

THÈSE DE DOCTORAT
DE L'UNIVERSITÉ PSL
Préparée à Ecole Normale Supérieure

Meaning in Context

Soutenue par

Amir ANVARI

Le 06/06/1987

Ecole doctorale n° 158

ED3C - 158

Spécialité

Sciences cognitives

Composition du jury :

Emmanuel, CHEMLA
M., Ecole Normale Supérieure *Examineur*

Amy Rose, DEAL
Mme., University of California, Berkeley *Rapporteur*

Cornelia, EBERT
Mme., Goethe-Universität Frankfurt *Rapporteur*

Danny, FOX
M., Massachusetts Institute of Technology *Examineur*

Philippe, SCHLENKER
M., Ecole Normale Supérieure *Directeur de thèse*

Benjamin, SPECTOR
M., Ecole Normale Supérieure *Co-Directeur de thèse*

RÉSUMÉ

Cette thèse contient une série d'enquêtes sur trois manières de comprendre la notion de contexte en tant que facteur interagissant avec le système linguistique afin de permettre, faciliter et, de manière générale, enrichir la communication. Les notions pertinentes sont les suivantes: contexte, contexte, situation, énonciation et signaux contextuels.

MOTS CLÉS

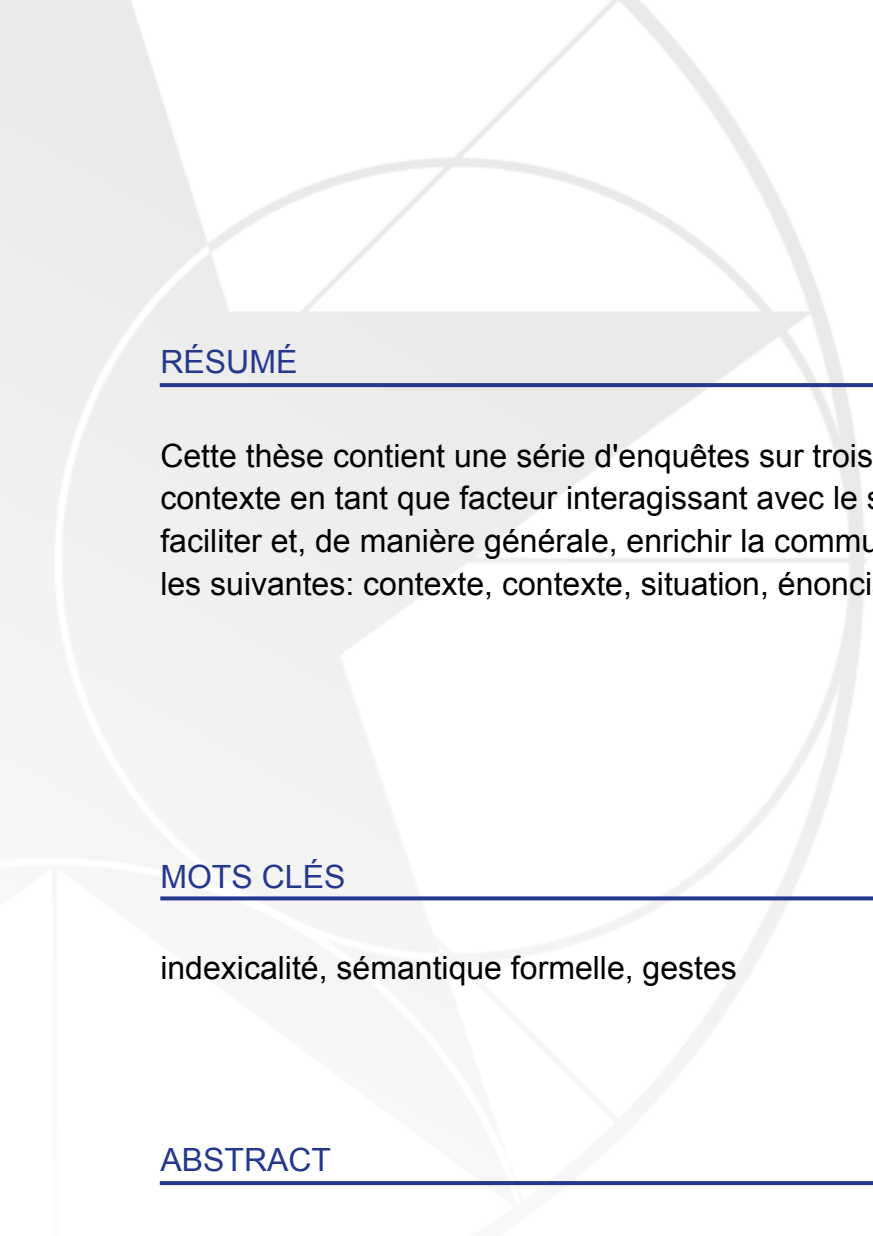
indexicalité, sémantique formelle, gestes

ABSTRACT

This dissertation contains a series of investigations into three ways in which the notion of context can be understood as a factor interacting with the linguistic system in order to enable, facilitate, and generally enrich communication. The relevant notions are context-qua-common ground, context-qua-situation of utterance, and context-qua-paralinguistic signals.

KEYWORDS

indexicality, formal semantics, gestures



Contents

1	Introduction	3
I	Part One: Context <i>qua</i> Common Ground	6
2	Logical Integrity	7
2.1	Introduction	7
2.2	Logical Integrity (first version)	10
2.3	<i>Maximize Presupposition!</i> and related phenomena	16
2.4	Mandatory implicatures	26
2.5	Logical Integrity (final version)	34
2.6	Loose ends and open problems	37
2.7	Conclusion	39
	Appendices	40
.1	Percus cases and <i>Presupposed Ignorance</i>	41
.2	On why <i>Presupposed Ignorance</i> and Magri's system cannot be innocently combined	51
.3	Some provisional remarks about <i>too</i>	56
3	A Problem for <i>Maximize Presupposition!</i> (locally)	60
4	The Problem of Reflexive Belief in Spanish (with Mora Maldonado and Andrés Soria)	62
4.1	Introduction	62
4.2	Why not <i>Maximize Presupposition!</i>	64
4.3	The Contrafactivity Hypothesis	70
4.4	The Enriched Excluded Middle Hypothesis	74
4.5	Some Further Empirical Observations	75
4.6	Conclusions	79
II	Part Two: Context <i>qua</i> Situation of Utterance	80
5	The Ban Against Illeism and Indexical Shift in Farsi	81
5.1	Introduction	81
5.2	Farsi and indexical shift	88
5.3	The operator-based account of indexical shift	92
5.4	Indexical shift bleeds the Ban Against Illeism	94

5.5	Indexical shift and question-embedding: a case study	96
5.6	Indexical shift feeds the Ban Against Illeism	98
5.7	Comparing the operator-based theory to the binding-based theory . .	103
5.8	Generalizing across sortal domains	107
5.9	Conclusion	109
Appendices		111
.1	Exceptions to the Ban Against Illeism	112
6	Copy Theory of Movement and Indexical Shift in Farsi	114
6.1	Introduction	114
6.2	Farsi and indexical shift	120
6.3	Two problems with copy theory and the operator-based approach . . .	124
6.4	The operator-based approach, deletion and indexical pronouns as re- stricted variables	127
6.5	Copy theory and the variable-based account of indexical shifting . . .	134
6.6	Conclusion	140
III Part Three: Context qua Paralinguistic Signals		141
7	On Co-nominal Pointing	142
7.1	Introduction	142
7.2	The exemplification hypothesis	147
7.3	The problem with <i>every</i> and <i>no</i>	151
7.4	A projection recipe for the exemplification hypothesis	152
7.5	Another look at <i>every</i> and <i>no</i>	154
7.6	Co-nominal pointing gestures and co-predicative iconic gestures . . .	155
7.7	Conclusion	159
8	Dislocated Cosuppositions	162
8.1	Introduction	162
8.2	The cosuppositional analysis	163
8.3	The <i>supervaluationist</i> analysis	165
8.4	Dislocated Cosuppositions	167
8.5	Conclusion	170

Chapter 1

Introduction

This dissertation contains a series of investigations into three ways in which the notion of *context* can be understood as a factor interacting with the linguistic system in order to enable, facilitate, and generally enrich communication. Correspondingly, the dissertation is organized in three parts. In this introduction, I will summarize and outline the rest of this dissertation. The relevant notions are context *qua* common ground (Part I), context *qua* situation of utterance (Part II), and context *qua* paralinguistic signals (Part III).

Thinking of *context* as the set of background assumptions (or, common ground, Stalnaker 1978, Lewis 1979b), in Chapter 2, I argue for a principle called Logical Integrity, which uniformly accounts for three classes of examples that have received distinct theoretical treatments in the literature; namely, *Maximize Presupposition!* (Heim 1991a and subsequent literature), *Presupposed Ignorance* (Spector & Sudo 2017b), and mismatching implicatures (Magri 2009b and subsequent literature). According to Logical Integrity a sentence is infelicitous if it has a logically non-weaker alternative that it contextually entails.¹

In a short squib in Chapter 3, an empirical problem is pointed out which, as far as I know, is problematic for every extant theory of the relevant paradigm.²

Finally, in another paper (co-authored with Mora Maldonado and Andrés Soria) the case of ‘reflexive belief’ in Spanish is investigated in detail. This construction, built by adjoining the reflexive pronoun *se* to the predicate *creer* (to believe), displays a puzzling presuppositional behavior. When unembedded, *creerse* triggers the inference that its complement clause is false. When embedded under negation, it triggers the inference that its complement clause is true. We first argue that the negative inference in the unembedded case is not due to the *Maximize Presupposition!* principle (or any other related mechanism, including Logical Integrity), carving out the empirical profile of *creerse* along the way. We then explore two alternative explanations. The first is that *creerse* presupposes the falsity of its complement, i.e., that *creerse* is anti-factive. The second is that *creerse* presupposes that the attitude holder is *wrongly* opinionated with respect to the embedded proposition. We argue that the first analysis fails unless it is supplemented with the syntactic account of neg-raising, whereas the second analysis faces empirical challenges.³

¹A shorter version of this paper has appeared in the proceedings of the 28th *Semantics and Linguistic Theory* conference. This, longer version has been accepted with major revisions by the journal, *Natural Language Semantics*.

²This squib has appeared in the 33rd issue of *Snippets*.

³This paper has appeared in the proceedings of the 23rd *Sinn und Bedeutung* Conference. I am

Thinking of *context* as the situation of utterance, encoding information about, e.g., the speaker, the time and location of utterance (Kaplan 1977, Lewis 1980), it is an interesting finding that in some languages there are constructions in which more than one context of utterance is relevant for interpretation, a phenomenon known as “indexical shift” (see Deal 2017 for a recent survey). In a language that allows indexical shift, direct discourse can be reported in indirect discourse using the same indexical expressions that were used in the original utterance. For example, in English, which does not allow indexical shift, if John asserts, “I am hungry”, his utterance must be reported as *John said he was hungry*, where the indexical first person pronoun is replaced with a third person pronoun referring to John (I abstract away from tense). In Farsi, which does allow indexical shift, on the other hand, John’s assertion can be reported using either a sentence that corresponds to *John said he is hungry* (like English) or a sentence that correspond to *John said I am hungry* (unlike English). Stated differently, the first person pronoun in Farsi can be interpreted either with respect to the actual situation of utterance or with respect to the *reported* situation of utterance. Part II is dedicated to this phenomenon. In a sequence of two papers, I build on data from Farsi to explore how indexical shift interacts with illeism (i.e., referring to the speaker or the addressee with a third person noun phrase), on the one hand, and movement, on the other.

Illeism, the act of referring to oneself in the third person, is generally infelicitous, a generalisation which I label the Ban Against Illeism (BAI). In Chapter 5, I provide evidence that, in Farsi, indexical shift both ‘bleeds’ and ‘feeds’ BAI. The former is the case because in an environment in which indexical shift is obligatory, third-person reference to the *actual* speaker is *acceptable*. Stated differently, one can construct cases of acceptable illeism by embedding speaker-denoting third person noun phrases in context-shifted environments. The latter is the case because in an environment in which indexical shift is obligatory, third-person reference to the *reported* speaker is *not* acceptable. Stated differently, third person noun phrases in Farsi become ‘anti-logophoric’ in context-shifted environments. In the paper, I point out that the former generalization can be easily captured on the so-called operator-based account of indexical shift (Anand & Nevins 2004; Anand 2006) coupled with the assumption that there is a competition-based principle of language-use according to which third person noun phrases compete with indexical (i.e., first or second person) pronouns and the latter are preferred if they convey the same meaning, which happens in particular when third person noun phrases refer to the actual speaker / addressee. The latter generalization, however, is not captured in this framework. This is because shifty indexicals and third person noun phrases utilize different modes of reference: shifty indexicals refer *de se* while third person noun phrases (at least can) refer *de re*. The problem is that *de se* and *de re* attitudes, even if they are ultimately ‘about’ the same individual, give raise to *distinct* truth conditions, both semantically and typically contextually as well, with the result that the competition-based principle, which, as a principle of language use, compares the meanings of alternatives at root, fails to apply to these cases. Therefore, I propose that the competition-based principle must be modified so as to compare noun phrases for covaluation, instead of sentences for equivalence. Once this modification is made, Sharvit’s (2010) independently motivated notion of ‘type-II covaluation’ can be used to derive the desired facts.

grateful to M. Maldonado and A. Soria for allowing me to reproduce with paper as a part of my dissertation. The authors made equal contribution to this paper.

In Chapter 6, I provide two pieces of evidence for the claim that the operator-based approach to indexical shift has over-generation problems when it is combined with the copy theory of movement (Chomsky 1993b and subsequent literature). I proceed to develop two techniques to solve the over-generation problem. One technique maintains the core tenets of operator-based account but introduces auxiliary assumptions to solve the problem. The introduced assumptions pertain to the representation of shift indexicals, on the one hand, and a fine-grained deletion operator at the LF interface which can target individual ϕ -features, on the other. The other technique proceeds from an entirely different theoretical premise. This time the binding-based approach to indexical shift (Schlenker 1999 and subsequent work) takes the center stage. I argue that the over-generation problems can be solved within this framework by a modification of an independently needed stipulation, i.e., the SHIFT TOGETHER constraint on indexical shift. According to this constraint, two indexicals that occur in the same ‘minimal domain’ must be evaluated with respect to the same context. The proposed modification is applied to the notion of ‘minimal domain’; the idea, in a nutshell, is that if an indexical occurs in a chain, its minimal domain is determined by the occurrence of the indexical at the *head* of the chain. In other words, one side-effect of movement, on this view, is to *extend* the minimal domain of indexicals that occur in the moved constituents.

Finally, thinking of *context* as paralinguistic signals, i.e., gestures that accompany or replace the spoken message, in Chapter 7, I zoom in on the interpretation of ‘co-nominal pointing’, i.e., those pointing gestures that accompany quantified noun phrases. I argue that co-nominal pointing ‘exemplifies’ plural discourse referents that are made salient by quantifiers for subsequent anaphoric pick-up (Nouwen 2003 a.o.). On this line of thinking, a direct parallelism is predicted between the interpretation of co-nominal pointing and the anaphoric potential of various determiners. In the paper, I provide evidence that this prediction is born out. In Chapter 8, I discuss the challenge that non-monotonic environments raise for the ‘cosuppositional’ account of iconic gestural enrichments, advanced in Schlenker 2018. The goal of this paper is to maintain the core of the cosuppositional analysis by introducing modifications that help improve on the predictions of the original analysis.⁴

⁴This paper has appeared in the proceedings of *the 21st Amsterdam Colloquium*.

Part I

Part One: Context *qua* Common Ground

Chapter 2

Logical Integrity

2.1 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 (1991a), the *Presupposed Ignorance* system in Spector & Sudo (2017a) and the *Mismatching Implicature* approach of Magri (2009b). 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 (1991a) 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 its presupposition is not known to be true. This principle was taken up in subsequent literature (Percus 2006a; Sauerland 2008b; Schlenker 2012b, 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.]

¹See the discussion in sections 2.2.2 and 2.5.

- 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.

- John has exactly two arms
- (3) John broke all of his arms
 \therefore John broke both of his arms

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 2.3.4, 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 restriction amount to in practice? Several arguments in the literature (see in particular Percus 2006a and Schlenker 2012b) 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 (2017a) (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.

- All students smoke
- (5) John is unaware that some students smoke
 ∴ 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 (38), 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 consequential mechanism of exhaustification is motivated in Magri (2009b,c, 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 2.4, (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 2.5 a modified version of the initial version of the generalization (as put forth in section 2.2) is motivated on the basis of a set of problematic data points discussed in the earlier sections. Section 2.6 is concerned with several loose ends and open problems for the present account.

2.2 Logical Integrity (first version)

2.2.1 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 $\mathcal{Y}_1, \dots, \mathcal{Y}_n$, if $\mathcal{X}_1(\mathcal{Y}_1) \dots (\mathcal{Y}_n) = 1$ then $\mathcal{X}_2(\mathcal{Y}_1) \dots (\mathcal{Y}_n) = 1$.

Another, strictly weaker notion of entailment, which we will have occasion to use in its bidirectional form in section 2.3, is that of Strawson-entailment (von Stechow 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 $\mathcal{Y}_1, \dots, \mathcal{Y}_n$, if $\mathcal{X}_1(\mathcal{Y}_1) \dots (\mathcal{Y}_n) = 1$ then $\mathcal{X}_2(\mathcal{Y}_1) \dots (\mathcal{Y}_n) \in \{\#, 1\}$.

This study relies on Schlenker’s (2009a) 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, its *local* context, denoted $lc(\alpha, \phi C)$. 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 v , and generalized quantifier \mathbb{Q} ,
- $lc(\psi, [\phi \wedge \psi], C) = lc(\psi, [\phi \rightarrow \psi], C) = \lambda w. w \in C \wedge \llbracket \phi \rrbracket^w = 1$
 - $lc(\psi, [\phi \vee \psi], C) = \lambda w. w \in C \wedge \llbracket \phi \rrbracket^w = 0$
 - $lc(\beta, [\mathbb{Q}(\alpha, \beta)], C) = \lambda w. \lambda x. w \in C \wedge \llbracket \alpha \rrbracket^w(x) = 1$
 - $lc(\phi, [v \text{ is (un)aware that } \phi], C) = \lambda w. \exists w' \in C : w = w' \vee w \in \text{DOX}_v^{w'}$

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 $\mathcal{Y}_1, \dots, \mathcal{Y}_n$, if $\mathcal{X}_1(\mathcal{Y}_1) \dots (\mathcal{Y}_n) = 1$ and $\mathcal{C}(\mathcal{Y}_1) \dots (\mathcal{Y}_n) = 1$ then $\mathcal{X}_2(\mathcal{Y}_1) \dots (\mathcal{Y}_n) = 1$.

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 , $\llbracket \phi \rrbracket \wedge C$ 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 worlds that are compatible with what v 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 v ’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.

2.2.2 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

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

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

examples, a subset of which was discussed in section 2.1, 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 2.5.

As pointed out in section 2.1, 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 $\exists w \in W : \llbracket \phi \rrbracket^w = 1 \wedge \llbracket \psi \rrbracket^w \in \{0, \#\}$. Then, ϕ is acceptable in context C only if $\exists w \in C : \llbracket \phi \rrbracket^w = 1 \wedge \llbracket \psi \rrbracket^w \in \{0, \#\}$.

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., $\phi \not\prec \psi$) 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., $\phi \not\prec_C \psi$) 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 (2017a) which, as pointed out in section 2.1, 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. $\llbracket \text{unaware} \rrbracket^w = \lambda P \lambda x : P(w) = 1. \neg B_x^w[P]$
 c. For any world w , individual x and proposition P , $B_x^w[P]$ 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., $\neg B_J[\exists]$). 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:

- (16) [Context: all students smoke.]

⁴See section 2.6.2 for a discussion of an immediate consequence of this prediction pertaining to homogeneity.

⁵‘ α is unaware that ϕ ’ can be analyzed in as $\underline{\phi} \wedge \neg B_\alpha(\phi)$, 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.

- ✓ 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) fares 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 ‘unaware. . . 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 all of them.
 b. #If all students smoke, John is unaware that some students smoke.
 c. #Either John gave his students different grades, or he gave some 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 presupposition that every

⁶The 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 2.5.

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 $S = \{\langle w, a \rangle : w \in C \wedge \llbracket \text{professor who has exactly two students} \rrbracket^w(a) = 1\}$. 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 $\langle w, a \rangle \in S$, $\llbracket \lambda x.x \text{ invited all of } x\text{'s students} \rrbracket^w(a) = \llbracket \lambda x.x \text{ invited both of } x\text{'s students} \rrbracket^w(a)$. 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 (2017a) for *Presupposed Ignorance*, and Schlenker (2012b) for *Mandatory Implicatures* (for the latter, see also section 2.3.2).⁷

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

- 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., $\beta \not\# \beta'$ but $\beta \vDash_C \beta'$).

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

⁷Although 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 $\lambda w. \exists w' \in C : w = w' \vee w \in \text{DOX}_J^{w'}$. 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*.

⁸In 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

sentence ruled out by (11) is ruled 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 2.5.

2.3 Maximize Presupposition! and related phenomena

2.3.1 Outline

In subsection 2.3.2 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 2.3.3 and it is argued, building in particular on Spector & Sudo (2017a) (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 2.3.4 closes by a discussion of two data points that are problematic for S&S’s proposal.

2.3.2 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, $\forall w \in C : \llbracket \phi \rrbracket^w = 1 \Leftrightarrow \llbracket \psi \rrbracket^w = 1$.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, $\forall w \in C : \llbracket \phi \rrbracket^w = \llbracket \psi \rrbracket^w$. 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. ALT = {The sun is shining}
 b. #All of John’s eyes are open. 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.

local contexts.

⁹This is the bidirectional version of contextual entailment as defined in section 2.2.1.

Furthermore, the two alternatives are equally informative in normal contexts in the sense that they cannot convey 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, $\forall w \in W : \{\llbracket \phi \rrbracket^w, \llbracket \psi \rrbracket^w\} \subseteq \{0, 1\} \Rightarrow \llbracket \phi \rrbracket^w = \llbracket \psi \rrbracket^w$.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, \dots, x_n of appropriate types, if $\llbracket \beta \rrbracket(x_1) \dots (x_n) \neq \#$ and $\llbracket \beta' \rrbracket(x_1) \dots (x_n) \neq \#$ then $\llbracket \beta \rrbracket(x_1) \dots (x_n) = \llbracket \beta' \rrbracket(x_1) \dots (x_n)$.

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

¹⁰This is simply the bidirectional counterpart of Strawson-entailment as defined in section 2.2.1.

¹¹One 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.

2.3.3 Non-Strawson-equivalent alternatives: Spector & Sudo (2017a) 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 2.3.2) 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 2.2.2, LI predicts the relevant facts without any analogous problems.

- (25) a. If MP is restricted by equal-informativeness, why is (23a) unacceptable?
 b. If MP is not 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 (2012b) 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 {✓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 (2017a), 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).

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

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 .

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 ‘ $exh^* \phi$ ’ is defined to be undefined if there is an innocently excludable alternative which is undefined (or if the prejacent itself is undefined).

- (28) Suppose $ALT_{IE}(\phi) = \{\psi\}$. Then,

$$\llbracket exh^*_{ALT_{IE}(\phi)} \phi \rrbracket^w = \begin{cases} \# & \text{if } \llbracket \phi \rrbracket^w = \# \text{ or } \llbracket \psi \rrbracket^w = \# \\ 1 & \text{if } \llbracket \phi \rrbracket^w = 1 \text{ and } \llbracket \psi \rrbracket^w = 0 \\ 0 & \text{if } \llbracket \phi \rrbracket^w = 0 \text{ or } \llbracket \psi \rrbracket^w = 1 \end{cases}$$

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.
 $ALT = \{\text{John is unaware that all students smoke}\}$

In S&S’s framework, the sentence $\phi = \text{‘John is unaware that some students smoke’}$ 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), ‘ $exh^* \phi$ ’ 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.]
 \checkmark John is aware that some students smoke.
 $ALT = \{\text{John is aware that all students smoke}\}$

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

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 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

2.3.4 Two challenges for S&S’s proposal

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 than 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).

