to modify is (47).

(47) Upon an utterance of ϕ, for any alternative ψ of ϕ to trigger a primary implicature,[31] it is necessary that ψ be relevant.

The assumption (47) coupled with the lemma (46) immediately entails (48).

(48) No innocently excludable alternative can ever trigger a primary implicature which is not as strong as a scalar implicature.

30 Note that for this reasoning to work, one needs to rely on the following modified version of (42e): whenever a cooperative speaker utters a sentence ϕ, any proposition- or property-denoting constituent of ϕ is relevant. Again, this assumption can be plausibly challenged.

31 Following Sauerland (2004), I assume primary implicatures are inferences of the form $\neg B_S\psi$, as opposed to scalar (or secondary) implicatures which have the form $B_S\neg\psi$ and ignorance implicatures which have the form $\neg B_S\psi \wedge \neg B_S\neg\psi$, where 'S' stands for the speaker. Note that under this conception scalar implicatures are strictly stronger than primary implicatures. This is why in (48) it is necessary to add "which is not as strong as as a scalar implicature".

The reason is this: a given innocently excludable alternative *either* is relevant in which case (according to (46)) it triggers a scalar implicature *or* is not relevant in which case (according to (47)) it triggers neither a primary nor a scalar implicature.[32] The problem, of course, is that innocently excludable alternatives in general *do* trigger primary implicatures that are not as strong as scalar implicatures. For example, consider (49).[33]

(49) [Context: John has been given eight cards. Upon receiving the cards, he throws a coin. Depending on which side of the coin comes up, he either looks at two of the cards or all of them. After this, he reports:] Some of my cards are hearts.
$\rightsquigarrow$ Either John has seen all of his cards and only some of them are hearts or he has only seen two of the cards and at least one of them is hearts.

In a nutshell, the inferences is that either John is ignorant about the truth of the alternative 'all of my cards are hearts' or he believes it to be false. This corresponds exactly to the primary implicature; a primary implicature, by definition, has the form $\neg \mathrm{B}\forall$ which is equivalent to $(\mathrm{B}\neg\forall) \vee (\neg\mathrm{B}\neg\forall \wedge \neg\mathrm{B}\forall)$.

Call the problem raised above the problem of Absence of Primary Implicatures, API. In the following subsection, I will consider a possible solution to API building on Meyer (2013).[34] I will show that while a modification of Magri's theory *à la* Meyer solves API it introduces another problem having to do with strictly local mismatching implicatures.

4.2.2 Magri *à la* Meyer

A possible way to solve API in Magri's framework is to adopt a system in which even primary implicatures are computed grammatically. Such a system has been proposed in Meyer (2013) under the title of the Matrix-K theory. The fundamental assumptions of this proposal are aggregated in (50).

(50) a. There is a covert exhaustivity operator, *exh*, that *may* be attached to any scope site.
b. *exh* operates solely on the basis of *logical* (as opposed to contextual) strength.
c. Every sentence is obligatorily parsed with a K-operator at root, where K is the doxastic operator associated with the speaker.[35]
d. The Principle of Epistemic Transparency, ET: a logical form $[\dots \mathrm{K}\phi]$ is ruled out unless for every $\psi \in \textsc{alt}(\phi)$, $[\dots \mathrm{K}\phi]$ *logically* entails either $\mathrm{K}\psi$ or $\neg\mathrm{K}\psi$.

I will use the particularly simple case of the existential quantifier 'some' for illustration.[36] According to the Matrix-K theory, exhaustification is optional while the K-operator is obligatory at root.

32 This observation about Magri's system was made independently by Meyer (2013) and Anvari (2015).

33 Example (49) due to B. Spector (pc). See also Dieuleveut et al. (2016) for experimental results on this very question, and (Meyer 2013: chapter 3) for further examples that (attempt to) make the same point.

34 It goes without saying that the discussion in the next section will not do justice to Meyer's proposal, which has to be evaluated on its own merits. My goal is merely to point out that the solution to API is not as straightforward as amending Magri's theory with Meyer's.

35 Thus, for any sentence ϕ, $[\![\mathrm{K}\phi]\!]^w = 1$ iff $\forall w' \in \textsc{dox}^w_{\mathrm{S}} : [\![\phi]\!]^{w'} = 1$, where S is the speaker.

36 The full force of Meyer's theory kicks in when disjunctions are taken into consideration. I use the simpler case of 'some' for expository purposes.

Consequently, any utterance of (51) is ambiguous between (at least) four syntactically distinct LFs, (51a) to (51d). The semantic value of these LFs are computed relatively straightforwardly, as indicated by the relevant equivalences.

(51) John gave an A to some of his students ($= \exists$).

a. $\mathrm{K}\exists$

b. $\mathrm{K}\mathit{exh}\exists \equiv \mathrm{K}[\exists \wedge \neg\forall]$

c. $\mathit{exh}\mathrm{K}\exists \equiv \mathrm{K}\exists \wedge \neg\mathrm{K}\forall$

d. $\mathit{exh}\mathrm{K}\mathit{exh}\exists \equiv \mathrm{K}[\exists \wedge \neg\forall]$

e. Context: John always gives the same grade to each of his students.

Now, the simplest parse (51a) is ruled out by the principle ET, since it does not *logically* entail either $\mathrm{K}\forall$ or $\neg\mathrm{K}\forall$, while the other three parses do satisfy ET. Furthermore, the most complex parse (51d) ends up being semantically equivalent with (51b) and can, therefore, be ignored. Consequently, Meyer predicts that any utterance of (51) can be either interpreted as $\mathrm{K}[\exists \wedge \neg\forall]$ or $\mathrm{K}\exists \wedge \neg\mathrm{K}\forall$. Now suppose (51) is uttered in the context of (51e). Both of the readings predicted by Meyer contradict background assumptions in this context. The crucial reason for this is that any proposition that is part of the commonground is, by definition, part of the beliefs of each interlocutor, including the speaker. Thus, if we adopt parse (51b) we infer that the speaker believes that John gave an A to only some of his students which clashes with the assumption that the speaker (like the other interlocutors) believes that John gave the same grade to his students. If the adopt parse (51c) we infer that the speaker believes that John gave an A to some of his students, which with the assumption that the speaker believes that John gave the same grade to each of his students, yields the inference that the speaker believes that John gave an A to all of his students, contradicting the entailment of (51c) that $\neg\mathrm{K}\forall$. Since no available parse of (51) is predicted to be felicitous in the context of (51e), the utterance as whole is predicted to be infelicitous in that context.

The upshot of all this, of course, is that while Meyer's account is close to the spirit (if not the letter) of Magri's account, unlike Magri, she does *not* predict that innocently excludable alternatives never give rise to primary implicatures; in the particular case of (51), such an inference is indeed predicted to be possible in general due to the availability of the parse in (51c). Importantly, *it is absolutely crucial in Meyer's system that exhaustification be optional*,[37] otherwise the only available parse for (51) would be (51d), which is undesirable as it would bring us back to square one.

Interestingly, while in Meyer's framework one tenet of Magri's system, namely the obligatoriness of exhaustification, is discarded, she *has to* double down on the other tenet of his system, namely blindness to contextual information. In Meyer's system, not only does *exh* exclusively rely on logical entailment, but the principle ET (50d) also has to rely exclusively on logical entailment. The latter is in fact crucial: had ET relied on *contextual* entailment, the parse in (51a) would become available in the context of (51e), since in that context (51a) *contextually* entails $\mathrm{K}\forall$, leading to the incorrect prediction that (51) is felicitous in the context (51e).[38]

37 More precisely, it is only absolutely crucial that exhaustification be optional below and above the matrix K. See also fn. 39.

38 The question of just why a pragmatic principle, or at least an ambiguity resolution heuristic, such as ET *must* be blind to contextual information is not addressed by Meyer and is rather implausible. One might wonder whether ET can be

However, an extension of Magri's system *à la* Meyer, as sketched above, cannot be the end of the story. Meyer's account comes at a cost: since her system relies on exhaustivity *not* being obligatory, her system cannot account for the cases where unacceptability is crucially due to embedded implicatures. To see this, consider (52) repeated from (45) above.