¹⁴S&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 2.2.2. 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.

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

- (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.

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 cannot be felicitously used in C.

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.

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 (1991a). (See Appendix .1 contains a detailed discussion.)

¹⁶Note that these data points are also problematic for standard *Maximize Presupposition!* as formulated in section 2.3.2. 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.

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

- (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) would have bocked (35a) as it has a stronger presupposition and it can be felicitously 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

¹⁸This 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).

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

{[?]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 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.
 \rightsquigarrow *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

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 2.5 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.
 \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)

Three features of the two sentences in (39) are crucial. First, note that the presup-

²⁰One 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.

²¹Here 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.

positions 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). 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 (they encode the same assertive content but) the presupposition of ψ is not weaker 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).

²²The 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 (2009b) account generalized to presuppositions. Magri’s account will be discussed in detail in section 2.4. The point of (39) is the problem it poses for MP and PI.

2.4 Mandatory implicatures

2.4.1 Basic cases and Magri’s (2009b) 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 2.2.2. 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 but not both 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 (2009b,c) 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)
- There is a covert exhaustivity operator, *exh*, that is *obligatorily* attached to every possible scope site.²³
 - exh* operates solely on the basis of *logical* (as opposed to contextual) entailment.
 - To any expression ϕ , a set of alternatives is associated, $ALT(\phi)$.²⁴ In general, the “domain” of *exh* is restricted to a subset of $ALT(\phi)$, i.e. the “innocently excludable” alternatives (Fox 2007), $ALT_{IE}(\phi)$. Intuitively, this is the collection of those alternatives that, in conjunction with the “prejacent” ϕ , can be consistently negated.
 - Let ϕ be some expression. *exh* ϕ always entails that ϕ is true. For any $\psi \in ALT_{IE}(\phi)$, *exh* ϕ entails that ψ is false if and only if ψ is relevant.

²³Any 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, $\bullet[\phi] \rightarrow \bullet[exh[\phi]] \rightarrow \bullet[exh[exh[\phi]]] \rightarrow \dots$. 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.

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

- 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. $\llbracket \text{exh } \phi \rrbracket^w = 1$ iff $\llbracket \phi \rrbracket^w = 1 \wedge \forall \psi \in \text{ALT}_{\text{nw}} : (\llbracket \psi \rrbracket^w = 0 \vee \neg \text{Rel}(\psi))$
 b. $\psi \in \text{ALT}_{\text{nw}}(\phi) \Leftrightarrow \psi \in \text{ALT}(\phi) \wedge \phi \not\# \psi$ (for ‘#’, see (7) from section 2.2.1)

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. $\llbracket (44a) \rrbracket^w = 1$ iff $\exists \wedge \neg \forall$. ($\not\#$ 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 (2012b). 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 (2012b).

- (45) No teacher who will assign the same grade to each of his students wants to give $\{\# \text{some}, \checkmark \text{all}\}$ 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

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

²⁶While 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.

²⁷Schlenker’s (2012b) 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

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 contextually equivalent with its alternative ‘ $\lambda x. x$ 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 ‘ $\lambda x. x$ has decided to give some of x ’s students an A’ is equivalent with its alternative ‘ $\lambda x. x$ has decided to give all of x ’s students an A’.

I therefore conclude with Schlenker (2012b) 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 1983b* or *à la Schlenker 2009a*), 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.

2.4.2 Problems with *Mandatory Implicatures*

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).

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.

²⁸For 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 committed to this assumption.

²⁹In 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.

³⁰Note 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.

- (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.

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 $\neg B\forall$ which is equivalent to $(B\neg\forall) \vee (\neg B\neg\forall \wedge \neg B\forall)$.

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.

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).

³¹Following 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”.

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

³³Example (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.

³⁴It 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.

- (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 $[\dots K\phi]$ is ruled out unless for every $\psi \in \text{ALT}(\phi)$, $[\dots K\phi]$ *logically* entails either $K\psi$ or $\neg K\psi$.

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. 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. $K\exists$
 b. $Kexh\exists \equiv K[\exists \wedge \neg\forall]$
 c. $exhK\exists \equiv K\exists \wedge \neg K\forall$
 d. $exhKexh\exists \equiv K[\exists \wedge \neg\forall]$
 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 $K\forall$ or $\neg K\forall$, 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 $K[\exists \wedge \neg\forall]$ or $K\exists \wedge \neg K\forall$. 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 $\neg K\forall$. 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*

³⁵Thus, for any sentence ϕ , $\llbracket K\phi \rrbracket^w = 1$ iff $\forall w' \in \text{DOX}_S^w : \llbracket \phi \rrbracket^{w'} = 1$, where S is the speaker.

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

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 $K\forall$, leading to the incorrect prediction that (51) is felicitous in the context (51e).³⁸

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.

$K \text{ no}(\alpha, \exists) \equiv \text{exh}K \text{ no}(\alpha, \exists) \equiv K\text{exh} \text{ no}(\alpha, \exists) \equiv \text{exh}K\text{exh} \text{ no}(\alpha, \exists)$

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

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, ‘no(α, \exists)’ logically entails ‘no(α, \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 ‘no($\alpha, \text{exh}\exists$)’), 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

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

³⁸The 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 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 $\text{exh}K(\text{exh})\phi$, 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 $K\phi$ and $K\text{exh}\phi$ are either ruled out by ET or are logically equivalent with $\text{exh}K\phi$ or $\text{exh}K\text{exh}\phi$, and (ii) $\text{exh}K\phi$ and $\text{exh}K\text{exh}\phi$ are never ruled out by ET.

³⁹B. 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.

raised by API does not have a trivial solution as things stand and is a major problem for Magri’s account.

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 whether Magri’s account be generalized (as Magri has suggested) to cover the *Maximize Presupposition!*-related phenomena that were discussed in section 2.2 adequately? My impression is that such extensions are not likely to work, in particular for 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 2008b).

- (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 (2017a), I will now show that disjunctions in general do in fact compete with their individual disjuncts,

⁴⁰There is a recent proposal, advanced in Marty (2017b), 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.

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.
 ALT = { $\underbrace{\text{John is unaware that Mary lives in Paris}}_L, \underbrace{\text{John is unaware that Mary lives in London}}_R$, .

Note that the alternatives *L* and *R* are cannot possibly be innocently excludable with respect to (54). The reason is that since the clausal complement of ‘unaware’ is Strawson-upward-entailing, truth 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.
 ALT = { $\underbrace{\text{John is unaware that Mary lives in Paris}}_L, \underbrace{\text{John is unaware that Mary lives in London}}_R$, .

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).

The reader is directed to Appendix .2 for a detailed discussion of what an analysis that aims to combine Magri’s proposal with Spector & Sudo’s.

2.4.3 Context and polarity

In Magri (2009b) 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 2.5 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, β)’, is strictly logically stronger than its ‘all’-alternative (56b), ‘every(\forall, β)’. Now, suppose *C* is a context in which it is taken for granted that no pro-

⁴¹This 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.

fessor 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:

- (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.

2.5 Logical Integrity (final version)

In subsection 2.4.3 it was established that, contra LI (repeated below from section 2.2.2), 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 $\beta \not\vdash_c \beta'$).

The moral of the data discussed in section 2.4.3 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 $\beta \not\vdash_c \beta'$).

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 2.3.4.

- (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 (38).

- (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).

2.6 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.

2.6.1 Presupposed ignorance, revisited

In section 2.3 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 (2017a) as evidence in favor of PI.

- (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. Mary lives in Paris or London \wedge John lives in Paris or London
 b. 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). The reader is directed to Appendix .3 for further discussion.

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

2.6.2 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 (2009b) 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.]
 \checkmark 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.

2.7 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.

In the following three appendices, I will briefly discuss various alternative analyses of some of the data discussed in this paper. These discussions were prompted by an anonymous reviewer of the journal of *Natural Language Semantics*, for which I am grateful.

Appendices

.1 Percus cases and *Presupposed Ignorance*

There are examples, discovered by Orin Percus (Percus 2010) and subsequently emphasized by the author, which involve a sentence being blocked by one of its alternatives which has a stronger presupposition despite the fact that the presupposition of that alternative is in fact not satisfied in the target context. For example,

- (73) [Context: a game of chess could end in a draw, in which case there are no winners / losers.] John and Mary played a game of chess . . .
 {# A, the} loser got angry.

These examples, call them the Percus cases, can be accounted for without further ado by LI but not PIP (or MP, which I will aside). The reason is that PIP crucially relies on the condition that the presupposition of the relevant competitors be satisfied in the target context. LI, on the other hand, relies on no such condition. On this basis examples like (73), and Percus cases more generally, appear to constitute strong evidence against PIP *as it is formulated by Spector & Sudo (2017b)*. The next question is whether there is a natural modification PIP which captures the Percus cases. This is what I turn to now.

The most salient option to modify PIP is along the lines of the schema in (74). The idea is that (74) accounts for (73) because the definite alternative of the second sentence of (74), namely [*the loser got angry*], can indeed be “felicitously uttered” and it has a stronger presupposition, therefore the indefinite alternative is blocked in (73).

- (74) A sentence S is blocked in context C if S has an alternative S' such that S' has a stronger presupposition and S' can be felicitously uttered in C.

Thus the problematic condition of PIP which required the presuppositions of the competitors to be contextually met is replaced in (74) with the condition that the competitors be felicitous in the relevant context. The condition “S' can be felicitously uttered in C” in (74) needs to be unpacked, which is why (74) is referred to as a schema. More specifically, “S' can be felicitously uttered in C” is here understood to be *weaker* than the original condition that the presupposition of S' be satisfied in C; its content is understood to be something along the lines that *either* the presupposition of S' is satisfied in C *or* . . . , where the blank needs to be filled in.

In Section 2.3.4 I pointed out that (74) makes the following prediction quite regardless of how felicitous-utterance is cashed out.

- (75) If S and S' are alternatives such that S' has a stronger presupposition then there is no context in which both S and S' can be felicitously uttered.

Example (35) in that section (originally due to Heim (1991a)), reproduced below, was meant to refute (74) *tout court* by problematizing (75) without actually getting into the nitty gritty details of what “felicitous utterance” in (74) actually amounts to.

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

The observation is that both of the sentences in (76) can be felicitously uttered in

any context in which it is not common ground just how many nosy neighbors I have. Thus even if I know that I have exactly one nosy neighbor, I can still use both of the alternatives in (76). This fact, it was argued out, is a clear counter-example to (75) and, therefore, a problem for (74) . . . *or so I concluded*. It could be argued that this conclusion is too rash.⁴³ Specifically, it appears that the judgments reported for (76) rely crucially on the choice of competitors, i.e., the articles. Once other cases are considered it turns out that the prediction in (75) is, on the face of it, not too far off the mark. Consider (77) for example. (Below I will use the symbol ‘%’ to report the judgment that a sentence is ‘weird to say’ as opposed to outright infelicitous. The reason for this distinction will be elaborated on below.)

- (77) [**Context:** it is not known how many sisters the speaker has. **Fact:** the speaker has two sisters and he knows it.] % All of my sisters are married.

On the assumption an utterance of the sentence [*both of my sisters are married*] by the same speaker in the context of (77) would have been felicitous then PIP modified along the lines of (74) predicts the weirdness of (77): the speaker could have asserted [*both of my sisters are married*] which as we just agreed would have been a felicitous utterance and which has a stronger presupposition, therefore the *all*-alternative in (77) is blocked. Furthermore, LI does *not* predict the weirdness of (77); from the stand point of LI, (77) should be as acceptable as (76) was above.

It seems then that a broader empirical view turns the table against LI and in favor of PIP as modified along the lines of (74). While granting that there is indeed something special about the example (76) *vis-à-vis* the choice of competitors, in this section I’d like to argue (i) that the Percus cases *do* pose a significant challenge to PIP (modified or otherwise), although one that is not quite as easy to pin down, and (ii) that the weirdness of examples like (77) is not problematic for LI once we take into account the role played by Gricean pragmatics as it interacts with LI.

Let us begin by observing that it is crucial that in (77) the speaker believes that he has two sisters. If, for some strange reason, he in fact does not know how many sisters he has although he has come to believe that all of them are married (or, more plausibly, if he believes that he has *n* married sisters, for some *n* larger than two), then it is totally fine for him to assert [*all of my sisters are married*] (in the context of (77)). On the assumption that the weirdness of (77) follows from PIP as modified in (74), from the contrast just noted it follows that whatever the condition “S’ can be felicitously uttered in C” in (74) amounts to it *must* make reference to the *private* epistemic state of the speaker (i.e., whatever the speaker in fact believes over and beyond what is common ground) *vis-à-vis* the presupposition of the stronger alternative, otherwise the prediction is made that [*all of my sisters are married*] is weird to say regardless of whether the speaker actually believes he has two sisters or not.⁴⁴ Thus, at a minimum, modified PIP must require that the speaker believes the presupposition of the presuppositionally stronger alternative to be true. The result, then, looks something this:

- (78) **Modified Presupposed Ignorance Principle (MPIP).** A sentence S is blocked in context C if S has an alternative S’ such that (i) the presupposition of S’ is

⁴³I thanks an anonymous referee of the journal of *Natural Language Semantics* for helpful comments.

⁴⁴The relevance of the speaker’s *private* belief-state is presumably the key factor that makes this utterance “weird to say” (marked ‘%’) as opposed to outright infelicitous (marked ‘#’).

stronger than S and (ii) the presupposition of S' is either satisfied in C or can be easily accommodated and the speaker believes the presupposition of S' to be true.

Two remarks are in order. First, we do not need to dwell on just when a presupposition “can be easily accommodated” here. In the next subsection I will borrow Chemla’s (2009) concept of *SPEAKER AUTHORITY* to cash this notion out. Here I will assume, for the sake of argument, that every example discussed involves a presupposition that can indeed be easily accommodated. Second, note that whenever a proposition is entailed by the common ground every interlocutor, including the speaker, believes that proposition to be true. Consequently, whenever the presupposition of S' is entailed by the global context C, it follows that the speaker believes the presupposition of S'. Furthermore, we can harmlessly stipulate that whenever the presupposition of a sentence is entailed by the context, that presupposition qualifies as “being easy to accommodate” in a vacuous sense. Putting these two observations together, condition (ii) of MPIP can be stated more elegantly: the presupposition of S' is easy to accommodate and is believed to be true by the speaker. There is no need to mention the global context explicitly. The disjunctive framing of condition (ii) in (78), then, is superficial. Below I will point out that the same cannot be said of the localized version of (78) (which is needed on empirical grounds).

Now, let us see now how MPIP accounts for the basic Percus cases. Consider (73) repeated below.

- (79) Yesterday, John and Mary played a game of chess. # A loser (of the game) got angry.

Since the speaker has asserted [*a loser got angry*], by Quality it follows that he believes that the game ended with at least one loser. But he knows that chess cannot end with more than one losers (if the game ended in a draw then technically nobody lost).⁴⁵ Therefore, he believes that there was exactly one loser. Therefore, he believes the presupposition of the *the*-alternative [*the loser got angry*] to be true. This presupposition is not satisfied in the context as the game could have ended in a draw. Nevertheless, it is a presupposition that can be easily accommodated (by assumption) and the speaker believes it to be true. It follows from MPIP, then, that the *the*-alternative blocks the sentence, whence the oddness of (73) / (79).

I will now discuss two shortcomings of MPIP both of which, it will be noted, involve *embedded* Percus cases. The first counter-example raises a challenge for MPIP specifically while the second raises a problem for how MPIP should interact, according to S&S guide-lines, with the exhaustification operator. As to the first counter-example, consider (80).

- (80) John and Mary played a game of chess. # I doubt that a loser (of that game) got angry (because they have a crush on each other).

Let us first see how LI accounts for the oddness of (80). The clause embedded under “doubt” contextually entails its *the*-alternative in its local context (= the set of all words *w* such that *w* is compatible with every proposition *p* such that it is common ground that the speaker believes that *p*, following Schlenker’s 2009 definition). This

⁴⁵I believe that if the speaker is not familiar with the rules of chess, and if this fact is common ground, the sentence becomes felicitous.

holds because the speaker believes, like everybody else, that if there is a loser in a game of chess (i.e., if a game of chess is not finished in a draw) then there is exactly one loser. As the entailment does not go through logically, LI predicts infelicity.

I claim that MPIP cannot account for the oddness of (80). The reason is this.⁴⁶ The *the*-alternative at root, [*I doubt that the loser got angry*], triggers the presupposition that (the speaker believes that) there was exactly one loser. This presupposition is not satisfied in the context of (80) as no assumption is made in the common ground regarding the speaker's beliefs *vis-à-vis* the outcome of the match. Let us assume, plausibly, that the presupposition is easy to accommodate. Is it the case that the speaker in fact believes this presupposition to be true? Is it the case that the speaker believes that there was exactly one loser / winner? Not necessarily. For example, he could very well think that John and Mary are equally smart and that there is no reason to assume that one of them can beat the other one in a fair game of chess. The crucial difference between (80) and (73) / (79) is that this time we cannot deduce from the asserted sentence and Quality that the speaker believes the presupposition of the alternative to be true. The assertion simply conveys the information that the speaker does not think that there was a loser who got angry: either nobody lost (the game ended in a draw) or the person who lost did not get angry. The speaker can assert this truthfully while at the same time believing that there is a very high likelihood that the game actually ended in a draw because Mary and John are more or less equally smart. MPIP does not predict (80) to be blocked in this kind of situation, contrary to fact, precisely because the speaker does not believe the presupposition of the alternative to be true.

Examples along the same line can be multiplied. I mention (81) and (82) without getting into the details.

- (81) John and Mary are about to play a game of chess in front of an audience from my department. I have no idea how the game is going to end, they both seem equally smart to me, but I know one thing for sure . . . # the audience will not applaud a winner (because they are a bunch of bores).
- (82) [Context: Seven volleyball matches are happening at the same time, there are seven referees. Each match has a referee, no referee watches more than one match. Each match could end in a draw or with one of the teams winning. At the end of each match, if there is a winner, the referee of that match can *but does not have to* nominate the winner to be considered as a candidate for the final prize.] # Interestingly, no referee nominated a winner (of the (only) match he was watching) for the prize.

Let us now move on to the second class of counter-examples. Consider the examples in (83). I will focus on (83a) but the same point can be made with (83b).

⁴⁶Does the *localized* version of MPIP predict the oddness of this sentence, if applied to the embedded clause relative to its local context? I will consider local MPIP below. For the particular case of (80), however, it is important to note that the effect of localization is nullified by the fact that the subject of the attitude in (80) is the speaker himself. The latter guarantees that the relevant local context is some subset of the global context-set (everything that is common ground is believed by the speaker but, naturally, not *vice versa*) and, consequently, the reasoning that we perform by applying MPIP at root applies *mutatis mutandis* to the embedded clause as well. MPIP cannot account for (80) even if it is localized.

- (83) a. # Some professors who brought two guests to the party brought all of their students.⁴⁷
 b. [Context: the teacher has published a list of students with their final grades.]
 Q: How did the students do on their final exam? Do they know the results yet?
 A: # Some students who have looked at the list think that they passed the exam (I don't know about the rest).

LI accounts for the oddness of (83a) as follows. Suppose we apply LI to the scope of the existential predicate in (83a). The local context of this expression, viewed extensionally, is the set of all professors who came with exactly two guests. Nothing is entailed by this local context about how many students these professors have in general. For example, it could be that one of the professors has ten students but came to the party with his two daughters. It follows that the presupposition of the *both*-alternative, [*came with both of their students*], is not satisfied in this local context. Nevertheless, in this local context [*came with all of their students*] contextually entails [*came with both of their students*]: if a professor came with two guests and, furthermore, if it is indeed true that he came with all of his students, then it *must* be the case that he has exactly two students. LI, therefore, predicts (83a) to be blocked.

What about MPIP? At root, the presupposition of the *both*-alternative is indeed stronger than the presupposition of (83a) regardless of specific assumptions about how the existential quantifier negotiates partiality in its scope (universal projection: every professor who came with two guests has exactly two students, existential projection: some professor who came with two guests has exactly two students). This presupposition is not satisfied in the global context. In order for MPIP to capture the oddness of (83a), then, it must be the case that the presupposition of the *both*-alternative is easy to accommodate and the speaker believes it to be true. Let us simply grant that both of these assumptions hold.⁴⁸ MPIP, then, predicts (83a) to be blocked. This is not the end of the story, however. The problem is that matrix exhaustification of (83a) generates a reading that bleeds MPIP. To see this, consider the parse in (84a) (embedded exhaustification in the scope of the quantifier is vacuous in this case). The prejacent of (84a) has at least the three alternatives that are listed in (84b).

- (84) a. EXH [some profs who brought two guests brought all their students]
 b. ALT = {
 [some profs who brought two guests brought both their students],
 [all profs who brought two guests brought all their students],
 [all profs who brought two guests brought both their students] }

Regardless of specific assumptions regarding presupposition projection from the scope of the existential quantifier, the [*some...both...*]-alternative is not innocently excludable with respect to the prejacent given how this notion is defined by S&S (this

⁴⁷This example is a variation on one of the examples discussed in Percus 2010.

⁴⁸If we assume existential projection from the scope of the existential quantifier, (83a) via Quality guarantees that the speaker believes the presupposition of the *both*-alternative. If, however, we assume universal projection then the assertion of (83a) plus Quality is not enough to guarantee that the speaker believes the presupposition of the *both*-alternative.

is because [*some...all...*] and [*some...both...*] are Strawson equivalent, again, regardless of whether the presupposition projects existentially, universally or otherwise). We can also ignore the [*all...all...*]-alternative as it does not have any effect of the present discussion. The [*all...both...*]-alternative is innocently excludable. The meaning of (84a) which would result from the negation of this latter alternative in conjunction with the prejacent is a proposition which presupposes that some / all (depending on specific assumptions about projection) professors who came with two guests have two students. This proposition clearly bleeds MPIP and we predict that (83a) should be felicitous with the parse in (84a). One could argue that as the presupposition generated by EXH is not satisfied in the global context, the parse in (84a) is not available due to presupposition failure. Putting aside the question of why accommodation should be unavailable in this case, we can simply assume that (83a) is uttered in a context in which it is entailed that every professor who came with two guests has two students, or even that every professor has two students. (83a) remains *infelicitous* in such contexts, contrary to what S&S would predict on the basis of how (M)PIP interacts with EXH.

Is it the case that if MPIP is applied locally it can account for the oddness of (83a)? After all, this is exactly how the oddness of this example was accounted for by LI. To see why this will not work, let us first localize MPIP.

- (85) **Local MPIP.** *S* is unacceptable in context *C* if it contains a constituent *A* such that (i) *A* has an alternative *A'* with a stronger presupposition which (ii) is either satisfied in the local context of *A* or can be easily accommodated and is believed by the speaker to be true.⁴⁹

One unfortunate aspect of Local MPIP, in comparison MPIP, is that its condition (ii), unlike condition (ii) of MPIP, must remain uncomfortably disjunctive. The reason is this. Quite generally, if a certain proposition *p* is entailed by a given local context it does *not* follow that the interlocutors and / or the speaker *ipso facto* believe *p* to be true: the whole point of local contexts is that they are in general richer than the global context. The conceptual consequence of this fact is that we can no longer ignore the first disjunct of condition (ii) of Local MPIP: we must make do with a principle that relies on a disjunctive condition.

Regardless of this conceptual issue, Local MPIP *cannot* account for the oddness of (83a). To see this, note first that, as pointed out above, the presupposition of the *both*-alternative of the scope expression in (83a) is not satisfied in its local context (i.e., the relevant local context does not entail that every professor who came with two guests has exactly two students). Furthermore, it does not even make sense to ask whether the speaker believes this the presupposition of *both*-alternative to be true: the presupposition triggered by the *predicate* [*came with both of their students*] is *not* a proposition and, *ipso facto*, is not the sort of thing that can be believed.