(52) #No teacher who will assign the same grade to each of his students wants to give some of them an A.
K $\text{no}(\alpha, \exists) \equiv exh\text{K no}(\alpha, \exists) \equiv \text{K}exh\text{ no}(\alpha, \exists) \equiv exh\text{K}exh\text{ no}(\alpha, \exists)$

α = teacher [λx [x will assign the same grade to each of x's students]] $\exists = \lambda x$ [x has decided to give some of x's students an A]

The problem raised by (52) is that (i) all four possible LFs for (52) in Meyer's system are logically equivalent, (ii) none of the LFs are ruled out by principle ET, and (iii) the meaning predicted for (52) does not generate any contradiction, contextual or otherwise. Regarding (ii), note that the principle ET in particular is satisfied because the scalar item $\exists$ is embedded in the scope of 'no', a DE environment. Since the K operator itself a UE environment, '$\text{no}(\alpha, \exists)$' logically entails '$\text{no}(\alpha, \forall)$', and therefore none of the available parses violate ET. In a nutshell, the diagnosis is this: to rule out the 'some'-sentence in (52) in this system, it is necessary that an embedded *exh* be present in the scope of 'no' (i.e., the parse '$\text{no}(\alpha,\ exh\exists)$'), but since exhaustification is optional in Meyer's system nothing forces this to happen.[39]

To recap, the situation is this: we noticed a rather fundamental problem with Magri's original system, namely API. We sought to remedy that problem by adopting Meyer's Matrix-K theory, but by doing so we lost the empirical coverage of Magri's original theory. I conclude, pending further research (fn. 39), that the challenge raised by API does not have a trivial solution as things stand and is a major problem for Magri's account.

4.2.3 Extension to *Maximize Presupposition!*

I will now switch back to Magri's original account, assuming that API is not an immediate problem. The question addressed in this section is wthether Magri's account be generalized (as Magri has suggested) to cover the *Maximize Presupposition!*-related phenomena that were discussed in section 2 adequately? My impressions is that such extensions are not likely to work, in particular for

somehow removed from her framework. I believe this is technically possible. One can simply drop ET and assume every sentence ϕ is obligatorily parsed as $exh\text{K}(exh)\phi$, with the *exh* above K being obligatory and the one below K optional. This "Matrix-*exh*K" system will make the same predictions as Meyer's original because (i) for any sentence ϕ, the two parses Kϕ and K*exh*ϕ are either ruled out by ET or are logically equivalent with *exh*Kϕ or *exh*K*exh*ϕ, and (ii) *exh*Kϕ and *exh*K*exh*ϕ are never ruled out by ET.

39 B. Spector (pc) points out the possibility of adopting Meyer's system but with the assumption that *only matrix exhaustification optional*: *exh* is obligatorily attached to every non-matrix scope site. See also fn. 37. Indeed, such a system can derive the oddness of (52) while solving the API. I leave a serious elaboration of a theory along these lines to the theoretician who has a stomach for it.

two rather foundational reasons: the fact that *exh* is an extensional operator and the fact that the "domain" of *exh* needs to be restricted to innocently excludable alternatives.[40]

First, since *exh* is an extensional operator an LF such as '*exh* all chairs in that room are broken' (assuming the relevant alternative is 'both chairs in that room are broken') will either presuppose or assert that there are more than two chairs in the room. This is in contrast to the traditional analysis of MP effects according to which an utterance of 'all chairs in that room are broken' is acceptable only if it is not common ground that there are exactly two chairs in that room; i.e., only if there is a world compatible with background assumptions in which there are more or less than two chairs in the room. Consequently, an MP inference as traditionally conceived is neither assertive (since it imposes a constraint on the common ground) nor presuppositional (since the constraint it imposes is that some piece of information *not* be entailed by the common ground, as opposed to presuppositions which are requirements that some piece of information be entailed by the common ground). Consider the following (see also Sauerland 2008).

(53) No professor invited all of his students.

a. No professor has more than two students and invited all of them.

b. Every professor has more than two students and None invited all of them.

Let us assume that the MP inference associated with the predicate '$\lambda x.\ x$ invited all of x's students' is that x has more than two students. If this inference was assertive (53) would be interpreted as (53a) and if it was presuppositional it would be interpreted as (53b). I can think of no context in which an utterance of (53) is understood as (53a) or (53b). Regarding the latter, of course it is true that any context that entails that every professor has more than two students is one in which (53) can be uttered felicitously but this is part of the much broader generalization that an utterance of (53) is felicitous as soon as the possibility is allowed that there is at least one professor who has more than two students. More to the point, it seems to me that no context in which there is ignorance about the number of students that each professor has, is one in which (53) can be interpreted as (53b).

Second, since it is universally assumed that the domain of the operator *exh* must be systematically restricted to innocently excludable alternatives (to avoid the "symmetry problem"), no account based on *exh* predicts competition with alternatives that are ruled out by such restrictions. Building on Spector & Sudo (2017), I will now show that disjunctions in general do in fact compete with their individual disjuncts, a fact that cannot be captured by any *exh*-based theory (see also Spector (2014) for a discussion several similar facts). As S&S point out, the sentence in (54) is unacceptable in the given context.

(54) [Context: Mary lives in Paris.]
#John is unaware that Mary lives in Paris or London.
$\textsc{alt} = \{\underbrace{\text{John is unaware that Mary lives in Paris}}_{\textsc{l}}, \underbrace{\text{John is unaware that Mary lives in London}}_{\textsc{r}}, \ldots\}$

Note that the alternatives L and R are can*not* possibly be innocently excludable with respect to (54). The reason is that since the clausal complement of 'unaware' is Strawson-upward-entailing, truth

40 There is a recent proposal, advanced in Marty (2017), which relies on a particularly radical modification of the exhaustivity operator. Marty's proposal covers much of the same ground as my own proposal here, however it too faces the two challenges raised in this subsection.

of (54) guarantees non-falsity of L and R: it is impossible for (54) to be true and L and R be false. Furthermore, it is also impossible for (54) to be true and L and R to be simultaneously undefined (since the latter entails that (54) is undefined as well). Consequently, it is not possible to attribute the oddness of (54) to any form of exhaustification.[41]

Note that LI immediately predicts the judgment reported in (54) on the basis of the competition of the sentence in (54) with L. In fact, LI even predicts the oddness of (55).

(55) [Context: Mary lives in France, we do not know where.]
#John is unaware that Mary lives in Paris or London.
$\text{ALT} = \{\underbrace{\text{John is unaware that Mary lives in Paris}}_{\text{L}}, \underbrace{\text{John is unaware that Mary lives in London}}_{\text{R}}, \ldots\}$

The difference between (54) and (55) is that in the latter it is not even commonground that Mary lives in Paris, only that she lives somewhere in France. In this context, as well as the context in (54), the target sentence contextually entails the L-alternative, which is furthermore logically non-weaker. Therefore, LI predicts the target sentence to be odd both in the context of (55) and the context of (54).

4.3 Context and polarity

In Magri (2009a) a pair of very insightful contrasts is discussed which cannot be captured by LI as it stands. Here the facts are described and in section 5 a modification of LI is presented which can accounted for them. The exposition in this section relies heavily on Spector (2014). First, consider the sentences in (56).

(56) a. Every professor who assigned an A to some of his students got a pay raise.
b. Every professor who assigned an A to all of his students got a pay raise.

In (56) the scalar items ‘some’ and ‘all’ are embedded in the restrictor of the universal determiner ‘every’, which is a downward-entailing environment: the ‘some’-sentence (56a), ‘every$(\exists,\beta)$’, is strictly logically stronger than its ‘all’-alternative (56b), ‘every$(\forall,\beta)$’. Now, suppose C is a context in which it is taken for granted that no professor gave an A to only some of his students. In such a contexts, two things happen: (i) the restrictor ‘professor who assigned an A to some of his students’ contextually entails its logically stronger alternative ‘professor who assigned an A to all of his students’ in C, and (ii) the sentence (56b), at root, contextually entails its logically stronger alternative (56a) in C. Given these facts, LI predicts (56a) to be unacceptable in context C because it contains a constituent which violates CC, namely, the restrictor; (56b) is also predicted to be unacceptable in C because it also contains a constituent which violates CC, namely, the whole sentence. The prediction then, is that in a context like C both sentences in (56) should be unacceptable. Unfortunately this prediction is only half-true:

41 This problem might have a solution if Meyer’s Matrix-K is adopted. Making this move, however, requires not only a solution to the problem raised in the previous subsection but also an extension of Meyer’s framework that can handle presuppositions.

(57) [Context: In this department, every professor assigns the same grade to all of his students.]
Every professor who assigned an A to {#some,✓all} of his students got a pay raise.

Intuitively, what makes (57) particularly interesting is that it is the 'all'-sentence which is judged as acceptable in the specified context, i.e. the logically weaker sentence.