I conclude that there seems to be no way to modify PIP to capture the Percus cases in their full generality, i.e., in both embedded and unembedded cases. Perhaps there is a way to modify PIP which satisfies the relevant desiderata and which I have not considered here. I believe it is fair to say that if there is such a modification the onus is on the proponent of PIP to formulate and defend it. I therefore conclude that the Percus cases constitute a strong argument against PIP.

⁴⁹Convention: the local context of a sentence at root is just the global context.

I would now like to look at the Percus cases from the vantage point of LI. Consider the following examples, slightly modified from the examples pointed out to the author by the reviewer.

- (86) a. [**Context:** we know that the speaker has at least two sisters but we don't know how many exactly. **Fact:** the speaker in fact has two sisters who are married and, naturally, the speaker knows this.] % All of my sisters are married.
- b. [**Context:** we do not know whether the speaker speaks French or not. **Fact:** the speaker does in fact speak French and, naturally, he knows it.] % John thinks that I speaker French.

In both cases a certain sentence is judged as “weird to say” by the speaker. This weirdness is not predicted by LI as, in the relevant contexts, the target sentences do not contextually entail their alternatives [*both of my sisters are married*] and [*John knows that I speak French*] respectively. In the previous subsection I pointed out a couple of arguments against the modified version of PIP, however it should be clear that MPIP (i.e., (78)) actually predicts the deviance of these two examples because in each case the presuppositionally stronger alternative is, intuitively, felicitous to assert and the speaker indeed believes the stronger presupposition (i.e., that he has two sisters and that he speaks French, respectively). Is it the case, then, that the data points in (86) are evidence in favor of MPIP and against LI, the arguments in the previous subsection against MPIP notwithstanding?

I would like to argue that an independently needed mechanism, suitably implemented, derives the deviance of the examples in (86) and therefore these examples are not a problem for LI. The independent mechanism that I have in mind is essentially the Gricean system of pragmatic reasoning modified so as to be sensitive to the assertion / presupposition distinction and their role in information-seeking discourse guided by neo-Stalnakerian considerations. To see where I am going with this, note that if I know that I have failed all of my students then it is certainly weird for me to assert that [*I failed some of my students*]. The explanation of this fact is straightforward: there is a more informative alternative, i.e., [*I failed all of my students*], which I know to be true. Not using the more informative alternative is tantamount to being uncooperative whence the weirdness that I sense in using the less informative alternative. My claim, in a nutshell, is that the weirdness of the examples in (86) is due to essentially same considerations but applied to presuppositions. The idea is simple and by no means novel (the neo-Stalnakerian view, Stalnaker 2002; Schlenker 2012a; Chemla 2008b). Presuppositions can be informative. Therefore, if the only difference between two alternatives is that one of them is more informative *because it has a stronger presupposition* (as is the case with examples in (86)) then the alternative with the stronger presupposition, *qua* the more informative alternative, must be used on Gricean grounds (if it is believed to be true) unless, of course, the informative presupposition is not something that the addressee(s) are willing to accept “quietly and without fuss” (von Stechow 2008).

As to the details, I note that Chemla's (2008b) account of anti-presuppositional inferences, originally conceptualized as supplementary to the *Maximize Presupposition!* principle, can be implemented within a general Gricean framework. The result is a system of rational communication that (i) supplements LI coherently and straightforwardly, (ii) is *a priori* plausible, (iii) generates anti-presuppositional inferences as

Chemla originally intended, (iv) derives the deviance of examples like (86) to boot.

I take it that communication is constrained by *a priori* rational considerations. The speaker has certain alternatives at his disposal and his task is to use that alternative which allows him to achieve the goal of information-exchange in the most efficient manner. Thinking in classic Gricean terms, certain principles have been isolated as being particularly pertinent to communication: Quantity, Quality, Relevance and (the ill-understood) Manner. Let me put aside the latter and put together the former three in one statement.

- (87) Given a set of relevant alternatives that you believe to be true, use the one that is most informative.

I would like to note that there is a fifth principle, pertaining to presuppositions going back to the work of Stalnaker. Abstractly speaking, suppose I have two objects at my disposal and I'd like to choose the one that achieves my goal most efficiently given the contingencies of the surrounding context. Suppose further that appropriate use of one of these objects requires certain conditions to be met by the context. I take it as a truism that if this condition is not met then that object cannot possibly be the rationally optimal choice. Specifically, then, thinking of presuppositions as preconditions that must be met by the context for sentences be appropriately usable, we can safely assume that the more complete rendition of (87) is (88).

- (88) Given a set of relevant alternatives that can be felicitously used and that you believe to be true, use the one that is most informative.

For (88) to make concrete predictions we need to spell out what felicitous utterance means. In the previous section, I analyzed this notion as meaning that the presupposition of the alternative can be easily accommodated and is believed to be true by the speaker. Here, we can take the next step and follow Chemla (2009) in assuming that a presupposition can be easily accommodated if and only if the speaker has authority over his interlocutors with respect to this presupposition. The idea, intuitively, is that the speaker has authority over the hearer with respect to p iff the sheer fact that the speaker asserts a sentence that presupposes that p is enough for the hearer to accept that p is true. Looking back at examples in (86), for example, I assume that the speaker has authority over his interlocutors with respect to how many sisters he has or whether he speaks French. We can, then, state (88) more precisely as follows.

- (89) Given a set of alternatives that satisfy the following conditions, use the one that is most informative.
- a. (You believe that) the alternatives are relevant.
 - b. You believe the alternatives to be true.
 - c. You believe that you have authority with respect to the presuppositions of the alternatives.

Two remarks are in order. First, note that the felicity condition (89c) only refers to authority, not speaker's belief in the presupposition of the alternative. This is because condition (89b) already requires the speaker to believe the presupposition of the alternative (and its assertive content). Second, (89) relies on the notion of informativity so we need to say a few words about this concept. As (89) is a principle of pragmatics, informativity ought to be defined with respect to the information that

is contextually encoded. The core intuition is that S_1 is more informative than S_2 in context C iff the context that results from C being (felicitously) updated with S_1 is richer than the context that results from updating C with S_2 . For our purposes we can define informativity with respect to what is common ground that the speaker believes.

- (90) S_1 is more informative than S_2 in context C iff S_1 is felicitous in C , and were the speaker to assert S_1 in C the resulting context would be properly included in the one that would result from uttering S_2 .

given what is common knowledge in C , if s believes S_1 then he believes S_2 as well but not *vice versa*.⁵⁰

Now, let us have a second look at (86a), repeated below (the same point can be made with (86b)).

- (91) [**Context:** we know that the speaker has at least two sisters but we don't know how many exactly. **Fact:** the speaker in fact has two sisters who are married and, naturally, the speaker knows this.] % All of my sisters are married.

Let us assume that the alternative [*both of my sisters are married*] is relevant. We can also safely assume that the speaker has authority with respect to how many sisters he has. Furthermore, note that [*both of my sisters are married*] is more informative than [*all of my sisters are married*] as the former gives us more information about the speaker's epistemic state than the latter; specifically, it tells us that the speaker believes that he has two sisters. It follows from (89) that the speaker uttered [*all of my sisters are married*] because he does not believe [*both of my sisters are married*] to be true. But the speaker believes [*all of my sisters are married*] to be true (since he asserted it, by Quality, i.e., (89b)). Therefore, the question is whether the speaker believes he has two sisters or not. If he does (as assumed in (86a) / (91)) then by using the *all*-sentence instead of the *both*-sentence he is being "uncooperative", thus the weirdness of the *all*-sentence.

Is there any reason to stipulate a principle like LI (or indeed MP / PIP) on top of (89)? In other words, doesn't (89) undermine the whole motivation behind LI to begin with? To see that a principle like LI is still needed, note that most of the examples discussed in the literature on MP / PIP, and in this paper on LI, involve the infelicity of sentences that are *not* more informative than their alternatives. The classic MP-effect in (92a) is one such case. Here the *both*-alternative is in fact not more informative because it is already established that the speaker has two sisters thus it is common ground that if the speaker believes the sentence [*all of my sisters are married*] he also believes the alternative [*both of my sisters are married*] (and *vice versa*, the two alternatives are equally informative in this context). The same applies to the Magri-case in (92b).

- (92) a. [Context: the speaker has two sisters.] # All of my sisters are married.
 b. [Context: John gave the same grade to his students.] # John gave some of his students an A.

Another type of example involves alternatives that are arguably semantically identical

⁵⁰In pleasant formalism: S_1 is more informative than S_2 in context C relative to speaker s iff $BEL_s(S_1) \not\subseteq_C BEL_s(S_2) \wedge BEL_s(S_2) \not\subseteq_C BEL_s(S_1)$.

at root, (93) (see Percus and Singh in particular for the relevant discussion, see also Anvari for a problematic case of the same genus). Thus given standard assumptions the sentence in (93a) denotes the same proposition that is denoted by its *both*-alternative [*if John has exactly two sisters then he is going to invite both of them.*]. If these two alternatives really are semantically identical then, naturally, neither can be more informative than the other one.

- (93) a. #If John has exactly two sisters then he is going to invite all of them.
 b. #If John gave the same grade to his students then he gave some of them an A.

The examples discussed in (92) and (93) cannot be accounted for by (89) only under the standard conceptualization of informativity. A dissenting view has been advanced by Schlenker according to which if S' is logically stronger than S then it is also more informative even in contexts in which the two end up being contextually equivalent (i.e., in contexts in which the truth of each guarantees the truth of the other). Suppose, for example, that we are interested in the alternatives [*a sun is shining*] and [*the sun is shining*]. The latter is logically stronger in the sense that the truth of the latter guarantees the truth of the former but not *vice versa*. Now, in a context in which it is entailed that there is exactly one sun, the two alternatives become contextually equivalent: any world in that context is one in which there is exactly one sun, therefore if [*a sun is shining*] is true in that world then [*the sun is shining*] is true as well (and the reverse always holds). Schlenker's intuition is that [*the sun is shining*] is still "more informative", in a technical sense, because there is a chance, greater than zero however small, that the interlocutors might not be "attending" to the proposition that there is exactly one sun even if they believe it to be true. If one adopts Schlenker's idea then (89) can indeed capture the data points in (92), as argued in detail by Schlenker himself (the same approach can in be applied to the examples in (93) at the level of local contexts.).

Importantly, even if one subscribed to Schlenker's view (89) still does not suffice to capture all the data points. The reason is that there are cases where *the blocked alternative happens to be the more informative one*. For example, in (94a) the *some*-sentence is in fact more informative (because it conveys more information about John's epistemic state) and yet it is blocked by the *all*-alternative. Similarly for (94b).

- (94) a. [Context: all students smoke.] # John is unaware that some of the students smoke.
 b. [Context: every professor in this department gives the same grade to all of his students.] # Not all professors gave an A to some of their students.

To sum up, it appears that no amount of conceptual gymnastics allows (89) to capture all the data points involved in this paper, the core reason being that in most of the relevant cases a sentence is blocked by an alternative which is not more informative than it.

I conclude, then, that on the one hand Schlenker's proposal does not cut quite deep enough and on the other hand LI can be straightforwardly combined with a formulation of Gricean pragmatics that is sensitive to presupposition / assertion distinction *à la* recent (neo-)Stalnakerian work. The resulting framework relies, on the one hand, on essentially Gricean reasoning to account for anti-presuppositional infer-

ences, much like Chemla’s proposal, and the general preference for more informative alternatives while relying on LI, on the other hand, for felicity-conditions in cases where the informativity condition of the Gricean system fails to mark a preference for the relevant alternative.

.2 On why *Presupposed Ignorance* and Magri’s system cannot be innocently combined

Consider the sentence in (95a) with the LF in (95b).

- (95) a. [Context: all students smoke.] # John does not know that some of the students smoke.
 b. NEG [John knows [some of the students smoke]]

Let us first approach (95) from the point of view of the system described by S&S. The *all*-alternative of the sentence in (95a), namely [*John does not know that all of the students smoke*], carries a stronger presupposition which is satisfied in the context of (95a). PIP therefore predicts the sentence in (95a) to be blocked *unless* there is a parse of this sentence in which exhaustification bleeds PIP.

It is plain that a parse with matrix exhaustification, (96a), is useless here because the scalar item “some” occurs in a Strawson DE environment, which means that at root [*John does not know that some of the students smoke*] does not have an innocently excludable alternative (given S&S’s definition of EXH). There are, however, (at least) two more options to consider given in (96b) and (96c).

- (96) a. EXH NEG [John knows [some of the students smoke]]
 b. NEG [EXH John knows [some of the students smoke]]
 c. NEG [John knows [EXH some of the students smoke]]

Now, (96c) can do anything but rescue the utterance of (95a) from oddness. The embedded EXH in (96c) *either* does not trigger any implicature, in which case the parse in (96c) is synonymous with the parse (95b) and is ruled out by PIP for the same reason, *or* it does trigger the *not-all* implicature in which case the parse at root ends up denoting a proposition which presupposes that some *but not all* students smoke; this presupposition contradicts background assumptions and, again, we predict oddness.

What about (96b)? Can it generate a reading that bleeds PIP? In fact it can. If the embedded EXH in (96b) negates the *all*-alternative of its prejacent, the clause embedded under negation is going to denote a proposition which presupposes that all students smoke and assert that John only knows that some of the students smoke. At root, then, we get the proposition which also presupposes that all of the students smoke (as negation is a presupposition-hole) and asserts that either John does not believe that some of the students smoke or he believes that all of them do. This proposition carries a strong presupposition and is, therefore, not ruled out by PIP. Hence we do not predict the utterance in (95a) to be odd unless the parse (96b) is somehow independently ruled out.

As it happens, S&S do have a reasonable hypothesis for why the parse in (96b) should be ruled out independently from PIP. It is a standard assumption that the exhaustification operator *cannot* be embedded in DE environments (such as the scope

of negation) unless certain conditions are met. The nature of these conditions is not entirely well understood (see Fox & Spector 2018 for a recent attempt and references therein) but it is fair to say that whatever those conditions are, they are not met in the case of (96b). S&S can therefore discard the parse in (96b) on the assumption that it falls under the generalization that exhaustification cannot be embedded in DE environments.^{51, 52}

Now, what does PIP+Magri predict for (95a)? It is a hallmark of Magri's system that the exhaustification operator is obligatorily inserted in every scope site, including DE environments. Empirical evidence for the latter assumption comes from the oddness of examples like (97).

- (97) [Context: every professor in this department gives the same grade to all his students.] # Not every professor gave some of his students an A.

The oddness of (97) cannot be accounted for by Magri unless at least one of the two parses in (98) is assumed to be available (and, indeed, unescapable). Note that in both cases EXH is embedded in a (globally) DE environment.

- (98) a. NEG [EXH [every professor [... some ...]]]
 b. NEG [every professor [EXH ... some ...]]

The same point can be made with the oddness of the sentence [*John did not go to the museum with Bill or Bob*] in a context in which it is known that Bob and Bill are Siamese twins.

The parse of (95a) according to PIP+Magri, then, must be (99). (Different occurrences of the exhaustivity operator are indexed for ease of reference.)

- (99) EXH₁ NEG EXH₂ John knows EXH₃ some of the students smoke

EXH₁ is inconsequential because it does not have an innocently excludable alternative given the semantics of the exhaustivity operator which is proposed by S&S.⁵³ Let me also ignore EXH₃ for now, I will return to it below.

Zooming in on EXH₂, then, what does *it* do? If we are following Magri then

⁵¹It is standardly assumed that the exhaustification operator cannot be embedded in DE environments *because* in such environments it would weaken the overall meaning of the sentence (e.g., this is the core intuition that Fox & Spector pursue). As it happens, this idea does not straightforwardly apply to (96a). The reason is that although in this parse exhaustification is indeed weakening the assertive content of the overall sentence (whence the disjunctive truth-conditions described above), it is at the same time strengthening the presuppositional component (from some students smoke to all students smoke). One might take this as evidence that the relevant notion of entailment needed to define DE environments, at least as far as the constraint on EXH is concerned, is Strawson entailment (see Sharvit 2017 for relevant discussion).

⁵²A different way to rule out the parse in (96b), one might argue, is to capitalize on the fact that its assertive content is hopelessly disjunctive. There might be independent pressure against such disjunctive meanings (recent work by Émile Enguehard and Emmanuel Chemla). I do not believe this kind of approach is pertinent here. The reason is that, with plausible assumptions, the parse [EXH [NEG [EXH [John knows [some of the students smoke]]]]] is also going to trigger the presupposition that all of the students smoke while at the same time trigger the assertion that John believes that all of the students smoke. This reading also bleeds PIP without having a disjunctive assertive content.

⁵³To be precise, EXH₁ is *either* vacuous *or* it generates a reading identical to the one pointed out in fn. 52 if EXH₂ or EXH₃ generate the *not-all* implicature. This does not affect the argument made in the body of the text because regardless of whether EXH₁ is vacuous or not, PIP+Magri predicts that (95a) should be acceptable in the target context.

we are committed to his assumption that relevance is necessary and sufficient to generate a scalar implicature. The question, then, is whether the alternative [*John knows that all of the students smoke*] is relevant or not. I submit that it follows from S&S's assumptions that this alternative must be relevant. The reasoning is as follows. According to S&S the alternative [*John does not know that all of the students smoke*] is a relevant alternative at root because that is the alternative that, according to S&S, blocks (95a) via PIP: if this alternative was not relevant, PIP as a principle of language use would be blind to it. But it is an axiom on relevance that it is closed under negation. Therefore, if [*John does not know that all of the students smoke*] is relevant then [*John knows that all of the students smoke*] is relevant too. But if the latter alternative is relevant then EXH₂ is bound to negate it because, recall, according to Magri relevance is necessary *and sufficient* to trigger an implicature. Therefore, the constituent [EXH₂ John knows (EXH₃) some of the students smoke] (ignoring EXH₃, see below) is going to trigger the presupposition that all students smoke and the assertion that John only believes that some students do. And again at root we get the presupposition that all students smoke which is strong enough to prevent PIP from blocking (95a) in the target context.

The picture, then, looks like this: either the *all*-alternative is not relevant at root in which case PIP cannot use it to block the *some*-sentence or it is relevant in which case EXH₂ is going to make sure that the presupposition that all of the students smoke is triggered, which is going to bleed PIP at root. Either way, the utterance (95a) is predicted to be felicitous.

The conclusion so far is that EXH₂ in (99) is going to make sure that PIP is bled (unless there is no relevant alternatives to apply to being with). It follows that the oddness of (95a) must be captured in some other way. One possibility is exhaustification below the attitude in (99), i.e., EXH₃. Notice that if EXH₃ triggers the *not-all* implicature we are going to get a presupposition which is contextually contradictory, and we can use this to explain the oddness of (95a). Now, the question is whether this operator is going to do that or not. As we are working within Magri's framework, we need to ask whether the alternative [*all students smoke*] is relevant for the embedded clause or not, which in turn forces us to ask what is the context that determines relevance for the embedded clause of (99). Fortunately, we need not get into this somewhat thorny issue here.⁵⁴ If [*all students smoke*] is not relevant for the

⁵⁴The two salient options are,

1. The context that determines what is relevant for the embedded clause is the global context. In this context [*some students smoke*] and [*all students smoke*] are contextually equivalent. As relevance is closed under contextual equivalence (Magri's assumption), [*all students smoke*] is relevant on the assumption that [*some students smoke*] is relevant to begin with (it is in fact not quite clear why the latter assumption should be true).
2. The context that determines what is relevant for the embedded clause is the local context of the embedded clause. This context, following Schlenker's 2009 procedure, is the union of the global context set (because "aware" is factive) with the set of all worlds *w* such that *w* is compatible with every proposition *p* such that it is common ground that John believes that *p*. Now, in this context [*some students smoke*] and [*all students smoke*] are *not* contextually equivalent simply because nothing is assumed in the common ground regarding John's epistemic state *vis-à-vis* the quantity of students who smoke. Furthermore, there is no other axiom on relevance that guarantees the relevance of this alternative in this context. Therefore, we have no reason to assume that [*all students smoke*] is relevant and we can proceed on the assumption that it is in fact not relevant.

embedded clause in (99) then EXH₃ in (99) is going to be vacuous, EXH₂ is going to bleed PIP, and the oddness of (95a) remains unaccounted for. If, on the other hand, [*all students smoke*] is relevant for the embedded clause in (99) then EXH₃ is going to trigger the *not-all* implicature, at root we are going to get the presupposition that some but not all students smoke which is contextually contradictory and we can explain the oddness of (95a).

Can we, then, conclude that all we need to do is to stipulate that [*all students smoke*] is relevant for the embedded clause in (99) and the problem for PIP+Magri disappears? Not quite, we now have another problem. The difficulty is to find a way to make sure that the *all*-alternative is relevant for the embedded clause in (99) but not for the embedded clause in (100b).

- (100) a. [Context: all students smoke.] John knows that some students smoke.
 b. EXH [John knows [EXH some students smoke]]

As S&S point out (100a) is felicitous (with focus on “some”, which I am going to ignore). If (95a) is ruled out *because* embedded exhaustification gives rise to a contextually contradictory presupposition, then why does the same thing not apply to (100a)? Put differently, if the *all*-alternative is relevant for the embedded clause of (99), then why is it *not* relevant for the embedded clause of (100b)?

It is entirely unclear how this question should be answered. Relevance is determined by QUD and the information encoded in the common ground. (95a) and (100a) are asserted against the same set of background assumptions and there is no obvious sense in which the QUD is different for the two cases. The problem is magnified by the assumption that relevance is closed under negation in conjunction with the fact that the two sentences are negations of each other. In a nutshell, PIP+Magri needs to assume that there is a crucial difference between the set of relevant propositions in (95a) and (100a) even though it is completely unclear why there should such a difference because (i) the common grounds are identical, (ii) the asserted sentences are negations of one another therefore the relevance of one forces the relevance of the other (so, presumably, any proposition that is made relevant by the utterance of one is also made relevant by the other), and (iii) there is no perceptible difference between QUDs of the two cases.

Once the general structure of the problem is made clear, it is not difficult to come up with other problematic examples.

- (101) [Context: every student smokes.] # If John is aware that some of the students smoke then we should fire him.

The oddness of (101) follows from LI applied at root. It also follows from PIP on the crucial assumption that the antecedent of the conditional is one of the environments in which EXH cannot be freely inserted. It is less clear if PIP+Magri can account for the oddness of (101), however. On the latter account, EXH is obligatorily inserted in the antecedent clause. The question boils down to whether the alternative [*John is aware that all of the students smoke*] is relevant or not. This time axioms on relevance to not guarantee that this alternative is relevant although it is rather difficult to think of a context in which the utterance of (101) itself is relevant, the *all*-alternative at root is relevant, but the *all*-alternative of the consequent is not. Regardless, the prediction made by PIP+Magri is that (101) is felicitous in any context on which [*John is aware that all of the students smoke*] is relevant which does not seem correct.

Let me briefly mention another example. Consider (102). The acceptability of this example is surprising on the assumption that the second sentence has [*John_F failed all of his students, too*] as an alternative which has a stronger presupposition (that Mary failed all of her students) which is satisfied in the relevant context.

(102) Mary failed all of her students. John_F failed most of his students, too.

S&S attempt to account for the acceptability of (102) on the assumption that the second sentence is parsed with matrix exhaustification.

(103) EXH [too [John_F failed most of his students]]

(103) is going to denote the proposition which presupposes that Mary failed all of her students and assert that John failed most but not all of his students. Evidence for this account comes from the fact that the *not-all* implicature in the assertive component is indeed intuitively perceived in (102). But if we look at this example from the lens of PIP+Magri, then the second sentence of (102) must receive the parse in (104) on the assumption that EXH is obligatorily inserted in every scope site.

(104) EXH₁ [too [EXH₂ [John_F failed most of his students]]]

Now, there are two possibilities. Either the matrix alternative [*John_F failed all of his students, too*] is relevant or it is not. S&S assume that it is, and indeed it would be strange if it were not since, as noted above, the *not-all* implicature is clearly perceived in (102). But if this alternative is relevant, then [*John failed all of his students*] is also relevant because the two are contextually equivalent: the presupposition of the former is satisfied in the relevant context and the two have the same assertive content. But if the latter alternative is relevant then EXH₂ is bound to negate it as relevance is sufficient to trigger an implicature. The result is that [EXH₂ [John failed most of his students]] is going to denote the proposition that John failed some but not all of his students. If we follow S&S in assuming that “too” is anaphoric to Mary it follows that “too” is now going to introduce the presupposition that Mary failed most but not all of her students which is contextually contradictory (EXH₁ becomes vacuous as it now lacks an innocently excludable alternative). In other words, PIP+Magri seems to be committed to the prediction that (102) is infelicitous, contrary to fact.

This concludes my discussion of the prospects of a naïve combination of S&S’s and Magri’s proposals. The moral of the preceding discussion is that these two systems, although technically compatible, cannot be easily combined without some empirical cost. Note that two assumptions in particular cause the problems mentioned above for PIP+Magri. These are that for an innocently excludable alternative to be negated by EXH relevance is (necessary and) sufficient and that EXH is obligatorily inserted in every scope site. These happen to be precisely the two assumptions that form the substance of Magri’s proposal.⁵⁵ I do not claim that there is no way to fix the problem raised above one way or the other. I do claim that doing so is no trivial task and until such a “bug-free” implementation of PIP+Magri is offered the presumption that PIP+Magri is a viable competitor to LI is unfounded.

⁵⁵Magri’s third assumption was that EXH is blind to contextual knowledge.

.3 Some provisional remarks about *too*

My aim in this subsection is to argue that the data involving the additive particle *too* are much more complex than predicted by S&S's proposal. The conclusion will be that their proposal is, at best, woefully incomplete. To start the discussion, let me begin with the contrast in (105).

- (105) a. [Context: all students smoke.] John is aware that some of the students smoke.
 b. [Context: all students smoke.] # If John is aware that some of the students smoke, he should take appropriate action.

The contrast in (105) has an elegant explanation in S&S's framework. The sentence (105a) competes with its *all*-alternative and is predicted by PIP to be blocked by the latter. However, matrix exhaustification can be used to bleed PIP in this case. The same strategy does not apply to (105b), however, on the plausible assumption that the antecedent of the conditional is not an environment in which the exhaustivity operator can be freely inserted. I will now point out two pieces of data that do not fit the same pattern, generating a serious problem for S&S's analysis. First, consider (106).

- (106) Mary has invited all of her students. . .
 a. John has invited most of his students, too.
 b. If John invites most of his students too, the room will be packed.

PIP predicts (106a) to be blocked by its *all*-alternative, [*John has invited all of his students, too*] as the latter has a stronger presupposition which is satisfied in the relevant context. However, similar to (105a), matrix exhaustification can rescue this sentence from oddness. Support for the claim that (106a) is rescued by matrix exhaustification comes from the observation that it is most naturally interpreted as implying that John has not invited all of his students. But notice that (106a) remains acceptable even if it is embedded in the antecedent of the conditional, as in (106b). The contrast between (106b) and (105b) is the first puzzle for S&S's system.

- (107) The puzzle of (106b) / (105b):
 a. If (105b) is odd because exhaustification cannot bleed PIP in the antecedent of the conditional, then why is (106b) acceptable?
 b. If (106b) is acceptable because exhaustification can bleed PIP in the antecedent of the conditional, then why is (105b) odd?

Second, consider the contrast in (108).

- (108) Mary speaks German and French. . .
 a. # John speaks (either) German or French, too.
 b. If John speaks German or French too, they can cooperate on this project.

Let us consider the three examples in (109).

- (109) a. # Mary was born in Paris. John was born in Paris or London, too.
 b. Mary invited all of her students. John invited most of his students, too.

- c. # Mary speaks German and French. John speaks (either) German or French, too.

The oddness of (109a) is, of course, predicted by the system laid out by S&S. PIP blocks the second sentence of (109a) as mentioned above and exhaustification cannot bleed PIP in this case because in the parse [EXH [TOO [John_F was born in Paris or London]]] the exhaustification operator has only one innocently excludable alternative, namely [*John was born in Paris and London*] which presupposes that Mary was born in Paris and London. Since this presupposition is negated by the context as a matter of principle (no individual can be born in two places),⁵⁶ the meaning generated by this parse is contextually contradictory. Either way, (109a) ends up being infelicitous.⁵⁷

As pointed out by S&S, examples (109b) and (109c) put together are problematic for their account. On the one hand, the felicity of (109b) seems to be due to matrix exhaustification bleeding PIP, since the latter predicts the alternative [*John invited all of his students, too*] to block the second sentence of (109b). Evidence in support of this idea is provided by the intuition that this sentence licenses the inference that John did not invite all of his students. On the other hand, the infelicity of (109c) suggests that matrix exhaustification is not able to bleed PIP in this case perhaps because the additive particle and the exhaustivity operator are both focus-sensitive operators and this (for whatever reason) prevents EXH from scoping above *too*. Note that in the parse [EXH [TOO [John_F speaks G or F]]] the exhaustification operator can negate the alternative [*John speaks G and F, too*] thereby triggering the presupposition that Mary speaks both German and French which is strong enough to bleed PIP and it known to be true as well. This is where S&S end their discussion: both (109b) and (109c) are in principle blocked by PIP but for some reason or other matrix exhaustification bleeds PIP only in the former case.

I would like to argue that the felicity of (109b) is not due to exhaustification (at least not necessarily) and that the infelicity of (109c) is in all likelihood not due to PIP. As to the former claim, note that the sentences in (110) are felicitous even though in (110a) the *not-all* implicature is explicitly rejected and in (110b), and perhaps also (110c), no implicature pertaining to the *all*-alternative is perceived at all. These data points, (110b) in particular, suggest at a minimum that exhaustification is not the only mechanism that can bleed PIP, otherwise all three examples in (110) are predicted to be odd.

- (110) Mary invited all of her students. . .
- a. John invited most, possibly all, of his students too.
 - b. If John invites most of his students too, the room will be packed.
 - c. If John is smart, he will fail most of his students too.

My argument for the second claim is slightly less concrete. The observation I'd like to make is that while the oddness of (109a) is robustly reproduced in embedded positions, (111), the oddness of (109c) disappears as soon as the second sentence of

⁵⁶The choice of example here is meant to put certain orthogonal complications under the rug. Specifically, note that in order to account for the infelicity of an example like #*[Mary speaks German, John speaks German or French too]* S&S need to assume either that the first sentence is read exhaustively (Mary speaks only German) or that matrix exhaustification of the second sentence is not possible because its presupposition, that Mary speaks both German and French, is not supported by the context.

⁵⁷The reader can verify that other parses, e.g., having EXH scope below the additive particle, or having two occurrences of EXH one above and one below the additive, do not change this conclusion.

(109c) is put in an embedded position, (112). This contrast suggests that whatever is behind the oddness of (112a) is not quite as “robust” as the mechanism behind the oddness of the sentences in (111).

- (111) Mary was born in Paris. . .
- a. # John was born in Paris or London too.
 - b. # If John was born in Paris or London too, they can cooperate on this project.
 - c. # If John is her brother, he was born in Paris or London too.
- (112) Mary speaks German and French. . .
- a. # John speaks German or French too.
 - b. If John speaks German or French too, they can cooperate on this project.
 - c. If John is her husband, (presumably) he speaks German or French too.

What can we make of this situation? The most conservative approach from S&S’s point of view would be to introduce one or more mechanisms, distinct from exhaustification, that can bleed PIP. This will allow them to deal with cases where PIP seems to rule out felicitous utterances. It is entirely unclear at the moment what the nature of these other mechanisms would be. The desiderata are highly intricate: one needs to predict the felicity of (109b) and (110) while predicting the oddness of, e.g., (113) and at the same time accounting for the contrast between (111b) and (112b).

- (113) [Context: all of the students smoke.] # John is unaware that most of the students smoke.

I would like to suggest an alternative route. Let us discard S&S’s proposal completely and look at the puzzling data from a fresh perspective. The felicitous cases are, of course, completely compatible with LI. For the infelicitous cases, it is clear that LI needs to be supplemented somehow. The contrast in (112), I’d like to suggest, is evidence that whatever is responsible for the oddness of (112a) is a speech-act level phenomenon which cannot quite reach the embedded positions and, therefore, is likely to be completely orthogonal to the present deliberations. What remains is the oddness of S&S’s original example in (111). I’d like to suggest the lesson that S&S took away from this example is incorrect: (111a) is not odd because it has an alternative with a stronger presupposition which is known to be true, rather it is odd because it has a *disjunctive* alternative the presupposition of which is known to be *false*. More specifically, I’d like to suggest the following generalization.

- (114) Let $\phi[\alpha \vee \beta]$ be a sentence, containing the disjunction $[\alpha \vee \beta]$, which presupposes that $[\alpha \vee \beta]$ is true. $\phi[\alpha \vee \beta]$ is odd in context C if $C \models \neg\alpha$ or $C \models \neg\beta$.

This constraint, certainly nothing more than a stipulation at this point, gains some *prima facie* plausibility if it is viewed as part of the intuition that felicitous use of disjunctions requires a certain kind of balance to be established between individual disjuncts. We can thus re-frame the generalization as follows.

- (115) Let $\phi[\alpha \vee \beta]$ be a sentence, containing the disjunction $[\alpha \vee \beta]$, which presupposes that $[\alpha \vee \beta]$ is true. If $\phi[\alpha \vee \beta]$ is felicitous in context C then *either* $C \models \alpha$ and $C \models \beta$ or $C \not\models \alpha$ and $C \not\models \beta$.

The idea is that if $\phi[\alpha \vee \beta]$ triggers the presupposition that $[\alpha \vee \beta]$, then it can be felicitously uttered only in contexts in which either both disjuncts are known to be true or neither is known to be true. The third possibility, that both disjuncts are known to be *false*, is automatically leads to infelicity on the assumption that $\phi[\alpha \vee \beta]$ presupposes that $[\alpha \vee \beta]$.

PIP and (115) can be empirically distinguished. Consider,

- (116) Mary was born in Paris. John was born in London. Bill was born in Paris or London, too (I think).

The felicity of this example is not immediately predicted by PIP. After all, the third sentence has no less than two alternatives with stronger, known-to-be-true presuppositions. The contrast between this example and (109a) / (111a) is entirely in line with the generalization above, however. Note that in this example the third sentence does *not* have a disjunctive alternative the presupposition of which is known to be false.⁵⁸

⁵⁸B. Spector (p.c.) points out that this example can be accounted for within the PIP analysis on the assumption that the relevant antecedent for *too* in the third sentence is the plurality consisting of Mary and John. That *too* allows for split antecedents along these lines is also suggested by the following example.

- (i) Mary solved the first problem. John solved the second problem. Peter, too, solved the first and second problems.

Chapter 3

A Problem for *Maximize Presupposition!* (locally)

Maximize Presupposition! (MP) is a principle of language use (rooted in Heim 1991a), which says that of two competing forms that are equally informative one must use the one which has a stronger presupposition, unless this presupposition is not satisfied in the particular context. This derives the oddness of (1) with the all-variant, because the assumed context supports the stronger presupposition of the both-competitor (that John has exactly two students).

- (1) Context: *John has two students.*
John invited {#all, both} of his students.

Either the above formulation of MP is on the wrong track or MP is not a speech-act-level principle. A convincing argument for this claim comes from an observation due to Percus (2006b). In (2), *all* is intuitively odd in the same way as it is in (1). However, the presupposition triggered by *both* in (2) is filtered through the antecedent of the conditional; the *both*-sentence presupposes that if John has two students, then he has two students, which is tautologous. Therefore, the presupposition of the both-sentence as a whole is not stronger than its *all*-alternative.

- (2) If John has exactly two students, he will invite {#all,both} of them.

Percus takes (2) as evidence that the standard formulation of MP is on the wrong track. His proposal is that whenever two alternatives are equally informative, the one which contains an occurrence of the presuppositionally stronger lexical item must be used. Thus in (2) the two alternatives are equally informative because neither can be true without the other being true as well; therefore, since *both* is a presuppositionally stronger lexical item than *all*, the *both*-sentence is preferred. Singh (2011) takes (2) as evidence that the standard formulation of MP is on the right track except that it needs to be relativized to local contexts (Heim 1983a; Schlenker 2009b). The local context of the *all* / *both*-consequents in (2) is the set of possible worlds that verify the antecedent, i.e. worlds in which John has exactly two students. Relative to this context, the stronger presupposition of the *both*-consequent is satisfied. The *all*-consequent is therefore predicted to be infelicitous, as desired.

The contrast in (3) is problematic for Singh's proposal.

- (3) I am critical of {#all,both} of the two mainstream presidential candidates.

Singh seems to predict no contrast between the two sentences in (3). This is because we do not expect the DPs [*both of the two candidates*] and [*all of the two candidates*] to differ with respect to the presuppositions they trigger. [*The two candidates*] already presupposes that there are exactly two candidates, and the same presupposition should, by standard assumptions, be triggered by [*both of the two candidates*] and by [*all of the two candidates*]. Note that [*all of the n candidates*] generally inherits the presupposition triggered by [*the n candidates*]:

- (4) a. Are you critical of (all of) the n mainstream candidates?
b. If you are critical of (all of) the n main stream candidates, who should we vote for?
 \rightsquigarrow *There are exactly n candidates.*

(3), though challenging, might ultimately be surmountable by a revision of Singh's account. One possibility, suggested by a reviewer, is to define the local context of both/all as the context that results after processing the restrictor. The tenability of this move, however, and the details of how it can work, must be left to future discussion. Percus's proposal, on the other hand, captures the contrast in (3) straightforwardly.

Chapter 4

The Problem of Reflexive Belief in Spanish (with Mora Maldonado and Andrés Soria)

4.1 Introduction

What we will refer to as the Reflexive Belief Construction in Spanish, RBC for short, is built by adjoining the reflexive pronoun to the predicate *creer* (*to believe*), as in (1).¹ Note that the reflexive pronoun must agree with the subject DP and therefore in (1) it appears in its third-person, singular form.

- (1) Juan **se** cree que está lloviendo.
'Juan REFL believes that it is raining'

We will zoom in on one particular interpretive aspect of RBCs which to our knowledge has not been discussed in any detail. An utterance of (1) naturally implies, on top of the expected entailment that Juan believes that it is raining, that it is in fact not raining, (2). Our goal is to understand the nature and the origin of this inference.^{2,3}

- (2) Juan **se** cree que está lloviendo.
'Juan REFL believes that it is raining'
↪ *it is not raining*

¹The reflexive pronoun in Spanish (and Romance languages in general) serves a variety of functions beyond its standard argumental role. In particular, the Spanish reflexive has been claimed to act as an aspectual (“telicity”) operator in certain constructions (Bogard 2006). RBCs, as discussed in this paper, have been recently categorized as involving another non-argumental use of the reflexive (DiTullio 2018), but their relation with other uses of the reflexive remains to be explicated.

²The data reported in this paper are based on introspective judgments of a number of native Argentinian and Spanish speakers. We believe the core generalizations carry over to other varieties of Spanish, excluding those in which *creerse* is not grammatical to begin with.

³Several recent studies focus on attitudinal predicates that show a similar “negative bias” as *creerse* (Kierstead 2014 for *akala* in Tagalog, Hsiao 2017 for *lih-tsun* in Taiwanese Southern Min, and Glass 2019 for *yǐwèi* in Mandarin). We were not able to access Kierstead’s article for the specifics. Comments on the differences between *creerse* and *yǐwèi* can be found in footnotes 9, 13 and 14. It appears that *creerse* and *lih-tsun* share substantial similarities, although there are at least two crucial differences (see footnotes 13 and 18), and data for a complete comparison are sorely lacking. Clearly much more detailed semantic fieldwork is required in this area.

The observation in (2) might appear rather underwhelming. After all, what appears to be essentially the same inference is often triggered by run-of-the-mill belief reports in both Spanish (3a) and English (3b).

- (3) a. Juan cree que está lloviendo.
 \rightsquigarrow *it is not raining*
 b. Juan believes that it is raining.
 \rightsquigarrow *it is not raining*

The standard account of the inferences in (3) involves the principle *Maximize Presupposition!* (hf. MP, Heim 1991b). Accordingly, our first order of business in the next section is to argue that the inference associated with RBCs and the one triggered by regular belief reports do not exhibit the same empirical characteristics. We begin with a brief exposition of the mechanics of MP in subsection 4.2.1, and in subsections 4.2.2 to 4.2.4 we discuss three ways in which RBCs depart from regular belief reports. Specifically, we will argue that (a) MP-type inferences are (in a sense) relatively easy to cancel while the inferences triggered by RBCs are surprisingly robust, (b) MP-type inferences are in general epistemically weak while the inference triggered by RBCs are epistemically strong, and finally, (c) MP-type inferences project existentially from the scope of universal quantifiers whereas the inference triggered by RBCs project universally from the same environments. We conclude that the mechanism underlying the negative inference in (2) is distinct from the one responsible for (3): while MP is potentially a good candidate for the latter, it cannot be extended the former.

Having established that the inference triggered by RBCs is not due to MP, we proceed to consider alternatives. We will entertain two hypotheses, laid out in sections 4.3 and 4.4. Both hypotheses are “presuppositional” in that both locate the source of the target inference in the definedness conditions associated with RBCs. They differ about the content of this presupposition. In section 4.3, we explore the idea that *creerse* is “contrafactive” in the sense that it presupposes that its complement clause is false. This, we will argue in subsection 4.3.1, allows us to explain the differences between RBCs and regular belief reports discussed in section 4.2, and more. However, the peculiar interpretation of RBCs when embedded under negation will force us to take the role of neg-raising seriously. In a nutshell, while unembedded RBCs trigger the inference that their complement clause is false, when embedded under negation they trigger the inference that the complement clause is *true*.⁴

- (4) Juan no **se** cree que está lloviendo.
 ‘Juan doesn’t REFL believe that it’s raining.’
 \rightsquigarrow *it is raining*

As we will argue in subsection 4.3.2, the contrafactivity analysis can only account for this “polarity reversal” effect under negation if it is coupled with the *syntactic* account

⁴There is an interaction between the inference reported in (4) and the mood of the embedded clause. Under negation, the predicate *creer* and its reflexive variant in Spanish license both indicative and subjunctive moods (Quer 2009). The inferences we are interested in arise *only* for indicative complements, not for subjunctive ones. Accordingly, we restrict the analysis to RBCs that embed indicative clauses. A similar behavior has been observed for factive and veridical predicates across Romance languages which also license both indicative and subjunctive complements under negation (Egré 2008).

of neg-raising (Collins & Postal 2014).⁵

In section 4.4, we explore an alternative analysis with the hope that it might keep all the good predictions of the contrafactivity analysis without forcing us to commit to the syntactic account of neg-raising. Adopting Gajewski’s (2007) theory of neg-raising, we stipulate that while a regular neg-raising predicate like *creer* (or *believe* in English) triggers the (soft) presupposition that the attitude-holder is opinionated with respect to the complement clause—what Gajewski, following Bartsch (1973), calls the “excluded middle presupposition”—*creerse* triggers the presupposition that the attitude-holder is *wrongly* opinionated with respect to the complement clause (we will call this the *enriched* excluded middle presupposition). While this second analysis covers much of the same ground as the first, it has two major shortcomings which are discussed in subsection 4.4.2. We conclude that the contrafactivity analysis, despite its reliance on the syntactic account of neg-raising, is nevertheless empirically superior and defer the future resolution of this dilemma to future work.

In section 4.5, we take a step back and provide a discussion of some further empirical issues. First, in subsection 4.5.1 we observe that RBCs can take other complements than propositional clauses. Interestingly, RBCs embed interrogative complements, but with two restrictions: only *wh*-questions can be embedded under *creerse* (to the exclusion of alternative and polar questions) and even this is only possible if *creerse* itself is embedded under negation.⁶ In subsection 4.5.2, we point out some differences between RBCs and their dative alternatives. Finally, in subsection 4.5.3, we report some preliminary evidence suggesting that the predicate *s’imaginer* in French has certain properties in common with RBCs. Section 4.6 concludes.

Before we move on, a cautionary remark is in order. Our attempt at analyzing the interpretation of *creerse* is non-compositional in that we merely attempt to provide an analysis for *creerse* considered as a unit. Ultimately, one would want to derive the semantics of *creerse* from an independently motivated lexical entry for *creer* plus whatever assumptions are necessary regarding the “reflexivization” process that *creer* would go through to generate *creerse*. Our hope is that this paper will contribute to this ultimate goal by providing, as a necessary first step, an adequate analysis of *creerse* as a whole.

4.2 Why not *Maximize Presupposition!*

4.2.1 Background on *Maximize Presupposition!*

Maximize Presupposition! is a principle of language-use which encodes a preference for alternatives with stronger presuppositions.⁷

⁵It is tempting to use the data point in (4) to argue that RBCs do not behave in the same way as regular belief reports, as, for example, *John does not believe that it is raining* does not normally trigger the inference that it is, in fact, raining. However, the behavior of regular belief reports when negated *vis-à-vis* the truth/falsity of their complement is complicated both empirically and theoretically and, to our knowledge, has not been discussed in any detail.

⁶Notice that the English predicate *surprise* also seems to be subject to these same two restrictions. We leaving this parallelism for future work.

⁷The principle is rooted in Heim (1991b). See Sauerland (2008a) and Schlenker (2012c) for various extensions. See Percus (2006b) and Chemla (2008a) in particular for relevant discussion on belief reports.

- (5) *Maximize Presupposition!* (hf. MP): Use the alternative that has the strongest presupposition, unless this presupposition is not known to be true.⁸

As an example, consider how MP might derive the contrast in (6).

- (6) John {#believes, knows} that Paris is in France.

Let us assume the toy semantics in (7). Note the only difference between *believe* and *know* is that the latter, being factive, triggers the presupposition that its complement is true.

- (7) a. $\llbracket \text{believe} \rrbracket^w = \lambda P_{st} \lambda x_e. \text{BEL}_x^w(P)$
 b. $\llbracket \text{know} \rrbracket^w = \lambda P_{st} \lambda x_e : P(w) = 1. \text{BEL}_x^w(P)$

According to (7a), the *believe*-sentence in (6) merely asserts something about John's doxastic state, something, furthermore, which we also know to be true. Consequently, the entry in (7a) alone is not sufficient to address the oddness of the *believe*-sentence in (6). But now suppose *believe* and *know* are competitors. It follows that an utterance of the *believe*-sentence in (6) prompts MP to compare it to the *know*-sentence as its alternative. As the *know*-sentence presupposes that Paris is in France while the *believe*-sentence is presupposition-less, and as this presupposition is known to be true, MP predicts that the *know*-sentence in (6) should block the *believe*-sentence, thereby deriving the latter's infelicity in any context in which it is common ground that Paris is in France. Assuming that contexts are sets of possible worlds that are compatible with background assumptions, MP predicts that any sentence of the form [*x believes that φ*] is felicitous in context C only if the presupposition of the *know*-alternative, i.e., [*x know that φ*], is not satisfied in C; i.e., only if there is at least one world in C in which *φ* is false.

MP as sketched so far is not entirely satisfactory. Although it yields appropriate results as far as felicity conditions go, it leaves something to be desired on the inferential side. To see this, consider the sentence in (8a). An utterance of this sentence is very likely to invite the inference that (the speaker believes that) Ann is not, in fact, 30 years old. Furthermore, in (8b), in which the first person pronoun replaces *Ann*, the parallel negative inference (that the speaker is not 30 years old) is much stronger.