Next, consider a different context C'. Suppose C' is a context in which the possibility of some professors giving an A to only some of their students *is allowed*, but it is nevertheless a context in which the two sentences in (56) are contextually equivalent. What are the predictions of LI in this context? (56a) is predicted to be acceptable, because it does not contains any non-weaker constituent that contextually entails its alternative (in its local context). (56b), on the other hand, is predicted to be unacceptable because, at root, it is logically weaker than (56a) but in C' it contextually entails (56a), a clear violation of CC. These predictions are indeed born out:

(58) [Context: Every year, the dean has to decide: if the college has made enough profit that year, he gives a pay raise to every professor who has assigned an A to at least some of his students; if there is not enough money, then no one gets a pay raise.]
Every professor who assigned an A to {✓some,#all} of his students got a pay raise.

Just how Magri manages to correctly predict the contrasts in (57) and (58) is not directly relevant here. My main concern is to modify LI to fix the incorrect prediction it makes regarding (57). This is what I turn to in the next section.

5 Logical Integrity (final version)

In subsection 4.3 it was established that, contra LI (repeated below from section 2.2), a sentence is *not* ruled out as soon as it contains a constituent that violates CC.

(59) **The Principle of Logical Integrity, LI.** (tentative)

a. *Projection principle.* A sentence ϕ is unacceptable in context C if it contains a property- or proposition-denoting constituent β which violates CC in its local context with respect to one of its alternatives β'.

b. *Core Condition, CC.* A property- or proposition-denoting expression β violates CC in (its local) context c w.r.t. β' iff β' is logically non-weaker than β and β contextually entails β' in c (i.e., $\beta \vDash \beta'$ but $\beta \nvDash_c \beta'$).

The moral of the data discussed in section 4.3 is that (59) is too strict: it rules out too many sentences/context pairs. My strategy in response to this problem is to restrict the alternatives that CC is sensitive to. In a nutshell, I'd like to implement the idea that if an alternative itself contains a CC violation it should be ignored. The result is the following modification to LI.

(60) **The Principle of Logical Integrity, LI*.** (final)

a. *Projection Principle.* A sentence ϕ is unacceptable in context C if it contains a property- or proposition-denoting constituent β which violates CC* in its local context with respect to one of its alternatives β'.

b. *Restrictive Core Condition, CC*.* A property- or proposition-denoting expression β violates CC* in (its local) context c w.r.t. β' iff (i) β violates CC in c w.r.t. β' and (ii) β' itself does not contain any constituent that violates CC in c w.r.t. any of its alternatives.

c. *Core Condition, CC.* A property- or proposition-denoting expression β violates CC in (its local) context c w.r.t. β' iff β' is logically non-weaker than β and β contextually entails β' in c (i.e., $\beta \vDash \beta'$ but $\beta \nvDash_c \beta'$).

As the reader can easily verify, both (59) and (60) rely on the same formulation of CC and the projection principle is also essentially the same. The only difference between (59) and (60) is the CC* layer in (60) which acts as a filter on the alternatives that CC "sees". Let me spell out the difference explicitly. (59) rules out too many sentences: as soon as a sentence contains a constituent which violates CC, the sentence is ruled out. (60) is more conservative. If a sentence ϕ contains a constituent β that violates CC with respect to its alternative β', (60) asks: does β' itself contain a constituent which violates CC? Only if the answer to that question is *no* is the sentence ruled out. Put differently, while LI (=(59)) CC-checks any constituent of ϕ w.r.t. *any* of its alternatives, LI* (=(60)) CC-checks any constituent of ϕ w.r.t. *only those of its alternatives that do not contain a* CC violation. As a consequence, (60) rules out fewer sentences than (59). I will now show that (60) captures three data points that were problematic for LI.

First, let us go back to the problematic data point in section 3.4.

(61) [Context: In this department, every professor assigns the same grade to all of his students.]

a. #Every professor who assigned an A to some of his students got a pay raise.

b. ✓Every professor who assigned an A to all of his students got a pay raise.

(61a) is ruled out by both LI and LI* because the restrictor 'professor who assigned an A to some of his students' contextually but not logically entails its 'all'-alternative. As pointed out before, (61b) is predicted to be unacceptable by LI because the sentence as a whole is logically weaker than its 'some'-alternative but contextually entails it. LI*, on the other hand, correctly predicts (61b) to be acceptable; the reason is that the relevant alternative, namely (61a), already contains a CC violation in its restrictor as pointed out just now. It is therefore not a "viable" alternative (it does not pass through the filter of CC*) and since there are no other alternatives (by assumption) that (61b) competes with, (61b) is predicted to be acceptable.