- (8) a. John believes that Ann is 30 years old.
 \rightsquigarrow *Ann is not 30 years old* (weak inference)
 b. John believes that I am 30 years old.
 \rightsquigarrow *I am not 30 years old* (strong inference)

Neither of the two data points in (8) are accounted for by MP as it currently stands. This is because MP predicts, at best, that the sentences in (8) can be used only in contexts in which either the question of Ann's or the speaker's age is not common ground. But the attested inference is stronger, namely, that it is common ground that *not p*. To strengthen this inference is to perform what Chemla (2008a) calls the *epistemic step*. It essentially consists of the assumption that the speaker is opinionated

⁸It is standard to assume that MP only compares alternatives that are contextually equivalent, although this has been challenged (Spector & Sudo 2017c, Anvari 2018). This issue is not directly relevant in this paper.

with respect to the truth of the complement (i.e. she has authority over *p*). Making that assumption permits strengthening the MP-inference from *it is not common ground that p* to *it is common ground that not p*, thereby arriving at (3b). Note this account also makes sense of the contrast between the sentences in (8). It is a safe assumption that the speaker has authority about how old he is. It is much less clear whether the speaker can safely be assumed to know how old Ann is.

Finally, as exactly the same mechanism can be plausibly applied to Spanish *creer* (*believe*) and *saber* (*know*), the question arises whether *creerse* behaves like regular *creer* or whether it instantiates a different empirical profile. This is what we turn to in the following subsections. We will argue that RBC-inferences are markedly different from the MP-type inferences associated with *believe/creer*. In a nutshell, we will argue that MP-type inferences are cancellable while RBC-inferences are not (subsection 4.2.2); that MP-type inferences are epistemically weak while RBC-inferences are not (subsection 4.2.3), and that the MP-inferences project *existentially* from the scope of universal quantifiers while RBC-inferences project *universally* in the same environments (subsection 4.2.4).

4.2.2 Cancellability

The Chemla-strengthening of MP inferences is a defeasible process. In particular, it can be cancelled by a continuation like ... *and he/she is right!*, as witnessed by the felicity of (9) in both English and Spanish. Note that in (9) it is the Chemla-strengthened inference (*it is common ground that Ann is not 30 years old*) that is cancelled, not the weaker MP-inference that underlies it (*it is not common ground that Ann is 30 years old*). The latter inference is triggered in the context against which the first sentence in (9) is evaluated while Chemla-strengthening affects the context as *updated* by the first sentence.

- (9) Juan cree que Ann tiene 30 años ... y tiene razón!
 ‘Juan believes that Ann is 30 years old ... and he is right!’

The first difference between RBCs and regular belief reports is that the inference triggered by the former *cannot* be cancelled in subsequent discourse.^{9,10}

- (10) # Juan *se* cree que Ana tiene 30 años ... y tiene razón!
 ‘Juan REFL believes that Ana is 30 years old ... and he is right!’

Before we move on, recall that Chemla-strengthening becomes particularly robust

⁹As pointed out by Glass (2019) (her example 9, repeated below), the negative inference triggered by *yǐwéi* is cancellable.

- (i) rénmén yǐwéi tō shì yìwànfùwēng ... ér tā díquè shì
 person-PL yǐwéi 3sg be billionaire ... and 3sg indeed be
 ‘People are under the impression (=yǐwéi) that she’s a billionaire ... and she actually is.’

¹⁰As Hsiao (2017) observes (her example 39, repeated below), the same pattern emerges with *liah-tsun*. Note the attitude-holder is first personal. We will return to this in subsection 4.2.3.

- (i) # gua2 liah8-tsun2 il si7 huan1-a2 ... ki5-sit8 il to7 si7
 I think he COP aborigine ... actually he exactly COP
 ‘I thought that he was an aborigine ... actually he is’

when the MP inference pertains to the speaker. Thus, one way to appreciate the force of the RBC-inference is to note that even the most robust cases of Chemla-strengthening are in general cancellable, in contrast to RBC.

- (11) *Juan cree que tengo 30 años . . . y tiene razón!*
 Juan believes that I am 30 years old . . . and he is right!

4.2.3 Epistemic Strength

Consider the example in (12). The sentence, in both Spanish and English, is felicitous only in contexts in which the question of whether it is raining is not settled. This is how things should be as predicted by MP. Furthermore, due to the fact that the attitude-holder in this example is the speaker, Chemla-strengthening is automatically cancelled as it would generate the inference that (the speaker believes that) it is common ground that it is not raining which immediately contradicts the asserted content of the sentence that the speaker believes that it is raining.¹¹

- (12) *Creo que está lloviendo.*
 ‘I believe that it is raining.’

The second difference between RBCs and regular belief reports is that RBCs are incompatible with first personal attitude-holders.^{12, 13}

- (13) # *Me creo que está lloviendo.*
 ‘I REFL believe that it is raining.’

(13) is barely interpretable, in a way that is somewhat reminiscent of Moore’s paradox in (14a) and, tellingly in light of the upcoming discussion, the unmarked reading of

¹¹Given standard assumptions, if x believes that it is common ground that P then x believes that P .

¹²In some cases first personal RBCs are fairly acceptable but with a rather different meaning.

- (i) *Soy tan ingenuo que me creo todo lo que dicen en las noticias.*
 ‘I’m so naïve that I REFL believe all that they say in the news.’

We suspect that RBCs have a second reading, on top of the one mentioned in the text, which effectively attributes to the attitude-holder a certain degree of gullibility or naïveté. This reading of *creerse* is, intuitively, rather similar to the English verb *buy* in an example like *He said he was a billionaire and she bought it*. This reading is difficult to access and becomes detectable only in certain environments such as the antecedent of conditionals (see also subsection 4.5.1). We have to leave this issue to future research.

¹³As pointed out by Glass (2019) (her examples 10 and 11, repeated below), *yǐwéi* is in fact compatible with first personal attitude-holders although in such cases it either signals that “the speaker [. . .] welcomes the hearer to disagree”, (ib), or forces a past-temporal understanding of the reported belief, (ia). Neither of these two readings/implications are available for *creerse*. As the example used already in footnote 10 demonstrates, *liah-tsun* seems to pattern similarly to *yǐwéi* in this regard (see also the discussion in Hsiao (2017)).

- (i) a. wǒ yǐwéi jīntiān yǒu ge jiǎngzuò
 I yǐwéi today have DL talk
 ‘I thought there was a talk today.’
 b. wǒ gèrén yǐwéi nǐ yīnggāi zhèyàng zuò
 I personally yǐwéi you should this-way do
 ‘Personally, I think you should do this.’

(14b).

- (14) a. # It is raining but I don't believe it.
 b. # I don't know that it is raining.

It appears that the inference triggered by RBCs, similar to assertive and presupposed content of sentences in general, is strongly tied to the speaker's personal beliefs, in contrast to MP inferences which pertain to speaker's belief about the common ground. As pointed out in footnote 11, beliefs about the common ground are in general stronger than personal beliefs: if x believes P is common ground then *ipso facto* x believes P , but not *vice versa*. Therefore, when negation is taken into account the situation reverses: if x does *not* believe that P is true (in particular, if x believes that P is false) then x does *not* believe that P is common ground. This is the sense in which, we believe, the negative inference triggered by RBCs is stronger than the MP inference associated with regular belief reports.

Finally, another piece of evidence (modelled after Glass 2019) that points in the same direction is the contrast in (15).¹⁴

- (15) a. *No sé si está lloviendo o no, pero Juan cree que está lloviendo.*
 'I don't know if it's raining or not, but Juan believes that it is raining.'
 b. # *No sé si está lloviendo o no, pero Juan se cree que está lloviendo.*
 'I don't know if it's raining or not, but Juan REFL believes that it is raining.'

Again, the RBC-inference that it is not raining clashes with the first sentence, which states the speaker's ignorance.

4.2.4 Projection

MP-inferences are known to project existentially from the scope of universal quantifiers (Sauerland 2008a, a.o.). More specifically, and putting aside epistemic issues for simplicity, a plain belief report embedded under a universal quantifier such as *each student* is felicitous as long as there is at least one student who has a false belief. This is demonstrated by the felicity of the sentences in (16) in the target context.

- (16) [Context: some students passed and the rest failed.]
*Cada estudiante cree que ha aprobado.*¹⁵
 'Each student believes that she passed'.

The third difference between RBCs and regular belief reports is that the former show a different behaviour in the same environment, as witnessed by the infelicity of (17a).

¹⁴Again, things are different with *yǐwéi*, as the felicity of Glass's example 7 reported below demonstrates.

- (i) wǒ bù zhīdào yǒu-méi-yǒu défēn, dànshì zhège qiúyuán yǐwéi défēn le
 I not know have-not-have score, but this-CL ball-player *yǐwéi* score ASP
 'I don't know whether the player scored or not, but he's under the impression that (= *yǐwéi*) he did.'

¹⁵The contrast between different universal quantifiers in both English (e.g., *each*, *every* and *all*) and Spanish (e.g., *cada*, *todo* / *todos*) is not relevant for our purposes as presuppositions triggered in the scope of all of these constructions project universally. Thanks to the editors of SuB 23 proceedings for bringing this issue to our attention.

The sentence becomes fully acceptable if the context is manipulated in an appropriate manner, as in (17b).

- (17) a. [Context: some students passed and the rest failed.]
 # *Cada estudiante se cree que ha aprobado.*
 ‘Each student REFL believes that she passed.’
- b. [Context 1: every student failed, or Context 2: it is not known whether there are any students who passed.]
Cada estudiante se cree que ha aprobado.
 ‘Each student REFL believes that she passed.’

Judgments become sharper with an example like (18). Given the context in (18), the sentence involving RBCs is clearly degraded, while the alternative with *creer* is fine.

- (18) [Context: in the aftermath of a swimming match, the only possible outcome of which is that exactly one of the competitors wins.]
- a. *Cada una de las nadadoras cree que ha ganado la carrera.*
 ‘Each swimmer believes that she has won the race’.
- b. # *Cada una de las nadadoras se cree que ha ganado la carrera.*
 ‘Each swimmer REFL believes that she has won the race’.

The infelicity of (17a) and (18b) (in their respective contexts) strongly suggests that the negative inference triggered by RBCs projects universally from the scope of universal quantifiers. Thus, the oddness of (18b) can be easily explained if the projected inference has the universal force that every swimmer lost the race. The felicity of (16) and (18a) (in their respective contexts), on the other hand, is compatible with the standard assumption that MP inferences project existentially from the scope of universal quantifiers (approximately, at least one student failed for (16), and at least one swimmer lost for (18a)).

4.2.5 Interim summary

Although RBCs and regular belief reports both tend to imply that their complement clause is false, there are at least three ways in which the falsity inference of RBCs is special: it is non-cancellable, epistemically strong and it projects universally from the scope of universal quantifiers. We take these three properties to be sufficient evidence for the claim that the mechanism underlying the negative inference of RBCs is distinct from the one associated with regular belief reports. *Maximize Presupposition!* may be the appropriate mechanism as far as regular belief reports go, but it is certainly inappropriate for RBCs.¹⁶

¹⁶There are implementations of *Maximize Presupposition!* that come closer to capturing RBC-inferences. Magri (2009a) and Marty (2017a) provide implementations of MP within the framework of (grammatical) exhaustification which generate inferences that are often stronger than those predicted by standard MP. Two points merit consideration. First, the empirical behavior of RBCs and plain belief reports is so different that if exhaustification is appropriate for former then it is not appropriate for the latter. Second, an exhaustification approach to RBCs, pushed to extreme for maximum coverage, becomes essentially a notational-variant of the contrafactivity analysis discussed in the next section, and will suffer from the same problem (namely, it also would have to rely on syntactic neg-raising). It is at the moment unclear to us whether anything can be gained by switching from MP as a principle of language-use to an implementation within the exhaustification framework as far as the data in this

In the following two sections we turn to the question of just what the mechanism behind RBCs is. We will consider two hypotheses. In section 4.3 we explore the hypothesis that RBCs are “contrafactive”, in that they presuppose the falsity of the clausal complement. In section 4.4 we discuss an alternative idea according to which RBCs trigger the presupposition that the attitude holder is *wrongly* opinionated with respect to the context of the clausal complement.

4.3 The Contrafactivity Hypothesis

4.3.1 The basic idea

The contrafactivity analysis boils down to the claim that *creerse* is contrafactive. Much like how *saber* (to know) is factive, in that it presupposes the truth of its complement, *creerse* is now claimed to be contrafactive, in that it is taken to presuppose the *falsity* of its clausal complement.

$$(19) \quad \llbracket \text{creer} \rrbracket^w = \lambda P \lambda x. \text{BEL}_x^w(P)$$

$$(20) \quad \llbracket \text{saber} \rrbracket^w = \lambda P \lambda x : P(w) = 1. \text{BEL}_x^w(P)$$

$$(21) \quad \llbracket \text{creerse} \rrbracket^w = \lambda P \lambda x : P(w) = 0. \text{BEL}_x^w(P)$$

If true, the contrafactivity of RBCs is rather remarkable as, to our knowledge, no contrafactive predicate has so far been attested in the literature. Indeed, [Holton \(2017\)](#) goes as far as to claim that no such predicate exists in natural language.¹⁷ The contrafactivity analysis characterizes the RBC-inference as a presupposition. The prediction, then, is that these inferences should behave like presuppositions do in general. Using *know* as baseline, in the rest of this subsection we demonstrate that this prediction is borne out.

In subsection 4.2.2 we pointed out that RBC-inferences, unlike MP-inferences, cannot be cancelled in subsequent discourse. The same holds true of the factive presupposition of *know*.

- (22) a. # Juan knows that Ana is 30 years old . . . and he is wrong!
 b. Juan believes that Ana is 30 years old . . . and he is right!
 c. # *Juan se cree que Ana tiene 30 años . . . y tiene razón!*
 ‘Juan REFL believes that Ana is 30 years old . . . and he is right!’

In subsection 4.2.3 we argued that the epistemic strength of RBC-inferences is stronger than that of MP-inferences. Again, the same is true for the factive presupposition of *know*. Looking at (23a) in particular, the speaker cannot use a sentence that carries the presupposition *P* if he or she is ignorant about whether *P* is true. This observation applies to *know* (with the presupposition that it is raining) exactly like it applies to *creerse* (with the presupposition that it is not raining) under the contrafactivity analysis.

paper are concerned, although the issue needs to be considered more carefully.

¹⁷Holton does in fact mention RBC in a footnote ([Holton 2017](#): p.250, n.3), but he does not elaborate further, reporting that his consultants were unsure about the exact content of the inference. *Liah-tsun* as analysed by [Hsiao \(2017\)](#) is certainly a candidate as well. But note that, the hall-mark of presuppositions being their projection profile, the relevant facts for *liah-tsun* are not yet known.

- (23) a. # I don't know that it is raining.
 b. I believe that it is raining.
 c. # *Me creo que está lloviendo.*
 'I REFL believe that it is raining.'
- (24) a. # I don't know whether it is raining or not, but John knows that it is raining.
 b. I don't know whether it is raining or not, but John believes that it is raining.
 c. # *No sé si está lloviendo o no, pero Juan se cree que está lloviendo.*
 'I don't know if it's raining or not, but Juan REFL believes that it is raining.'

In subsection 4.2.4 we pointed out that RBC-inferences project universally from the scope of universal quantifiers. This behavior is, of course, the hall-mark of presuppositions.

- (25) *Each student knows that he passed the exam.*
 ~> *every student passed the exam*
- (26) *Cada estudiante se cree que ha aprobado.*
 'Each student REFL believes that she passed.'
 ~> *every student failed the exam*

Not only can the contrafactivity analysis account for the data so far discussed, it also makes further predictions that happen to be true. Specifically, presuppositions are known to project in polar questions and from the scope of existential modals. The same holds for RBC inferences as well.

- (27) a. Does John know that it is raining?
 ~> *it is raining*
 b. *¿Se cree Juan que está lloviendo?*
 'Does Juan REFL believe that it is raining?'
 ~> *it is not raining*
- (28) a. John might know that it is raining.
 ~> *it is raining*
 b. *Puede que Juan se crea que está lloviendo.*
 'might that Juan REFL believe that it is raining.'
 ~> *it is not raining*

We take these observations to lend strong *prima facie* support to the contrafactive analysis.

4.3.2 The Polarity Reversal Problem

Before uncorking the champagne and toasting to contrafactivity however, the analysis faces a simple and possibly devastating difficulty. As pointed out briefly in section 4.1, the polarity of RBC inferences reverses under negation.¹⁸

- (29) a. *Juan se cree que está lloviendo.*

¹⁸As pointed out by Hsiao (2017) (her example 34, repeated below), *liah-tsun* cannot be negated. This is, perhaps, the most crucial difference between *liah-tsun* and *creerse*.

- ‘Juan REFL believes that it’s raining.’
 \rightsquigarrow *it is not raining*
- b. *Juan no se cree que est  lloviendo.*
 ‘Juan doesn’t REFL believe that it’s raining.’
 \rightsquigarrow *it is raining*

The problem that (29b) raises for the contrafactivity analysis is plain: just like polar questions and existential modals, negation is a presupposition hole: any presupposition triggered in the scope of negation should either project or be locally accommodated. Neither of these two possibilities seem to account for (29b).

- (30) The contrafactivity analysis predicts (29b) to mean either . . .
- a. *it is not raining and it is not the case that Juan believes that it is raining* (if the contrafactive presupposition projects) or . . .
 - b. *if it is not raining, then Juan believes that it is raining* (if the contrafactive presupposition is locally accommodated)

Similar results are obtained when we embed RBCs under a negative universal quantifier.

- (31) *Ning n estudiante se cree que ha suspendido.*
 No student REFL believes that she has failed.
 \rightsquigarrow *every student has failed.*

Let us call this the Polarity Reversal (hf. PR) Problem. How serious is the PR problem? Recall that the contrafactive account made several good predictions. It seems a worthwhile enterprise to try to rescue it somehow. To do that, we will explore an intuitive subsidiary hypothesis: given that non-reflexive *creer* is a neg-raising predicate, perhaps *creerse* is also neg-raising and the PR problem is only a problem as long as neg-raising is ignored.

It is well-known that doxastic predicates like *believe* and *creer* are neg-raising predicates, in the sense that under negation they give rise to interpretations that are stronger than predicted (Horn 1978). Thus the inferences from (32a) to (32b) seems intuitively valid although the latter, given otherwise standard assumptions about negation and the meaning of *believe*, should also be compatible with John being totally ignorant about whether it is raining or not.

- (32) a. *John doesn’t believe that it’s raining.*
 b. *John believe that it’s not raining.*

On the syntactic account of neg-raising (Collins & Postal 2014) in (32a) the negation is actually interpreted at LF in the embedded clause but raises to the matrix clause where it is pronounced. In a nutshell, two copies of negation are present, the higher is pronounced and the lower is interpreted.

- (33) LF of (32a): John NEG believe that [it is NEG raining]

- (i) # il bo5 liah8-tsun2 a1-ing1 tsa1-hng1 kah4 ong5-sian1-sinn1 tso3-hue2.
 he neg think A-ing yesterday with Wang-Mr. be.together
 ‘He didn’t mistakenly think that A-ing was with Mr. Wang yesterday.’

Now, what if this mechanism is also available for *creerse*?

- (34) Juan no (se) cree que está lloviendo
 LF: Juan NEG (se) cree que [NEG [está lloviendo]]

This assumption coupled with the claim that *creerse* is contrafactive, now predicts the polarity reversal behavior noted above.

- (35) *Juan no se cree que está lloviendo.*
 Assertion: Juan believes that it is not raining
 Presupposition: It is not the case that it is not raining

Thus the syntactic account of neg-raising can solve the contrafactivity analysis' PR problem. There are, however, several arguments against the syntactic account of neg-raising (see Romoli 2013 in particular and Collins 2019 for a recent argument, and references therein) and it seems prudent to ask whether the PR problem really forces the contrafactivity analysis to commit to the syntactic account.

On the other hand, there are several *semantic* analyses of neg-raising available (the homogeneity-based account of Križ 2015, the scalar implicature-based account of Romoli 2013, the presuppositional account of Gajewski 2007). As far as we can see, none of the semantic accounts can solve the PR problem. The reason, in a nutshell, is that in order to solve the PR problem via neg-raising a connection must be established between the doxastic state of the attitude-holder and the truth/falsity of the complement clause. The syntactic account accomplishes this by interpreting the negation in the scope of the attitude predicate, thereby enriching the attitude-holder's reported doxastic state while at the same time reversing the polarity of the contrafactive presupposition. The semantic accounts, on the other hand, fail to establish this connection. We illustrate this with Gajewski's presuppositional account.

Abstracting from orthogonal concerns, Gajewski's proposal boils down to the claim that *believe* (and *creer*) triggers the presupposition that the attitude-holder is opinionated with respect to the complement clause.

- (36) $\llbracket \text{creer/believe} \rrbracket^w = \lambda P \lambda x : \text{BEL}_x^w(P) \vee \text{BEL}_x^w(\neg P). \text{BEL}_x^w(P)$

Nothing changes in the unembedded case. But under negation, the excluded middle presupposition projects thereby strengthening the truth-conditions of the sentence.

- (37) *John not believe that it is raining.*
 PRESUPPOSITION: $\text{BEL}_j^w(r) \vee \text{BEL}_j^w(\neg r)$
 ASSERTION: $\neg \text{BEL}_j^w(r)$
 \Rightarrow true iff $(\text{BEL}_j^w(r) \vee \text{BEL}_j^w(\neg r)) \wedge \neg \text{BEL}_j^w(r)$ iff $\text{BEL}_j^w(\neg r)$

Now suppose we superimpose the excluded middle presupposition on the contrafactive presupposition.

- (38) $\llbracket \text{creerse} \rrbracket^w = \lambda P \lambda x : P(w) = 0 \wedge (\text{BEL}_x^w(P) \vee \text{BEL}_x^w(\neg P)). \text{BEL}_x^w(P)$

The prediction is clearly inadequate. As made explicit below, we effectively predict (39) to be true iff Juan knows that it is not raining.

- (39) *Juan no se cree está lloviendo.*
 LF: Juan NEG se-believe [it is raining]

PRESUPPOSITION:

1. $BEL_j^w(r) \vee BEL_j^w(\neg r)$
2. $r(w) = 0$

ASSERTION: $\neg BEL_j^w(r)$

\Rightarrow true iff $r(w) = 0 \wedge (BEL_j^w(r) \vee BEL_j^w(\neg r)) \wedge \neg BEL_j^w(r)$ iff $r(w) = 0 \wedge BEL_j^w(\neg r)$

It seems, then, that the contrafactivity hypothesis only works if it is supplemented with the syntactic account of neg-raising. In the next section we explore an alternative analysis which does remove the reliance on the syntactic account, but is empirically less successful than the contrafactivity analysis.

4.4 The Enriched Excluded Middle Hypothesis

4.4.1 The basic idea

According to Gajewski's account, regular belief reports trigger the (soft) presupposition that the attitude-holder is opinionated with respect to the embedded proposition. We would like to suggest that the special ingredient of *creerse* is a richer excluded middle presupposition. Specifically, the idea is that *creerse* triggers the presupposition that the attitude-holder is *wrongly* opinionated with respect to the embedded proposition: if the proposition is true the attitude-holder believes it to be false and if the proposition is false then the attitude-holder believes it to be true.

- (40) a. $\llbracket \text{creer/believe} \rrbracket^w = \lambda P \lambda x : BEL_x^w(P) \vee BEL_x^w(\neg P). BEL_x^w(P)$
b. $\llbracket \text{creerse} \rrbracket^w = \lambda P \lambda x : (P(w) = 0 \wedge BEL_x^w(P)) \vee (P(w) = 1 \wedge BEL_x^w(\neg P)). BEL_x^w(P)$

(40b) predicts that RBCs, when unembedded, are true iff the complement clause is false and the attitude-holder believes that it is true. Furthermore, under negation the enriched excluded middle presupposition will project. We then predict the sentence to be true iff the complement clause is true and the attitude-holder believes that it is false. Thus the enriched excluded middle analysis seems to make correct the prediction for the basic cases and the cases involving embedding under negation. The reader can verify that an analysis along the lines of (40b) will also derive the projection facts, discussed in subsection 4.2.4, and the fact that RBC inferences are epistemically strong, as discussed in 4.2.3.

We would, however, like to point out two shortcomings of this analysis, one possibly minor and one major. We discuss the latter in the next subsection. The minor difficulty is that neg-raising inferences are known to be defeasible. In Gajewski's framework this means that the original excluded middle inference is a *soft* presupposition. If the enriched excluded middle that we have postulated for *creerse* is in the same category then we cannot explain why RBC inferences cannot be cancelled in subsequent discourse, as discussed in 4.2.2. Some justification, then, is needed to explain why the enriched excluded middle inference is a *strong* presupposition even though the excluded middle itself is a soft presupposition.

4.4.2 The problem with Polar Questions

The major difficulty pertains to projection. As pointed out in at the end of section 4.3.1, RBC inferences project from polar questions.

- (41) *¿Se cree Juan que está lloviendo?*
 ‘Does Juan REFL believe that it is raining?’
 ↗ *it is not raining*

This fact cannot be explained by (40b). At best, we predict the project inference to be that the attitude-holder is wrongly opinionated with respect to the embedded proposition. In fact, it remains a mystery why this inference itself surfaces in the enriched form that we have observed.

4.4.3 Interim summary

The theoretical situation is perplexing. We have an analysis which is empirically rather successful, namely the contrafactivity analysis, but necessitates commitment to the syntactic account of neg-raising, which we think is not desirable. On the other hand, we have sketched an alternative analysis which does not rely on the syntactic account but at the cost of empirical coverage. We leave the ultimate resolution of this dilemma to future work.

4.5 Some Further Empirical Observations

4.5.1 Question embedding and other complements

The data discussed above involve cases where *creerse* embeds propositions. Crucially, *creerse* can also take complements that plain *creer* does not admit; namely, interrogatives and small clauses. We briefly discuss these in turn.

Let us start by considering what is maybe the most surprising of these observations: when they are negated, RBCs can embed *wh*-interrogatives.

- (42) *Juan no se cree {quién vino, dónde es el concierto, qué hay que hacer para entrar}.*
 Juan not REFL believe {who came, where is the concert, what one has to do to get in}.

As illustrated in (43a), the ability to embed interrogatives is only available under negation. Furthermore, as illustrated in (43b) and (43c), even when *creerse* is embedded under negation it can only embed *wh*-questions to the exclusion of polar and alternative questions.

- (43) a. **Juan se cree {quién vino, dónde es el concierto, qué hay que hacer para entrar}.*
 b. **Juan no se cree si está lloviendo.*
 Juan doesn’t REFL believe whether it’s raining.
 c. **Juan no se cree si el auto es rojo o azul.*
 Juan doesn’t se believe whether the car is red or blue

Now, what does, e.g., the *who* case of (42) actually mean? Intuitively, the sentence is true iff there is at least one person *x* who in fact came but Juan believes that *x* did not come. Thus, consider a context where Mary and Ann came, and Bill did not. Sentence (42) would then be true as long as John either believes that Mary did not

come or he believes that Ann did not come. Importantly however, having a false belief regarding a *negative* answer to the question does not license the construction. Thus if Juan believes that Mary, Ann *and* Bill came the sentence is not acceptable.¹⁹

We should additionally note that the use of an embedded interrogative requires the answer to the embedded question to have been recently introduced into the common ground. Intuitively, (42) can only be used when someone has told Juan who came, but he does not believe it and has different beliefs about it.²⁰ It's not surprising then that *dative* belief constructions in Spanish, obtained by adjoining the dative pronoun to the predicate *creer*, pattern with RBCs in this respect:

- (44) *Juan no me cree quién vino.*
'Juan doesn't 1sg:DAT believe who came.'

(44) seems to have analogous truth conditions to (42), modulo the fact that the former sentence implies that it's the speaker who has provided an answer to the question.

Besides interrogatives, *creerse* can also embed small clauses, where the reflexive pronoun the subject of the small clause, as in (45). This alternative is also attested for reflexive belief constructions in other Romance languages, such as French and Italian.

- (45) *María se cree muy mala.*
'María REFL believes herself to be very bad.'
↪ *María falsely believes that she is very bad.*

While superficially similar, the inference triggered by (45) is weaker than the ones analyzed so far: it can be cancelled —it allows the continuation *and she is right!*—, and does not arise under negation.

Given that, in these cases, the reflexive is presumably generated in the embedded clause (as subject) and only raises to the matrix clause to receive case (i.e. raising to object, Chomsky 1993a), we consider these constructions plain belief reports, and the corresponding inference is just an MP-type inference.

4.5.2 Dative belief construction

As observed above, the predicate *creer* can also take a dative pronoun as complement. Could RBC be thought of as a special case of this *Dative Belief Construction* (henceforth, DBC), also existent in English?²¹

- (46) a. Juan believed you that it was raining.
b. *Juan te creyó que estaba lloviendo.*
'Juan 2sg:DAT believed that it was raining'

The DBC in (46) is typically taken to mean *believed you when you tell him* (Chemla 2008a). The RBC in Spanish could then be analyzed by saying that the *inducir*

¹⁹The facts pertaining to the ability of *creerse* to embed *wh*-questions exclusively only when it is embedded under negation are potentially significant in the context of recent attempts in deriving the selectional restrictions of various predicates from semantics assumptions (Mayr 2017 and Theiler et al. 2019, a.o.). For space limitations we have to leave this to future work.

²⁰This suggests that perhaps in such constructions the second meaning of *creerse* (see footnote 12) is the culprit.

²¹The claims made here about the inferential pattern of Spanish DBC should not be extended to English. English examples are only used for clarity purposes.

and receiver of the belief are the same person. Sentence (1), repeated below, would then be interpreted as *Juan believes himself when he tells himself that it's raining*.

- (47) *Juan se cree que está lloviendo.*
 'Juan REFL believes that it is raining.'

Treating the RBC as a special case of DBC would give us the possibility of expanding our predictions to other uses of the predicate *believe* in Spanish and potentially other languages. However, Spanish RBCs and DBCs display important differences. To begin with, unlike our original RBC example in (1), repeated above, unembedded DBCs in Spanish are quite deviant in present tense:

- (48) ?? *Juan te cree que está lloviendo.*
 'Juan 2sg:DAT believes that it is raining'

Present tense DBCs are however fully acceptable when embedded under negation and in polar questions.

- (49) a. *Juan no te cree que está lloviendo.*
 'Juan doesn't 2sg:DAT believe that it is raining'
 b. *¿Te cree Juan que está lloviendo?.*
 'Does Juan 2sg:DAT believe that it is raining?'

Negative sentences such as (49a) trigger the inference that it is, in fact, raining. Crucially, this inference does not arise from the polar question in (49b), indicating that it is not as strong as the RBC-inference.

Dative and reflexive constructions also differ in their ability of taking non-human subjects. This can be taken to be a direct consequence of the fact that DBCs have a speech-act ingredient (believing *p* as a result of *being told p*):

- (50) *El perro no se/#te cree que está lloviendo.*
 'The dog does not REFL/2sg:DAT believe that it is raining.'

Last but not least, the reflexive pronoun in Spanish is known to have other non-argumental uses (cf. footnote 1). In particular, the reflexive can be attached to other attitudinal predicates, which cannot take dative arguments. This is the case of the predicate *pensar* (*think*):²²

- (51) *Juan se piensa que está lloviendo.*
 'Juan REFL thinks that it is raining.'
 ~> *it's not, in fact, raining*

Crucially, (51) gives rise to an inferential pattern analogous to our RBC, at least in the positive form. Indeed, *creerse* and *pensarse* have been treated as alternatives to each other (DiTullio 2018). The construction with *pensar* was left outside the scope of this paper because it presents some minor differences with RBC, specifically under

²² We should note that Spanish also has a reflexive variant of the factive predicate *saber* (*to know*). Reflexive *saber* can only take nominal complements not propositional ones:

(i) *Juan se sabe la lección*
 Juan REFL knows the lesson.

The reflexive in these cases functions as a telicity marker: the sentence above implies that he knows the lesson *completely*.

negation. However, its existence suggests that providing a semantic account of the “reflexivization” process is an urgent *desideratum*. We hope to address this issue in future work.

4.5.3 Cross-linguistic overview: The case of French *s’imaginer*

We have tackled the semantics for *creerse* as a unit. The presence of a contrafactive presupposition, however, seems to be tightly linked to the “reflexivization” process that allows deriving *creerse* from *creer* —and possibly *pensarse* from *pensar*. Indeed, it was brought to our attention that French also displays a RBC, built by adjoining the reflexive pronoun to the predicate *imaginer* (*imagine*; p.c. Paul EgrÃ¶). As illustrated in (52), French *s’imaginer* gives rise to similar inferences to Spanish *creerse*.

- (52) a. *Jean s’imagine qu’il pleut.*
 ‘Jean REFL thinks it’s raining.’
 \rightsquigarrow *it is not raining.*
- b. *Jean ne s’imagine pas qu’il pleut.*
 ‘Jean doesn’t REFL think it’s raining.’
 \rightsquigarrow *it is raining.*
- c. *#Je m’imagine qu’il pleut.*
 ‘I REFL think it’s raining.’
- d. *#Jean s’imagine que j’ai une soeur et il a raison!*
 ‘Jean REFL thinks I have a sister, and he’s right!’

These examples provide further evidence suggesting that the contrafactive inference arises, at least partially, as a result of adjoining the reflexive pronoun: in Spanish and French, *different* predicates give rise to similar inferences in their reflexive variant.

However, “reflexivizing” an attitudinal predicate is not enough to make it contrafactive: as observed, many predicates that have reflexive variants do not give rise to a contrafactive inference. Besides the already mentioned *saberse* —which cannot take propositional complements—, it’s worth mentioning that Spanish *imaginar* (*imagine*) also has a reflexive alternative *imaginarse*. Unlike *s’imaginer* in French, Spanish *imaginarse* is equivalent to plain *imaginar* —it triggers MP-type inferences.

- (53) a. *Me imagino que est lloviendo.*
 ‘I REFL think it’s raining.’
- b. *Juan se imagina que tengo una hermana, y tiene razn!*
 ‘Juan REFL thinks I have a sister, and he’s right!’

A compositional account of contrafactivity would then require modelling the semantic import of the reflexive pronoun as a function of the predicate to which is attached: it is adjoining the reflexive pronoun to only *certain* predicates which results in the addition of a contrafactive presupposition. We believe the data presented here is too sparse to draw a sound generalization about the semantic import of the reflexive.

4.6 Conclusions

This article contributes to the research on belief reports by bringing what appears to be a ‘contrafactive’ predicate into the picture. We have investigated the puzzling inferential behaviour of Reflexive Belief Constructions in Spanish. We have shown that RBC-inferences cannot be accounted for by standard treatments of belief reports, in terms of MP-inferences. Instead, we explored two hypotheses: that *creerse* is contrafactive and that *creerse* carries an *enriched* excluded middle presupposition. The latter has wider empirical coverage but succeeds only if it is supplemented with the syntactic account of neg-raising. The latter does not rely on syntactic neg-raising, but has narrower empirical coverage. Future research needs to address this dilemma.

For the sake of simplicity, and as a *necessary* first step, we have provided here a non-compositional account of RBC: we gave a meaning to the predicate *creerse* as a whole, without assessing the semantic import of the reflexive pronoun. As observed in section 4.5, a more sophisticated, compositional account of RBCs needs to be made to account for both cross-linguistic and within-language data.

Part II

Part Two: Context *qua* Situation of Utterance

Chapter 5

The Ban Against Illeism and Indexical Shift in Farsi

5.1 Introduction

Illeism is the act of referring to oneself, and, by extension, to the addressee, in the third person. More specifically, illeism is constituted by any utterance of a sentence containing a noun phrase (hf. NP) that is intended to refer to the speaker or the addressee but is grammatically marked as third person, as exemplified in (1) and (2).^{1,2}

- (1) a. #**[Jimmy_i:]** {Jimmy, he_i} is hungry.
b. #**[Jimmy_i:]** Elaine said that {Jimmy, he_i} is hungry.
- (2) a. #**[Jimmy to Jerry_i:]** {Jerry, he_i} is an idiot.
b. #**[Jimmy to Jerry_i:]** Elaine said that {Jerry, he_i} is an idiot.

Illeism defined as such is in general infelicitous, at least in quotidian linguistic exchange.³ Therefore, a generalization along the lines of (3) seems to be called for. (Some apparent counter-examples to (3) are discussed in Appendix .1. None of those cases threaten the substance of the discussion that follows, as far as I can see.)

¹In a framework in which the third person is not grammatically / syntactically represented, the last sentence should be rephrased as follows: illeism is constituted by any utterance of a sentence containing a noun phrase that is intended to refer to the speaker or the addressee but is *not* grammatically marked as first or second person respectively. For simplicity, I will assume that the third person is grammatically represented. Nothing hinges on this assumption, but see fn. 6, 18 and 27.

²On notation: for any proper noun N (and M) and string ϕ , the notation “**[N_i (to M_j):]** ϕ ” is meant to signify an utterance of the string ϕ by N (to M) evaluated against only those assignment functions that map the index i to N (and j to M). Note that there is a difference between “# **[N_i (to M_j):]** ϕ ” and “**[N_i (to M_j):]** # ϕ ”; the former encodes the claim that any utterance of ϕ by N (to M) is infelicitous while the latter encodes the claim that any utterance of ϕ (by anyone (to anyone)) is infelicitous.

³In literary usage illeism is possible and generates rhetorical effects of a rather elusive kind, such as a larger-than-life effect. One (rather literal) example of this is due to Charles de Gaulle who in the first volume of his memoirs comments on the 1962 assassination attempt on his life, writing “of the 150-odd bullets aimed at us, 14 strike our vehicle. Yet none of us is hit. May De Gaulle therefore go on pursuing his road and his vocation!” (taken from [here](#), where other examples are recorded; thanks to François Recanati (p.c.) for bringing this example to my attention). Note the fact that in literary contexts illeism leads to secondary inferences or rhetorical effects is further evidence for the presence of some sort of pressure against nonchalant usage of it. See also the discussion on ‘imposters’ in Appendix .1.

- (3) Using third person NPs to refer to the speaker or the addressee of the utterance is unacceptable.

(3) is merely a generalization. What principle(s) does it follow from? A plausible candidate is a principle which encodes a *ceteris paribus*-preference for indexical (i.e., first and second person) pronouns over third person NPs. More specifically, one can stipulate a rule according to which whenever a third person NP can be replaced by an indexical pronoun without changing the meaning of the utterance, using the third person NP is illicit.

- (4) **Ban Against Illeism.** (hf. BAI) Let ϕ and ϕ' be two sentences such that the only difference between them is that a third person NP in ϕ is replaced by an indexical pronoun in ϕ' . In any context in which ϕ and ϕ' are contextually equivalent ϕ is unacceptable.^{4,5} (cf. (62) below)

Consider again (1a), repeated below. This example is now predicted by BAI to be blocked by (5b). Clearly the sentence in (5a) conveys the same proposition as (5b) in any context in which Jimmy is the speaker, as is the case here by assumption. Therefore (5b) is preferred to (5a) by BAI.

- (5) a. #**[Jimmy_i:]** {Jimmy, he_i} is hungry.
 b. **[Jimmy:]** I am hungry.

In the above formulation of BAI, I do not mean to imply that BAI is a *sui generis* principle. One very plausible hypothesis is that BAI itself should be derived from the *Maximize Presupposition!* principle (Heim 1983a and much subsequent work) applied to the domain of ϕ -features. I refrain from taking this step because the discussion that follows is largely neutral regarding the mechanism underlying BAI (but see the discussion in the concluding Section 5.9).⁶

In English, BAI rules out illeism *en bloc*: it predicts there to be no *acceptable* case of illeism (in non-literary contexts, see fn. 3). To see this, note that (quotation aside) in

⁴The exact definition of contextual equivalence is not crucial here. I assume the standard definition according to which two propositions p and q are contextually equivalent in C iff p and q are true in the same C -words.

⁵I claim no originality in proposing this rule. Indeed, a principle or generalization along these lines has been assumed by many researchers. Podobryaev's (2014) *Elsewhere 3rd person* rule is just one example.

⁶Another possibility is to assume that the third person feature is also indexical but in a way that is complementary to the first and second persons, as in (i) (ignoring gender and number). This is essentially a lexicalist alternative to the competition-based rule in (4).

- (i) $\llbracket \text{he}_x \rrbracket^{c,g,\dots} = \#$ if $g(x)$ is the author or the addressee of c ; if defined, $\llbracket \text{he}_x \rrbracket^{c,g,\dots} = g(x)$.

The lexicalist analysis suffers from two problems. First, it requires the third person feature to be syntactically represented, which is contested (e.g., Harley & Ritter 2002 a.o.). Second, it incorrectly predicts that a sentence like *every student in this class thinks she is smart*, on a reading where the pronoun is bound, should trigger the inference that the speaker and the addressee are not students in the relevant class, with the auxiliary result that the sentence *every student in this class thinks she is smart, including me* is incorrectly predicted to be infelicitous. The mentioned inference would result from a universal projection of the presupposition introduced by the third person feature on the bound pronoun on this analysis, analogous the attested inference of this sentence that the class consists only of female students, which results from the universal projection of the presupposition introduced by the gender feature on the same pronoun (note that *every student in this class thinks she is smart, including me* is indeed infelicitous if the speaker is male). (See also fn. 18.)

English indexical pronouns *always* take their value from the *actual* context of speech. Consequently, any use of a third person NP that refers either to the speaker or the addressee is predicted to be blocked by BAI. However, if BAI is on the right track then we should expect more nuanced predictions cross-linguistically. Specifically, we should expect an interaction between (in-)felicity of illeism and availability of *indexical shift* (see Deal 2017 for a recent survey and references therein). This is because in such languages indexical pronouns, in indirect discourse, may refer to the *reported* speaker / addressee instead of the *actual* speaker / addressee. If one manages to construct examples in which ‘indexical shift’ is obligatory, that is, environments in which indexical pronouns are forced to shift, then one would expect on the basis of BAI that in such environments third person reference to the actual speaker / addressee should be possible as the competitors (i.e., indexical pronouns) are clearly not suited for the job in such environments.

In Section 5.2, I lay out the basic facts pertaining to indexical shift building on data from Farsi and, in Section 5.3, I sketch the operator-based account of indexical shift (Anand & Nevins 2004; Anand 2006; Deal 2017). Against this background, in Section 5.4, I use the SHIFT TOGETHER constraint on indexical shift (Anand & Nevins 2004; Anand 2006) to construct cases in which indexical shift is obligatory and I provide evidence that the prediction made by BAI is indeed born out: illeism is acceptable in context-shifted environments. In Section 5.5, I use an interaction between indexical shift and question-embedding to further corroborate this prediction. In Section 5.6, I discuss a generalization, motivated on the basis of Farsi data, according to which third person reference to the *reported* speaker / addressee is unacceptable in a context-shifted environment. I will explain why this second generalization is not immediately captured by the BAI and I will suggest a modification to BAI which allows it to derive this generalization as well. According to this modification, BAI compares NPs for covaluation (not sentences for equivalence) and the relevant notion of covaluation is one which is ‘blind’ to the *de re* / *de se* distinction (‘type-II covaluation’ of Sharvit (2010)). In Section 5.7, I use this paradigm to compare the operator-based account of indexical shift to the binding-based account of indexical shift (Schlenker 1999, 2003; von Stechow 2004). My conclusion will be that the former has an advantage over the latter as far as the present data-set is concerned. In Section 5.8, I provide evidence that the pattern discussed *vis-à-vis* individual-denoting NPs generalizes to temporal and locative adverbials; for example, *Monday* and *today* compete in the same way that *Jimmy* and *I* do. Section 5.9 concludes the paper. In the remainder of this section I will summarize the rest of the paper, leaving blanks to be filled in the following sections.

As just mentioned, in languages that allow indexical shift, indexical pronouns embedded in *indirect* discourse may take their value from the *actual* or *reported* context of speech. For instance, consider the schematic example in (6) from Farsi (Indo-European, also referred to as Persian; see Section 5.2 for detailed discussion of indexical shift in this language).

- (6) [Sajjad to Qazal:] Leila told Mina that I hate you. (Lit. from Farsi)

In one reading of (6) the indexicals take their value from the *actual* context of speech (like English) and refer to Sajjad and Qazal respectively. In the other reading the indexicals take their value from the *reported* context of speech (unlike English) and refer to Leila and Mina respectively.

- (7) *Possible readings for (6):*
- a. Leila told Mina that Sajjad hates Qazal.
 - b. Leila_i told Mina_j that she_i hates her_j.

One crucial empirical characteristic of indexical shift is the **SHIFT TOGETHER** constraint (Anand & Nevins 2004; Anand 2006) according to which all indexicals within the same minimal domain (the embedded clause in (6)) take their value from the same context. Thus (6) does not allow for a reading in which the first (second) person pronoun is shifted but the second (resp. first) person pronoun is not.

- (8) *Impossible readings for (6):*
- a. Leila told Mina_j that Sajjad hates her_j.
 - b. Leila_i told Mina that she_i hates Qazal.

Theoretically, one can account for indexical shift by stipulating a covert ‘context-shifting’ operator (Anand & Nevins 2004; Anand 2006) which is licensed by only certain attitude predicates (such as speech-report predicates) and, if present, takes scope over the whole embedded clause. If this operator is present, (9a), then every indexical it c-commands is shifted and if it is absent, (9b), then no indexical is shifted (for the detailed semantics of this operator see Section 5.3).

- (9) **[Sajjad to Qazal:]**
- a. Leila told Mina [OP [I hate you]]
‘Leila_i told Mina_j that she_i hates her_j’
 - b. Leila told Mina [∅ [I hate you]]
‘Leila told Mina that Sajjad hates Qazal’

Using **SHIFT TOGETHER** one can guarantee the syntactic presence of **OP** and one can thereby construct environments in which indexical shift is *obligatory*. This type of environment in turn allows testing the predictions of **BAI**. Consider the schematic construction in (10), where the embedded third person NP refers to the actual speaker *A* and the second person pronoun is shifted to refer to reported addressee *y*.

- (10) **[A to B:]** x told y that [... [3rd NP] ... [2nd-pro] ...]
-

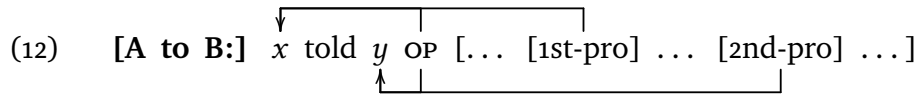
From the latter assumption, coupled with the **SHIFT TOGETHER**, we can conclude that the context-shifting operator is syntactically present and the LF underlying (10) looks something like (11).⁷

- (11) **[A to B:]** x told y OP [... [3rd NP] ... [2nd-pro] ...]
-

Thus the third person NP in (11) is in the scope of the context-shifting operator. Therefore, replacing the third person NP with a first person pronoun results in a structure like (12) in which the first person pronoun is also in the scope of the

⁷A more accurate LF for (10) would at least take into account the fact that the embedded third person NP would refer to the actual speaker *de re*. Here we can afford to ignore this complication because what ultimately matters is that the interpretation of (11) is distinct from that of (12). As discussed below, the same luxury cannot be afforded for cases where indexical shift feeds **BAI**.

context-shifting operator; it follows that, much like the second person pronoun, the introduced first person pronoun *must* shift.