Having worked through one example, let me note in passing that one way to conceptualize CC* in (60) is to say that the more deeply embedded violations of CC are more severely punished than the less embedded ones. Given the choice between (61a) and (61b) in the context of (61), (61a) is "less optimal" because it contains a CC violation in its restrictor while the smallest constituent of (61b) that violates CC is the sentence itself.

Second, let us go back to the problem raised in section 3.4.2.

(62) [Context: two students solved all of the math problems and the rest solved none.]

a. #Both students who solved some of the math problems passed.
$\rightsquigarrow$ Exactly two students solved some of the math problems (presupposition)

b. ✓Both students who solved all of the math problems passed.
$\rightsquigarrow$ Exactly two students solved all of the math problems (presupposition)

As a reminder, the problem was this: the sentences in (62) encode the same assertive content but carry logically independent presuppositions. Consequently, neither of the two alternatives logically entails the other one. But since in the context specified in (62) both presuppositions are satisfied, both sentences contextually entail each other in that context. Therefore, LI predicts both sentences to be unacceptable (they "cancel each other out"). LI*, on the other hand, correctly breaks the tie in favor of (62b). The reason is similar to above: not only does (62a) violate CC at root, its restrictor 'students who solved some of the math problems' also violates CC because it entails its 'all'-alternative contextually but not logically. (62b) is, therefore, predicted to be acceptable precisely because the only alternative to the sentence as a whole is (62a) which already contains a CC violation.

Third, and finally, let us go back to a problem pointed out in fn. 6.

(63) [Context: all students smoke, John does not know about the smoking situation of the students but (we know that) he firmly believes that it is not the case that some but not all of the students smoke, he thinks that either all students smoke or none do.]

a. #John is unaware that some students smoke.

b. ✓John is unaware that all students smoke.

LI predicts (63b) to be unacceptable because, in the context of (63), (63b) contextually entails (63a) even though it does not do so logically. LI* correctly predicts (63b) to be acceptable for the same reason as above: the only relevant alternative is (63a) which already contains a CC violation in the embedded clause and, therefore, does not pass through the CC* filter in (60).

6 Loose ends and open problems

In this section I will reflect on those cases that I am aware of which are problematic for the analysis LI(*).[42] These are ordered from more to less severe.

6.1 Presupposed ignorance, revisited

In section 3 it was shown that LI can match the prediction of *Presupposed Ignorance* for the following case.

(64) [Context: Mary leaves in Paris.]

a. #John is unaware that Mary leaves in Paris or London.

b. ✓John is unaware that Mary leaves in Paris.

There is one particularly recalcitrant construction, quite similar to (64), which can easily be accounted for by PI but not LI(*); indeed, this data point was put forth by Spector & Sudo (2017) as evidence in favor of PI.

42 From here on I will use 'LI(*)' to signify that the difference between LI and LI* does not matter for the case at hand.

(65) Mary lives in Paris...

  a. #John$_F$ lives in Paris or London, too.

  b. ✓John$_F$ lives in Paris, too.

To see why LI(*) cannot capture the oddness of (65a), note once the anaphoric element of 'too' in (65a) and (65b) is resolved to Mary, their meanings can be represented as in (67) where underlining marks for presuppositionality.

(66) a. $\underline{\text{Mary lives in Paris or London}} \wedge$ John lives in Paris or London

  b. $\underline{\text{Mary lives in Paris}} \wedge$ John lives in Paris

Now, (66a) clearly does not logically entail (66b); LI(*) therefore predicts that it can be used only if the context allows for the possibility of (66a) to be true and (66b) to be false or undefined. This boils down to the requirement that there must be world compatible with the context in (65) in which Mary lives in Paris and John lives in London. The problem is that this possibility is by no means ruled out in the context of (65), which is why (65a) is predicted to be acceptable, contrary to fact. Note that we can also look at the expression 'John lives in Paris or London' which is embedded under 'too'. Since the local context of this expression is the same as the global context, the requirement triggered by LI(*) is that it must be contextually possible that John lives in London and not Paris, which, again, is not problematic. Furthermore, the oddness of (65a) cannot be traced to redundancy: in the relevant contexts, neither 'John lives in Paris or London, too' nor 'John lives in Paris or London' are contextually equivalent with the alternatives 'John lives in Paris, too' and 'John lives in Paris'.

I conclude that (65) is a genuine puzzle for LI(*), and indeed, of all the analyses that were reviewed in this article, only PI can derive the oddness of (65): since (65b) is an alternative to (65a), the latter cannot be used if the presupposition of the former is satisfied which is what is the case in (65).