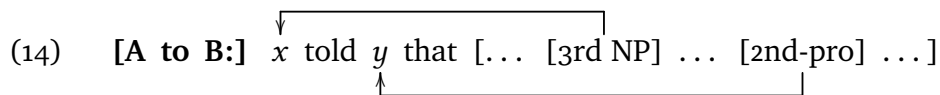


(11) and (12) are anything but synonymous / semantically equivalent. Intuitively, the embedded clause in (11) predicates something of the actual speaker A while the embedded clause in (12) predicates something of the reported speaker x .⁸ As semantic equivalence of the competitors is a necessary condition for BAI to get activated, it does not predict (12) to block (11). The schema in (10), therefore, gives us a recipe to construct cases of *acceptable* illeism in languages that allow indexical shift (and obey SHIFT TOGETHER). The resulting prediction is given in (13) and Section 5.4 is dedicated to the corroboration of this and related predictions. The reader should note that the prediction in (13) follows immediately from the competition-based principle given in (4) and the SHIFT TOGETHER constraint on indexical shift; no auxiliary assumptions need to be made.

(13) **Indexical Shift can bleed BAI.** In an environment in which indexical shift is obligatory, a third person NP *may* be used to refer to the *actual* speaker.

To begin with, in Section 5.4 prediction (13) is verified using fact from Farsi. I then turn to examples in which indexical shift is *optional*, in the sense that there is no independent reason to assume that the context-shifting operator is present in the structure, and I will argue that indexical shift can be used in those examples to bleed BAI even though it does not have any detectable interpretive consequences, i.e., is semantically vacuous. One result of this observation is that it gives us a way to make sure that the context-shifting operator is syntactically present even if its presence cannot be semantically detected. In Section 5.5 I turn to how indexical shift interacts with question-embedding. I point out that in Farsi the (responsive) predicate *goftan* (to say / tell) does not allow indexical shift if it embeds a question. I will not feign a positive hypothesis regarding why this is the case here, but, building on the discussion in Section 5.4, I will argue that the reason is most likely syntactic in nature. I will do this by constructing examples in which *goftan* cannot embed a question while the context-shifting operator is syntactically present but semantically vacuous.

Granted that indexical shift can bleed BAI, can it also *feed* it? Consider the schematic construction in (14). The only difference between (11) and (15) is that in the latter the embedded third person NP refers to the *reported*, instead of actual, speaker.



Again, because the second person pronoun is assumed to refer to the reported ad-

⁸In principle it is possible that A and x are / denote the same individual, but this either leads to an unacceptable LF or a case in which indexical shift feeds BAI, as discussed below. To see this note that if x is marked as third person then BAI itself guarantees infelicity. If x is the first person pronoun then the nature of the embedded third person NP becomes relevant. If this NP is a proper noun then we get a Condition C violation (because the third person NP is c-commanded by a co-referring pronoun, i.e., x) and we predict unacceptability. The case where the third person NP is a pronoun effectively reduces to the schema in (14) discussed below.

dressee, the context-shifting operator must be present. The LF underlying (14), then, approximates (15).

(15) [A to B:] x told y OP [... [3rd NP] ... [2nd-pro] ...]

Now suppose we replace the third person NP with a token of the first person pronoun. This takes us from (15) to its BAI-competitor (16).

(16) [A to B:] x told y OP [... [1st-pro] ... [2nd-pro] ...]

As a brief inspection of the arrows in (15) and (16) makes clear, this time the embedded clauses in the competing LFs (15) and (16) *seem* to predicate something of the same individual, namely the reported speaker x . Consequently, one might think, the competing LFs are semantically equivalent and BAI predicts (16) to block (15). In other words, (17) below appears to be a prediction of the present framework. This, however, is not technically the case. While there is indeed a sense in which the embedded clauses in (15) and (16) are both *about* the reported speaker, the vehicles used to refer to the reported speaker utilize different modes of reference; the third person NP in (15) refers (or at least can refer) to the reported speaker *de re* while the first person pronoun in (16) obligatorily refers to the reported speaker *de se*. Consequently, the propositions denoted by these two LFs diverge at root: (16) is true only if the reported speaker asserted something that was knowingly about himself while (15) can also be true if the reported speaker asserted something about some individual that happens to be himself but he does not realize this (e.g., by asserting something about a child in a photograph without realizing that the photograph was taken of the speaker in his infancy).

(17) **Indexical Shift can feed BAI.** In an environment in which indexical shift is obligatory, a third person NP may *not* be used to refer to the *reported* speaker.⁹

Section 5.6 is dedicated to the putative prediction (17). I will provide evidence that (17) is indeed empirically accurate in Farsi. I will furthermore point out that *optional* indexical shift cannot feed BAI: if the presence of the context-shifting operator is not independently forced, it cannot generate unacceptability by feeding BAI. Given these facts, one might want to modify BAI to make sure that (17) follows from it.¹⁰ I will

⁹This statement is somewhat reminiscent of Anand's (2006) *De Re Blocking* generalization (which itself generalizes the 'Oneiric Reference Constraint' of Percus & Sauerland (2003b)). According to Anand's formulation, a *de se* pronoun cannot be c-commanded by a *de re* pronoun if both happen to be 'about' the same individual (this is a rough formulation, details do not matter here). In light of this generalization, one may re-formulate (17) as follows: a pronoun cannot refer to a coordinate of the reported speech context *de re* if there is a pronoun in its syntactic vicinity (e.g., same embedded clause) which refers to a (possibly different) coordinate of the reported speech context *de se*. As suggestive as this line of thinking is, there is an important difference between the two generalizations. *De Re Blocking* crucially relies on *c-command*, while, as far as I can see, this is not the case with (17). This suggests that the two generalizations have different sources, but of course the issue needs to be investigated more closely.

¹⁰An alternative is to keep BAI as is and try to provide an alternative explanation of why (17) (stated in pre-theoretical terms: a third person pronoun, embedded in an attitudinal, context-shifted environment, cannot refer to the attitude-holder) holds in Farsi. I have to leave this option for future work, but I would like to point out one *prima facie* promising avenue. It has been observed that epithets, embedded in attitudinal environments, cannot be anteceded by the attitude-holder: **Melvin_i claims*

first discuss in detail why (17) in fact does *not* follow from the current formulation of BAI, given standard assumptions regarding *de re* / *de se* reference. The conclusion will be that one way or the other BAI must be made ‘blind’ to the *de re* / *de se* distinction. I will then argue that this can be done on the assumptions (i) that BAI compares NPs for covaluation (not clauses for semantic equivalence) and (ii) that the relevant notion of covaluation is the one proposed by Sharvit (2010) (‘type-II covaluation’) to account for certain binding theoretic effects in English.

In sections 5.7 and 5.8, I turn to two secondary questions. Up to Section 5.7, the operator-based approach to indexical shift is assumed without argument to the exclusion of the competing approach (or family of approaches), which I will refer to as the binding-based approach (Schlenker 1999, 2003; von Stechow 2004). As is well-known (Anand 2006; Deal 2017), while the operator-based approach derives SHIFT TOGETHER in a principled way on the assumption that the context-shifting operator can only take proposition-denoting complements (e.g., on the assumption that the operator is a complementizer head), the canonical formulations of the binding-based approach can only do so via an imported stipulation. Nevertheless one might ask whether the present dataset yields any information that can be used to further evaluate the relative merits of these theories. In Section 5.7 I argue that it does. Specifically, the operator-based approach (in conjunction with BAI and the auxiliary assumption about *de re* / *de se* synonymy mentioned above) can capture the relevant facts without further ado, while the binding-based approach does not. The main result will be the three-part observation that (i) while the binding-based approach as it stands fails for a number of cases, (ii) if it is supplemented with SHIFT TOGETHER as an imported stipulation its empirical coverage improves significantly but (iii) nevertheless there is at least one kind of example which even the improved version of the theory does not capture (although this example is unproblematic for the operator-based approach).

Up to Section 5.8, the database consists solely on facts pertaining to indexical pronouns and third person NPs, i.e., individual-denoting expressions. The question arises whether the phenomenon under discussion generalizes across sortal domains. In Section 5.8, I seek to address this question by having a brief look at temporal and locative adverbials. I will argue that, for example, *today* and *Monday* compete in very much the same way that *I* and *Jimmy* do, once certain complicating factors having to do with the under-specificity of these expressions (locative *here* in particular) are controlled for. In particular, I will argue that there is an interaction with indexical shift in the former case much like the latter. If correct, this observation has theoretical consequences. The data pertaining to individual-denoting expressions support two views of BAI; one view according to which BAI encodes a competition between marked (first and second person) vs. unmarked (third person) features, and one view

that the bastard_i was honest. The unacceptability of such cases may be understood to be due to a Condition C violation (coupled with the assumption that epithets are r-expressions). But evidence to the contrary comes from the observation that epithets *can* have a c-commanding antecedent in cases that do not involve attitudinal predicates, consider for example *John_i ran over a man (who was) trying to give the idiot_i directions,* which is acceptable. On this basis, Dubinsky & Hamilton (1998) (from whom these examples have been borrowed, and to whose paper the reader is directed for relevant references; see also Patel-Grosz 2014) argue that epithets are pronouns *that come with an anti-logophoricity requirement.* The reason that the former example is unacceptable, in their view, is that “the nonlocal antecedent in [that] case is the perspective-bearer (i.e., the one from whose perspective the attributive content of the epithet is evaluated).” I would like to suggest that there might be a way to unify (17) with the anti-logophoricity of epithets.

according to which the competition is between indexical expressions (first and second person pronouns) and their non-indexical counterparts (third person NPs generally). The data from temporal and locative adverbials strongly support the latter view. This, and other issues, are taken up in the concluding Section 5.9.

5.2 Farsi and indexical shift

Farsi belongs to the class of languages in which embedding indexical pronouns, that is, first and second person pronouns, in the clausal complement of speech-report predicates leads to a systematic ambiguity that boils down to whether the pronouns take their value from the actual or the reported situation of speech. Thus in the Farsi sentence (18) the first and second pronouns can either refer to the actual speaker and addressee, Reading 1, or the reported speaker and addressee, Reading 2. As already pointed out in the previous section, ‘mixed readings’ are not possible; e.g., it is not possible for the first person pronoun in (18) to refer to the actual speaker (as in Reading 1) if the second person pronoun refers to the reported addressee (as in Reading 2).

- (18) Leilā be Minā goft barāt ketāb xaridam.
 L to M told for-2sg book bought-1sg
Reading 1: ‘Leila told Mina that I bought a book for you’
Reading 2: ‘Leila_i told Mina_j that she_i bought a book for her_j’

Example (18) does not rule out the possibility of the embedded clause being quoted as in *Leila told Mina, “I bought a book for you”*. Other constructions are incompatible with such an analysis. Consider (19), which is based on an example from Schlenker 1999.

- (19) Nashnidam Leilā goft chi biyār.
 not-heard-1sg L said what bring-IMP.2sg
 ‘I did not hear what Leila told me to bring’

In (19) the embedded second person pronoun is shifted to refer to the reported addressee, i.e., the actual speaker. This, however, cannot be due to quotation as the latter would generate the reading *I did not hear that Leila said, “(you) bring what?”*. Similarly, in (20) the embedded clause contains two indexical pronouns and the *wh*-word *chi* (what). The sentence is again ambiguous between a shifted and non-shifted reading if it is parsed as a matrix question.¹¹ Here too quotation would generate the wrong interpretation, i.e., *Leila told Mina “what have I bought for you?”*¹²

¹¹As *goftan* is a responsive predicate, one might expect that the *wh*-word can also be interpreted below *goft* in (20). Interestingly, *goftan* does not license indexical shift if it embeds a question. In other words, while *Leila told Mina what I bought for you* (where the indexicals refer to coordinates of the actual context of speech) is a possible reading of (20), *Leila_i told Mina_j what she_i bought for her_j* is not a possible reading of (20). See Section 5.5 for the relevant discussion.

¹²Schlenker (1999) provides the following example from Lazard 1992 (who cites literary sources).

- (i) be shomā xabar dād ke kojā xāham raft
 to you.pl news gave that where will-1sg go

Schlenker’s informant finds this sentence deviant (on the shifted reading) while Lazard reports it as OK. My own judgments, and the one informant that I have consulted, is that the sentence is OK with

- (20) Leilā be Minā goft barāt chi xaridam?
L to M told for-2sg what bought-1sg

Reading 1: ‘What did Leila tell Mina that I bought for you’

Reading 2: ‘What did Leila_i tell Mina_j that she_i bought for her_j’

(21) makes the same point.

- (21) Un ketābe ke Leilā be Minā goft barāt xaridam xeili gerune!
That book-DEF that L to M told for-2sg bought-1sg very expensive-is

Reading 1: ‘The book that Leila told Mina I bought for you is very expensive’

Reading 2: ‘The book that Leila_i told Mina_j she_i bought for her_j . . .’

Indexical shift is possible even if the wording of the embedded clause is not faithful to the original utterance being reported. For example, consider the dialog in (22) between Ali and Mina, followed by the dialog in (23) between Mina and Leila.

- (22) a. **[Mina to Ali:]** “Where did you buy this shirt?”
b. **[Ali to Mina:]** ino Nāzanin xarid barām, nemidunam az
This-OM N bought-3sg for-1sg not-know-1sg from
kojā.
where
‘Nazanin bought this (shirt) for me, I don’t know from where’
- (23) a. **[Leila to Mina:]** “Did you ask Ali where he bought his shirt?”
b. **[Mina to Leila:]** āre, nemidunest, goft xāharam barām
yes, not-knew-3sg, said-3sg sister-1sg for-1sg
xaride.
bought-3sg
‘Yes, he didn’t know. He said that his sister bought it for him.’

Ali’s assertion in (22b) is the original utterance that is reported to Leila by Mina in (23b). Note that while Ali refers to Nazanin, his sister, by name Mina does so using the possessive. Mina’s utterance in (23b) is particularly acceptable in a context in which Leila does not know Ali’s sister by name, hence motivating Mina to use a description. (24) below is another example in a similar vein.

- (24) a. *Context: A new store has opened up in the Tajrish neighborhood of Tehran called Kereshmeh. Leila tells her sister, “I’m going to Kereshmeh to buy a shirt” (miram kereshmeh ie pirhan begiram). Later, Leila’s grandmother asks Leila’s sister where she is. Leila’s sister, knowing that her grandmother would not recognize the shop by name, says:*
b. ie maqāzeye jadid bāz shode tu Tajrish, Leilā goft miram unjā
one store-EZ new open has-become in T, L said go-1sg there
lebās bexaram.
clothes buy.INF-1sg.
‘[A new store]_j has opened in Tajrish, Leila_i said she_i is going there;
clothes-shopping.’

the shifted interpretation of the indexical only if the sentence is parsed as a matrix question: *where did x tell you that x will go?*. See also fn. 11 and the discussion in Section 5.5.

Examples (19) to (24) are problematic for an account based on *clausal* quotation. However, they can be accounted for if partial quotation is admitted; i.e., if it is possible for quotation to target constituents smaller than clauses, including single pronouns. The immediate problem for partial quotation is to account for the SHIFT TOGETHER constraint. It is not clear how an account based on partial quotation can make sure that distinct indexicals in the same minimal domain (the embedded clause) are such that either neither shifts or both shift (see Deal 2017 for an elaboration of this argument). Regardless, the account based on (clausal or partial) quotation can be pushed to rather implausible extremes. Consider the example in (25). This text is taken from a newspaper interview¹³ with a lawyer whose client has been sentenced to death on the basis of a certain *fatwa* (Islamic ruling). The lawyer explains in (25a) how he personally visited two of the most important clerics who have supported the relevant *fatwa* and has explained his client's situation to them. In sentence (25b) the lawyer reports their response.

- (25) a. man xodam xedmat-e Āyatollah ... va Āyatollah ... residam va barāye har do nafar tozih dādam ke chenin chizi bude ...
 'I myself visited *Ayatollah* ... and *Ayatollah* ... and explained the situation to both of them. ...'
- b. har do nafar goftand mā fatvāie koli dādim va darbāreye
 each two person said 1pl fatwa-EZ general gave-1pl and about-EX
 in shaxse xas hokm nadādim.
 this person-EZ particular judgment not-gave-1pl
 'Both of them said they have issued a general fatwa and have not given a judgment about this particular person.'

The crucial aspect of this example is the fact that the *plural* first person pronoun is shifted to refer to the reported speakers, i.e., the two *Ayatollahs* that the lawyer has spoken to. It is, however, completely clear from the context that the lawyer has talked to the *Ayatollahs* separately and is simply reporting the gist of what they had conveyed to him. The problem is that if this sentence is to be analyzed via quotation, the pronouns cannot be quoted, as the plural feature cannot be attributed to the *Ayatollahs*.¹⁴ It must be the case that quotation in this example targets only the first person *feature* of these pronouns. The same point can be made with the following example.

- (26) bā setā az doktorāye bimārestan sohbat kardam, har se
 with three-many from doctors-EZ hospital talk did-1sg, each three
 goftand az nazare mā mo'af hasti.
 said-3pl from view-EZ 1pl exempt are-2sg
 'I talked to three doctors from the hospital, all three said that from their point of view I am exempt (from military service).'

Another angle from which this problem can be viewed is that of ellipsis. First, note

¹³Accessible from <http://magiran.com/n3335357>.

¹⁴It might be argued that the plural feature can indeed be attributed to the *Ayatollahs* as they might have used the plural first person as an honorific (this is common for individuals that speak from a position of high authority or prestige). First, the example (28) below does not suffer from this potential confound. Second, intuitively, the utterance in (25b) does *not* commit the speaker, i.e., the lawyer, to the claim that the *Ayatollahs* used the honorific form.

that in general the number marking on shifted indexicals must agree with the number marking on the subject / indirect object of the attitude. In (27) this is shown with respect to shifted first person.¹⁵

- (27) a. #Ali goft goshnamune
 A said-1sg hungry-1pl-are
 b. #Ali o Minā goftan goshname
 A and M said-3pl hungry-1sg-is

(28) shows that agreement is not necessary when ellipsis is involved. If this was not the case one would expect, e.g., (28a) to be as unacceptable as (27a) since the reported speaker in the second disjunct is singular while the elided material contains a shifted plural first person.¹⁶

- (28) a. yā Ali o Minā goftan goshnamune yā Leilā
 or A and M said-3pl are.hungry-1pl or Leila
 ‘Either Ali and Mina said they are hungry or Leila said she is hungry’
 b. yā Leilā goft goshname yā Ali o Minā
 or L said-3sg is.hungry-1sg or A and M
 ‘Either Leila said she’s hungry or Ali and Mina said they are hungry’

The ‘disappearance’ of ϕ -features on pronouns in environments involving ellipsis is of course expected (Heim 2008; Kratzer 2009, a.o.), although standard examples involve *bound* pronouns while the pronouns in (28) are arguably not bound. However, the acceptability of the examples in (28) is quite problematic, as far as I can see, for an account of indexical shift based on quotation, cf. (29).

- (29) a. #They said “we are hungry” and Bill did too
 b. #Bill said “I am hungry” and they did too.

Before I move on one additional property of indexical shift must be noted. Shifty indexicals in Farsi obligatorily receive *de se* interpretations, as illustrated by the contrast in (30) (based on an example from Anand 2006).

- (30) a. doktor goft behem kudum āmbulo bezanin?
 Doctor said to-1sg which shot-OM give.IMP-2pl
 ‘Which shot did the doctor say we should give him’
 b. doktor goft behesh kudum āmbulo bezanim?
 Doctor said to-3sg which shot-OM give-1pl
 ‘Which shot did the doctor say we should give him’

¹⁵(27a) is marginally acceptable if the shifted plural is interpreted as referring to Ali and some other individuals, e.g., his friends. The reading is only marginally acceptable because out of the blue the identity of those other individuals is not settled. Furthermore, (27b) is also marginally acceptable if the embedded clause is interpreted as involving quotation and the subjects are distributed over, *Ali and Mina each said, “I am hungry”*. This reading is marked and forces the speaker to somehow signal that quotation is involved, via non-standard intonation or air-quotes. Neither of these two observations threaten the claim made in the text as far as I can see.

¹⁶The example (28b) potentially suffers from the potential confound that the second disjunct might be re-analyzed with a distributivity operator generating a reading similar to the one noted in fn. 15. As far as I can see, the example in (28a) does not suffer from any potential confound *vis-à-vis* the claim made in the text.

The sentence in (30a), which involves indexical shift, can be used to report the situation in (31a), but not the one in (31b). The sentence in (30b) (which does not involve indexical shift), on the other hand, can be used to report both situations in (31).

- (31) a. The doctor examines several patient’s reports. For each report he dictates his prescription to the two nurses that are present. Since the reports *are not anonymous*, the doctor recognizes his own chart and, after examining it says “give *me* a shot of Penicillin”. Later, one of the nurses asks the other one (30).
- b. The doctor examines several patient’s reports. For each report he dictates his prescription to the two nurses that are present. Since the reports *are anonymous*, the doctor does not recognize his own chart and, after examining it says “give *this patient* a shot of Penicillin”. Later, one of the nurses asks the other one (30).

5.3 The operator-based account of indexical shift

The simplest implementation of the operator-based approach to indexical shift (Anand & Nevins 2004; Anand 2006) rests on the following assumptions. First, the interpretation function is relativized to a context parameter c , an index parameter i , and a contextually-supplied assignment function g : $\llbracket \cdot \rrbracket^{c,i,g}$. Second, contexts are formalized as tuples of objects of appropriate sorts. For our purposes (abstracting from the temporal dimension) a context c is a triple consisting of a possible world, denoted $\text{WORLD}(c)$, an individual speaker/author, denoted $\text{AUTH}(c)$, and an individual addressee, denoted $\text{ADDR}(c)$. Finally, it is assumed that indexes and contexts are ‘homologous’; each index i also consists of a world $\text{WORLD}(i)$, an author $\text{AUTHOR}(i)$, and an addressee $\text{ADDRR}(i)$. I will occasionally refer to world-author-addressee tuples as formal objects in abstraction from the role they play in interpretation (as context or index) as “centers”. Within this broad framework, a plausible lexical entry for the predicate *goftan* (to say / tell) is (32).

- (32) a. $\llbracket x \text{ be } y \text{ goft } \phi \rrbracket^{c,i,g} = 1$ iff $\forall i' \in \text{SAY}_{x,y}^i : \llbracket \phi \rrbracket^{c,i',g} = 1$
- b. $i' \in \text{SAY}_{a,b}^i$ iff i' is compatible with what a told b in $\text{WORLD}(i)$.

Intuitively, i' is compatible with what a told b in w iff, from a ’s point of view, if her utterance is true then i' could be the context in which her utterance took place.

The proposed lexical entries for indexical pronouns are given in (33). Notice it is not assumed that indexical pronouns are variables; that is, unlike the third person pronoun, the interpretation of indexical pronouns is assignment-independent.

- (33) $\llbracket \text{1sg-pro} \rrbracket^{c,i,g} = \text{AUTH}(c)$, $\llbracket \text{2sg-pro} \rrbracket^{c,i,g} = \text{ADDRR}(c)$

The last ingredient is the context-shifting operator. Since indexical pronouns take their value from the context parameter, they are predicted to refer to the actual speaker / addressee unless this parameter is somehow manipulated. In the operator-based approach this manipulation is effected by a covert operator, OP , which is syntagmatically defined in (34).

$$(34) \quad \llbracket \text{OP } \phi \rrbracket^{c,i,g} = \llbracket \phi \rrbracket^{i,i,g}$$

That the system so far sketched accounts for context-shifting is established by the following toy examples.