6.2 Homogeneity

As a reader who is familiar with the literature covered in this paper must have noticed, I have so far completely avoided any discussion of homogeneity. To see the relevance of homogeneity, consider (67).

(67) [Context: talking about four brothers and sisters. Obviously they share their last name.]

  a. #Some of these kids have a beautiful last name.

  b. #All of these kids has a beautiful last name.

What is surprising here, from our perspective, is the oddness of (67b). Magri (2009a) proposes that (67b) is unacceptable because it competes with (68).

(68) These kids have a beautiful last name.

The key here is that the plural definite 'these kids' in (68) triggers the homogeneity inference that either each of the kids have a beautiful last name or that none of them does. Magri assumes without argument that homogeneity inferences are presuppositional. Given this assumption, it is then natural to conclude that the oddness of (67b) is a *Maximize Presupposition!*-type effect: since it is common knowledge that children inherit the last name of their father, it is common knowledge that siblings have the same last name, which in turn yields the homogeneity "presupposition" associated with (68). From the point of view of L(*) the problem here is that the homogeneity inference of (68) is entailed by the assertive meaning of (67b). In fact, (67b) logically entails (68). It immediately follows that the oddness of (67b) cannot be attributed to (68) if LI(*) is the correct hypothesis.

Having pointed out that (67b) is a genuine problem for LI(*) let me now spell out some considerations that suggest this is not necessarily a bad thing. First, as has been convincingly argued in Križ (2016), Križ & Spector (2017), homogeneity is *not* a presupposition. For example, unlike regular presuppositions, (69a), the homogeneity inference triggered by the plural definite in the scope of 'not all' does not project universally to root. An account based on *Maximize Presupposition!* therefore is a *non sequitur*.

(69) a. Not all students have stopped smoking.
$\rightsquigarrow$ Every student used to smoke

b. Not all students read the books.
$\rightsquigarrow$ There is at least one student who read none of the books
$\not\rightsquigarrow$ At least one student read all of the books and the rest read none

Second, homogeneity violations appear to be not as robust as the other cases of oddness that were discussed in previous section, in the sense that structurally similar sentences are sometimes (in fact quite often) acceptable. The most obvious example is the one that we are very familiar with:

(70) [Context: John always gives the same grade to all his students.]
✓This semester, he gave all of his students an A.

It is clear that the homogeneity inference associated with 'John gave his students an A', that either John gave all his students an A or he gave none of them an A, is supported in the context of (70); but the 'all'-sentence is impeccable. Intuitively, it seems like homogeneity violations lead to oddness only when the relevant pieces of background assumption are (at least partially) deeply entrenched in common knowledge. This is stark contrast to other constructions discussed in this paper. For these other examples oddness ensues regardless of whether the relevant piece of background assumption is conveyed by previous discourse or is part of some deeply rooted system of common knowledge:

(71) a. #John broke all his arms.

b. #If John has only two fingers, then he broke all of them.

Third, I'd like to make the following observation: quite often, at least in English, 'all'-type sentences are structurally more complex than their definite alternatives. This is already the case with (69b) and (70). Furthermore, the definite alternatives are also "semantically simpler" in the sense that a DP like 'all students' has the semantic type of a generalized quantifier while a definite like 'the students' arguably has a lower type. Under the assumption that there is a general pressure to use syntactically/semantically "simpler" alternatives whenever possible, one might postulate the following principle.

(72) Whenever two sentences ϕ and ψ are contextually identical *relative to common knowledge* and ψ is structurally/semantically “simpler”, then one must use ψ.

In a nutshell, I’d like to propose that homogeneity violations are the result of competition of alternatives with various degrees of syntactic/semantic complexities relative to common knowledge. This does not make any commitments regarding the precise analysis of homogeneity, which is an advantage. It could be a theorem of a general theory of oddness based on structural redundancy, or it could be a *sui generis* principle that regulates the use of homogeneity triggering expressions (as far as anyone knows homogeneity itself is a *sui generis* phenomena, it does not seem far fetched that it might be guided by *sui generis* principles (as disappointing as this might be theoretically)). See Singh (2009) for some arguments to the same effect.

7 Conclusion

A generalization is proposed that yields a unified analysis of three types of acceptability patterns; going by the title of the associated theories, these where *Maximize Presupposition!* cases, *Presupposed Ignorance* cases, and *Mismatching Implicature* cases. Furthermore, it is argued that the predictions made by this generalization are either effectively identical with the competing piece-meal theories, or are superior. To my knowledge this is the first generalization that has been able to cast such a wide net. Nevertheless some problems remain, pointed out in the previous section, which require further research.

References

Anvari, Amir. 2015. Scalar implicatures and anti-presuppositions: a unified approach. Term Paper (supervisor: B. Spector).

Chemla, Emmanuel, Paul Égré & Benjamin Spector. 2017. Characterizing logical consequence in many-valued logic. Journal of Logic and Computation 27(7). 2193–2226.

Dieuleveut, Anouk, Emmanuel Chemla & Benjamin Spector. 2016. Weak and strong quantity implicatures–an experimental investigation. Ms. LSCP & IJN.

von Fintel, Kai. 1999. Npi licensing, strawson entailment, and context dependency. Journal of Semantics 16(2). 97–148.

Fox, Danny. 2007. Free choice and the theory of scalar implicatures. In Presupposition and implicature in compositional semantics, 71–120. Palgrave Macmillan.

Gajewski, Jon. 2004. On analycity in natural language. Manuscript, MIT.

Gajewski, Jon & Yael Sharvit. 2012. In defense of the grammatical approach to local implicatures. Natural Language Semantics 20(1). 31–57.

Heim, Irene. 1983. On the projection problem for presuppositions. In Wccfl 2, 114–125.

Heim, Irene. 1991. Semantics: an international handbook of contemporary research chap. Artikel und definiheit. de Gruyer, Berlin.

Katzir, Roni & Raj Singh. 2015. Economy of structure and information: oddness, questions, and answers. In Eva Csipak & Hedde Zeijlstra (eds.), Proceedings of sinn und bedeutung 19, .

Križ, Manuel. 2016. Homogeneity, non-maximality, and all. Journal of Semantics 33(3).

Križ, Manuel & Benjamin Spector. 2017. Interpreting plural predication: homogeneity and non-maximality. Manuscript, IJN.

Magri, Giorgio. 2009a. A theory of individual-level predicates based on blind mandatory scalar implicatures. Natural Language Semantics 17(3). 245–297. http://dx.doi.org/10.1007/s11050-009-9042-x.

Magri, Giorgio. 2009b. A theory of individual-level predicates based on blind mandatory scalar implicatures: MIT dissertation.

Magri, Giorgio. 2011. Another argument for embedded scalar implicatures based on oddness in downward entailing contexts. In N. Li & D. Lutz (eds.), Salt 20: Proceedings of the 20th annual conference semantics and linguistics theory, 564–581.

Marty, Paul P. 2017. Implicatures in the dp domain: MIT dissertation.

Meyer, Marie Christine. 2013. Ignorance and grammar. Cambridge, MA: MIT dissertation.

Percus, Orin. 2006. Antipresuppositions. In Ayumi Uyema (ed.), Theoretical and empirical studies of reference and anaphora: Toward the establishment of generative grammar as an empirical science, 52–73. Japan Society for the Promotion of Science.

Percus, Orin. 2010. Antipresuppositions revisited. Handout of a talk given at Universite de Caen.

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

Sauerland, Uli. 2008. Implicated presuppositions. In A. Steube (ed.), Sentence and context: Language, context, and cognition, .

Schlenker, Philippe. 2009. Local contexts. Semantics and Pragmatics .

Schlenker, Philippe. 2012. Maximize presupposition and gricean reasoning. Natural Language Semantics 20(4). 391–429.

Schlenker, Philippe. to appear. Gesture projection and cosuppositions. Linguistics & Philosophy .

Sharvit, Yael & Jon Gajewski. 2008. On the calculation of local implicatures. In Wccfl, vol. 26, 411–419.

Singh, Raj. 2009. Maximize presupposition! and informationally encapsulated implicatures. In Proceedings of sinn und bedeutung, vol. 13, .

Singh, Raj. 2011. Maximize presupposition and local contexts. Natural Language Semantics 19(2). 149–168.

Spector, Benjamin. 2014. Scalar implicatures, blindness, and common knowledge - comments on magri (2011). In Salvatore Pistoia Reda (ed.), Pragmatics, semantics and the case of scalar implicatures, 146–169. Palgrave Macmillan.

Spector, Benjamin & Yasu Sudo. 2017. Presupposed ignorance and exhaustivity: how scalar implicatures and presuppositions interact. Linguistics and Philosophy .