- (35) a. $\llbracket \text{L told M that I am hungry} \rrbracket^{c,i,g} = 1$ iff
 $\forall i' \in \text{SAY}_{L,M}^i : \llbracket \text{I am hungry} \rrbracket^{c,i',g} = 1$ iff
 $\forall i' \in \text{SAY}_{L,M}^i : \llbracket \text{hungry} \rrbracket^{c,i',g}(\llbracket \text{I} \rrbracket^{c,i',g}) = 1$ iff
 $\forall i' \in \text{SAY}_{L,M}^i : [\lambda x. x \text{ is hungry in } \text{WORLD}(i')](\text{AUTH}(c)) = 1$ iff
 $\forall i' \in \text{SAY}_{L,M}^i : \text{AUTH}(c) \text{ is hungry in } \text{WORLD}(i')$
- b. $\llbracket \text{L told M that OP I am hungry} \rrbracket^{c,i,g} = 1$ iff
 $\forall i' \in \text{SAY}_{L,M}^i : \llbracket \text{OP I am hungry} \rrbracket^{c,i',g} = 1$ iff
 $\forall i' \in \text{SAY}_{L,M}^i : \llbracket \text{I am hungry} \rrbracket^{i',i',g} = 1$ iff
 $\forall i' \in \text{SAY}_{L,M}^i : \llbracket \text{hungry} \rrbracket^{i',i',g}(\llbracket \text{I} \rrbracket^{i',i',g}) = 1$ iff
 $\forall i' \in \text{SAY}_{L,M}^i : [\lambda x. x \text{ is hungry in } \text{WORLD}(i')](\text{AUTH}(i')) = 1$ iff
 $\forall i' \in \text{SAY}_{L,M}^i : \text{AUTH}(i') \text{ is hungry in } \text{WORLD}(i')$

Note that the operator-based approach immediately captures SHIFT TOGETHER: if OP is present every indexical it c-commands is shifted and if it is absent no indexical is shifted. Furthermore, there is no reason to think that OP is opaque to grammatical transformations, such as extraction. Hence the problems raised in the previous section for the quotation account are not problematic for the operator-based approach.¹⁷

Finally, it follows from the definition of OP, the lexical entries for the verb *goftan* and the indexical pronouns, that shifted indexicals will always be interpreted *de se*. To see this, consider the two scenarios in (31), repeated below.

- (36) a. The doctor examines several patient's reports. For each report he dictates his prescription to the two nurses that are present. Since the reports *are not anonymous*, the doctor recognizes his own chart and, after examining it says "give *me* a shot of Penicillin" (*be man penisilin bezanin*).
- b. The doctor examines several patient's reports. For each report he dictates his prescription to the two nurses that are present. Since the reports *are anonymous*, the doctor does not recognize his own chart and, after examining it says "give *this patient* a shot of Penicillin" (*be in bimar penisilin bezanin*).

The doctor's utterance in (36b) is compatible with any center i' in which ADDR(i'), i.e. the two nurses, comply with AUTH(i')'s, i.e. the doctor's, demands by giving a shot of Penicillin to some individual a in WORLD(i'), a being the person whose file AUTH(i') is examining in WORLD(i'). As there is uncertainty as to the identity of this person, there are centers i' compatible with the doctor's utterance in which a is some individual other than AUTH(i') and there are centers i' in which a happens to be the same individual as AUTH(i'). However, since in situation (36a) the doctor explicitly refers to himself using the first person pronoun, the centers i' compatible with his instructions are precisely those in which the nurses comply with AUTH(i')'s

¹⁷Except for the puzzle raised by shifty plural pronouns in Section 5.2, which does not have an immediate solution in the operator-based approach either. I will have to leave this problem future research.

demands by giving a shot of Penicillin to AUTH(*i'*). Now consider the sentence in (37) asserted by one of the nurses to the other one. Since the first person pronoun used in the embedded position is shifted by assumption, for any index *i'* compatible with what the doctor said, it will refer to AUTH(*i'*). Consequently, this sentence is only compatible with the situation in (36a) and not (36b).

- (37) doktor goft behem penesilin bezanin
 Doctor said to-1sg Penicillin give.IMP-2pl
 ‘The doctor told us that we (should) give him a shot of Penicillin.’

5.4 Indexical shift bleeds the Ban Against Illeism

Consider again prediction (13) from Section 5.1, repeated in (38). As discussed there, this prediction can be tested using the schema in (10), repeated in (39).

- (38) **Indexical Shift can bleed BAI.** In an environment in which indexical shift is obligatory, a third person NP may be used to refer to the actual speaker.

- (39) [A to B:] x told y that [... [3rd NP] ... [2nd-pro] ...]
-

The embedded clause in (39) by assumption contains (i) a third person NP that refers to the actual speaker and (ii) a second person pronoun that is shifted to refer to the reported addressee. Assumption (ii) guarantees that the context-shifting operator c-commands the embedded clause. Thus the LF underlying (39) is approximately (40).

- (40) [A to B:] x told y OP [... [3rd NP] ... [2nd-pro] ...]
-

The embedded clause constitutes an environment in which indexical shift is “obligatory” in the following sense. If BAI compares an LF of the form in (40) to an alternative that is exactly same except that the third person NP is replaced with a token of the first person pronoun then the presence of the context-shifting operator in the alternative LF guarantees that the introduced first person pronoun will be “automatically” shifted to refer to the reported speaker. I.e., the LF of the relevant alternative is approximately (41).

- (41) [A to B:] x told y OP [... [1st-pro] ... [2nd-pro] ...]
-

Thus the conjunction of assumption (i) with the LFs (40) and its BAI-alternative (41) allows us to test prediction (38): if a sentence constructed along the lines of (39) is felicitous then we have evidence that “obligatory” indexical shift bleeds BAI.

Evidence that this prediction is indeed born out comes from the sentence in (42). Given the facts reviewed in Section 5.2, in principle this sentence should be ambiguous between a reading in which the second person pronoun is shifted, Reading 1, and a reading in which it is not, Reading 2. In fact only the former reading is intuitively available and the latter is markedly deviant (to the extent that it can be accessed to begin with). This is as things should be if (i) obligatory indexical shift bleeds BAI

(rendering Reading 1 acceptable) and (ii) no other factor would interact with BAI (rendering Reading 2 unacceptable due to BAI violation).

- (42) **[Sajjad to Qazal:]** Leilā be Minā goft Sajjād azat asabāniye.
 L to M told S from-2sg angry.is-3rd
 ✓**Reading 1:** ‘Leila told Mina_j that Sajjad is angry at her_j’
 ✗**Reading 2:** ‘Leila told Mina that Sajjad is angry at Qazal.’

The sentences in (43) and (44) show the same pattern. These sentences involve extraction from the embedded clause and as such guarantee that clausal quotation is not a confounding factor.

- (43) **[Sajjad to Qazal:]** un chizi ke Leilā be Minā goft bāyad vase Sajjād
 that thing that L to M told should for S
 bexari xeili gerune!
 buy.IMP-2sg very expensive.is
 ✓**Reading 1:** ‘The thing that Leila told Mina_j she_j should buy for Sajjad is very expensive.’
 ✗**Reading 2:** ‘The thing that Leila told Mina Qazal should buy for Sajjad is very expensive.’
- (44) **[Sajjad to Qazal:]** Leilā be Minā goft vase Sajjād chi bexar?
 L to M told for S what buy.IMP-2sg
 ✓**Reading 1:** ‘What did Leila tell Mina_j that she_j should buy for Sajjad?’
 ✗**Reading 2:** ‘What did Leila tell Mina that Qazal should buy for Sajjad?’

As control, (42) could be compared with (45).

- (45) # **[Sajjad to Qazal:]** Leilā be Minā goft Sajjād azash asabānie
 L to M told S from-3sg angry.is-3sg
Intended: ‘Leila told Mina_j that Sajjad is angry at her_j’

The only difference between the two sentences is that in the latter the second person pronoun of (42) is replaced with the third person pronoun referring to Mina, the reported addressee. In this example indexical shift is irrelevant as there is no indexical pronoun in the embedded clause to be shifted. Strikingly, (45) is sharply infelicitous.

According to prediction (38), indexical shift bleeds BAI in environments in which the former is obligatory. What about environments in which indexical shift is optional, that is, examples in which there is no way to guarantee that the indexical pronoun that replaces a third person NP is “automatically” shifted? The framework established so far, i.e., the operator-based account of indexical shift coupled with BAI, does not make any specific predictions for such cases. It could be that indexical shift can only bleed BAI if its presence is independently forced, as in the examples discussed in the previous section. It could also be that, as far as BAI goes, indexical shift can be thought of as a *rescuing mechanism* which, to the extent that its presence is not independently *prohibited*, can bleed BAI to rescue some sentences from a BAI-violation. The choice between the two options must be resolved empirically. This is what I turn to now.

Consider the example in (46).

- (46) **[Sajjad to Qazal:]** Leilā be Minā goft Sajjād asabānie.
 L to M told S angry.is

‘Leila told Mina that Sajjad is hungry.’

This sentence does not contain any indexical expression, but it does contain a third person NP in the embedded clause that refers to the actual speaker. Assuming no interference, then, the sentence is predicted to be blocked by BAI. In fact the sentence is felicitous. This can be explained on the assumption that indexical shift can bleed BAI even if its presence is not otherwise forced. Evidence that it is indeed indexical shift that makes (46) acceptable comes from the unacceptability of (47).

- (47) # [Sajjad to Qazal:] Leila fek-kard Sajjad asabānie.
 L thought S angry.is
 ‘Leila thought that Sajjad is hungry.’

The difference between (46) and (47) is that the predicate *goftan* in the former is replaced by the predicate *fekr-kardan* (to think) in the latter. As indexical shift in Farsi is *not* possible with *fekr-kardan*, the contrast between the two examples supports the claim that the felicity of (46) is due to indexical shift.

Notice that example (46) does not contain any indexical expression. This is no accident: no indexical expression can be present in the clause embedded by *goftan* if indexical shift is to be optional. But if no indexical expression is present then indexical shift is *vacuous* as there is nothing for it to shift. The felicity of (46), therefore, supports the idea that OP can bleed BAI even if, strictly speaking, it is semantically vacuous.¹⁸

5.5 Indexical shift and question-embedding: a case study

Suppose we can construct a sentence *S* which is in principle structurally ambiguous between two LFs, ϕ_1 and ϕ_2 . Furthermore, let us assume ϕ_1 does *not* allow indexical shift while ϕ_2 does. If the conclusion of the previous section, namely, that indexical shift can bleed BAI even if the presence of the context-shifting operator is not forced for independent reasons, is correct, then we predict that if *S* contains a third person NP that refers to the actual speaker / addressee, then ϕ_1 will not be an acceptable parse for *S*; that is, we predict that in such cases the only acceptable reading of *S* is the one expressed by ϕ_2 . In this section, I would like to argue that this prediction is correct. Evidence for this claim comes from the interaction between question-embedding and indexical shift. Specifically, *goftan* is a responsive predicate and, as such, in general embeds questions. Interestingly, *goftan* does *not* license question-embedding and indexical shift simultaneously: if it embeds a question, then it does not allow indexical shift. I will not make any positive claims regarding why this is the case (but see the end of this section for a brief discussion); but I will use this generalization to test the prediction sketched just above.¹⁹ Schematically, then, I will

¹⁸This conclusion, if true, is rather surprising on the view that the distribution of covert operators in general must be subject to an economy condition which punishes their semantically vacuous occurrences. One possible solution would be to make sure that the context-shifting operator is *not* vacuous in the relevant cases. For example, if we postulate that the third person feature is semantically interpreted along the lines sketched in 6, then this *desideratum* is met.

¹⁹There is a lacuna in the literature on how indexical shift and embedded questions interact. Consequently, as the moment it is unclear whether the incompatibility of question-embedding and indexical

consider sentences of the following form.

(48) [A_i to B:] *x* told *y* [. . . [3rd NP]_{*i*} . . . *wh* . . .]

In principle, a sentence of this form should be ambiguous between a matrix-question reading (*wh* > *goftan*) and an embedded-question reading (*goftan* > *wh*). But since the embedded clause contains a third person NP that refers to the actual speaker, we predict that the sentence is going to violate BAI unless indexical shift bleeds BAI (assuming that there is no other mechanism, relevant here, that can bleed BAI). However, if the generalization is correct that indexical shift is incompatible with the embedded-question reading, then the only acceptable parse would have to be the one that expresses the matrix-question reading. Below I will flesh out this reasoning and provide evidence that the prediction is correct.

First, consider the sentence in (49) which involves a *wh*-word embedded in the clausal complement of *goftan*. This sentence is ambiguous (depending on intonation) between a matrix question reading and an embedded question reading, as *goftan* is a responsive predicate and Farsi is an *wh*-in situ language.²⁰

(49) Leilā be Minā goft ki da'vatesh kard?
 L to M told who invite-EZ-3sg did-3sg
 ✓Reading 1: 'Who did Leila_{*i*} tell Mina invited her_{*i*}?'
 ✓Reading 2: 'Leila_{*i*} told Mina who invited her_{*i*}'

The same ambiguity arises if the third person pronoun in (49) is replaced with the first person pronoun but, interestingly, only if this pronoun is *not* shifted; if the pronoun is shifted the embedded question reading becomes impossible:

(50) Leilā be Minā goft ki da'vatam kard
 L to M told who invite-EZ-1sg did-3sg
 ✓Reading 1: 'Who did Leila tell Mina invited me?' [wh > say, -shift]
 ✓Reading 2: 'Leila told Mina who invited me.' [wh < say, -shift]
 ✓Reading 3: 'Who did Leila_{*i*} tell Mina invited her_{*i*}?' [wh > say, +shift]
 ✗Reading 4: 'Leila_{*i*} told Mina who invited her_{*i*}.' *[wh < say, +shift]

The absence of Reading 4 in (50) is evidence that, at least in Farsi (cf. fn. 19), if the speech-report predicate *goftan* embeds a question then it does not license indexical shift. Next, consider (51) which has the form of the schema in (48).

(51) [Sajjad to Qazal:] Leilā be Minā goft ki Sajjād-o da'vat kard?
 L to M told who S-OM invite did-3sg
 ✓Reading 1: 'Who did Leila tell Mina invited Sajjad?'
 ✗Reading 2: 'Leila told Mina who invited Sajjad.'

The complement of *goftan* does not contain any indexical expression in (51). Instead it contains a third person NP that refers to Sajjad, the actual speaker. Interestingly, the

is specific to Farsi or whether it generalizes to other languages. For example, one might suspect the fact that Farsi is an *in situ* language plays a role here. If so, the prediction is made that this generalization does not apply to *ex situ* languages (that license indexical shift). Naturally, I have to leave this question to future work.

²⁰I will construct examples with the *wh*-word *ki* (who). As far as I can see, nothing hinges on this. Every interrogative form, including polar and alternative questions, can be used to make essentially the same point.

embedded question reading is judged unacceptable while the matrix question reading is readily accessible. This fact is entirely expected if the context-shifting operator is a rescuing mechanism *via-à-vis* BAI: as the sentence contains a third person NP that refers to the actual speaker, the only way for it to be felicitous is for indexical shift to bleed BAI (assuming there is no other way to escape the wrath of BAI). But indexical shift is impossible if the *wh*-item *ki* (who) is interpreted below the predicate *goftan*. Consequently, the only acceptable LF for (51) is one in which *ki* is interpreted at root and the context-shifting operator is inserted to rescue the sentence from a BAI violation.

On the one hand, the observations made above corroborate the conclusion of the previous section, that indexical shift is a readily available mechanism to bleed BAI (but, of course, only in environments that license indexical shift to begin with), even if semantically vacuous. On the other hand, these data might shed some light on the reason why question-embedding and indexical shift are incompatible (in Farsi). Specifically, note that since in (51) there is no indexical expression, the context-shifting operator would be vacuous. The reason that the presence of this operator is incompatible with question-embedding, then, appears to be syntactic in nature: it is the sheer presence of the context-shifting operator that interferes with question-embedding, even if, as is the case in (51), the former does not have any detectable semantic consequences.²¹

5.6 Indexical shift feeds the Ban Against Illeism

Consider the schema (52) repeated from Section 5.1.

(52) [A to B:] x told y that [... [3rd NP] ... [2nd-pro] ...]

This schema contains (i) a third person NP that refers to the reported speaker and (ii) a second person pronoun that refers to the reported addressee. The latter guarantees that the LF underlying any sentence of the form (52) contains an occurrence of the context-shifting operator in the embedded clause.

(53) [A to B:] x told y [OP [... [3rd NP] ... [2nd-pro] ...]]

As it stands, (53) does not violate BAI with respect to the coordinates of the *actual* context of speech as it does not contain any NP that refers to any of those coordinates. Does it violate BAI with respect to the *reported* context of speech? Specifically, suppose the third person NP of (53) is replaced with a token of the first person pronoun. Notice that since OP is syntactically present, the introduced first person pronoun is bound to shift.

(54) [A to B:] x told y [OP [... [1st-pro] ... [2nd-pro] ...]]

BAI predicts that (54) blocks (53) only if the two are semantically equivalent. Strictly

²¹This conclusion can be challenged if, for example, we assume that the third person feature is indexical along the lines sketched in fn. 6.

speaking, this is not the case. The reason is that in (54) the first person pronoun refers to the reported speaker *de se* through the mechanism of context-shifting while in (53) the third person NP refers (or can refer) to the reported speaker *de re*.²² Therefore, the two alternatives are not in fact semantically equivalent: while (54) entails (53)²³ the reverse is not the case.

In the rest of this section, I will first provide evidence that sentences of the form in (52) are in fact unacceptable in Farsi. On the assumption that the deviance of these examples is in fact due to BAI (as no alternative explanation is readily available, cf. fn. 10), the theoretical task is to modify BAI in order to derive the desired effects. I will propose two modifications. I will argue that BAI, instead of comparing clauses for semantic equivalence, needs to compare noun phrases for covaluation. Not any notion of covaluation will do, however. What BAI seems to be sensitive to, is a notion of covaluation that is *not* fine-grained enough to register the semantic difference between *de re* and *de se* reference. I will argue that the relevant notion of covaluation is in fact the one that is proposed by Sharvit (2010), under the heading ‘type-II covaluation’ (equally appropriate is ‘intensional covaluation’), to account for certain puzzling binding-theoretic effects in English.

Let us first verify that sentences that instantiate the schema in (52) are indeed unacceptable. Consider (55).

- (55) Leilā be Minā goft barāt ketāb xaride.
 L to M told for-2sg book bought-3sg
 ✗ **Reading 1:** ‘Leila_i told Mina_j that she_i bought her_j a book.’
 ✓ **Reading 2:** ‘Leila told Mina_j that she_k bought her_j a book.’
 ✓ **Reading 3:** ‘Leila_i told Mina that she_{i,k} bought you a book.’

Reading 3 establishes that if the indexical second person pronoun is *not* shifted, the third person pronoun may or may not refer to the reported speaker, Leila. The contrast between Readings 1 and 2 establishes that if the indexical pronoun does shift then the third person pronoun can no longer refer to the reported speaker: it must refer to some salient individual distinct from the reported speaker (or addressee, for that matter). Example (56) makes the same point with extraction, making sure that quotation is not a confounding factor.

- (56) Leilā be Minā goft barāt chi xaride?
 L to M told for-2sg what bought-3sg
 ✗ **Reading 1:** ‘What did Leila_i tell Mina_j she_i bought for her_j?’
 ✓ **Reading 2:** ‘What did Leila tell Mina_j she_k bought for her_j?’
 ✓ **Reading 3:** ‘What did Leila_i tell Mina she_{i,k} bought for you?’

Again, as witnessed by Reading 3, if the embedded second person pronoun is *not* shifted the third person pronoun may or may not refer to the reported speaker. If, however, the embedded indexical pronoun is shifted the third person pronoun may

²²Percus & Sauerland (2003a) argue that third person pronouns can receive a dedicated *de se* parse. This does not affect the argument made in the text: if Sauerland & Percus are correct, then third person pronouns can be ambiguously represented either as *de se* pronouns or *de re* pronouns. The former is straightforwardly blocked by BAI in the relevant environments. The problem is the existence of the *de re* LF of third person pronouns, which is uncontroversial.

²³I assume following much previous literature that *de re* LFs can in principle generate *de se* readings via the ‘self’ acquaintance relation.

refer to some salient individual (Reading 2) but not the reported speaker (*Reading 1). The reader will note that in both (55) and (56) Reading 1 is precisely the one that corresponds to the schema in (52) with a shifted indexical and a non-indexical pronoun referring to coordinates of the reported situation of speech. I conclude that when indexical shift is made obligatory by the presence of another shifted indexical, illeism can indeed feed BAI, generating oddness.

I noted above that indexical shift can bleed BAI even if it is not obligatory. The conclusion was that indexical shift can be thought of as a mechanism that is always available to bleed BAI, rescuing certain sentences from oddness. If this line of thinking is correct then we would expect *optional* indexical shift *not* to feed BAI. This is indeed the case as the acceptability of the sentence below demonstrates.

- (57) [Sajjad to Qazal:] Leilā_i be Minā goft p_rθ_i asabanie.
 L to M told pro angry-is-3SG
 ‘Leila told Mina that she is hungry.’

This examples makes it clear that the third person pronoun in and out of itself is quite capable of referring to the reported speaker / addressee; i.e., it is not the case that the third person pronoun in Farsi has an ‘anti-logophoric’ requirement (cf. fn. 10).²⁴ This observation in conjunction with the contrast between Readings 2 and 3 of (55) and (56) above establishes that indexical shift plays a crucial role in turning the third person pronoun into an anti-logophoric term.

The examples above might allow for an analysis on the basis of the assumption that there is a strong bias for *de re* pronouns to refer *de se* (i.e., via the ‘self’ acquaintance relation). If so, then the absence of Reading 1 in (55) and (56) follows from BAI since a third person pronoun that refers to the reported speaker / addressee *de se* can be replaced with a shifted first or second person pronoun without changing the overall meaning of the utterance. However, this approach has two short-comings. First, it does not explain why (57) is acceptable. After all, here too, *if the context-shifting operator is inserted*, the third person pronoun could be replaced with a first person pronoun without changing the meaning. Perhaps the context-shifting operator is simply not allowed in such cases. The second problem is much more significant. Rather surprisingly, third person reference to the attitude holder in a context-shifted environment is unacceptable even if the original utterance being reported is explicitly non-first personal. Consider the following context, slightly modified from (31b) in section 5.2.

- (58) *Context: The doctor examines several patient’s charts. For each chart he dictates his prescription to a nurse. Since the charts are anonymous, the doctor does not recognize his own chart and, after examining it says “give this patient a shot of Penicillin”.*

²⁴Schlenker (1999) notes that in Amharic third person reference to the attitude-holder is usually unacceptable (see fn. 25):

- (i) John jigna nə-w alɛ
 John hero be.PRES-3smO say.PERF.3sm
 ‘John_i said that he_{*i,j} is a hero.’

The contrast between (57) and this example suggests that the phenomenon under discussion in Farsi is not quite the same as the Amharic case. See also fn. 25.

Even if this is the context that is being reported, the utterance in (59) is judged sharply infelicitous.

- (59) #doctor be xānum parastār goft behesh kudum āmpulo bezan?
 Doctor to lady nurse said to-3SG which shot give-2SG.IMP
Intended: ‘Which shot did the doctor_i tell the nurse_j that she_j should give him_i?’

The unacceptability of (59) is surprising because it seems to suggest that a mere preference for *de se* interpretation of *de re* pronouns is not enough. If it was, one would expect (59) to be entirely acceptable as the situation being reported is one in which attitude-holder’s attitude is explicitly non-first personal. Indeed the only way to express the intended reading of (59) in the context of (58) is to avoid using indexical shift altogether (i.e., to replace the embedded second person pronoun with a third person pronoun referring to the nurse).²⁵

Having established that indexical shift feeds BAI, we now need to account for it. As noted above the main obstacle in doing to is the fact that the alternatives that BAI needs to evaluate are *not* semantically equivalent. One alternative involves an indexical pronoun that shifts and refers to, say, the attitude-holder *de se* while the other alternative involves a third person noun phrase that refers to the attitude holder *de re*. I do not see a way of solving this problem while maintaining that BAI applies only at the clausal level. However, if we allow ourselves to check BAI at the level of individual NPs then we might be able to get a grip on the problem. Specifically, consider the following puzzle. The (English) sentence (60) is intuitively true in the following context (taken from Sharvit 2010): “Sarah Palin, who is running for president, wakes up from a coma and suffers from severe memory loss: she doesn’t remember that she is running for president and perhaps doesn’t even know who she is. McCain visits her in the hospital, and she says to him: ‘I don’t know who to vote for.’ While the two of them look at a picture of her in the newspaper, he says to her: ‘You must vote for this woman.’ Palin, who doesn’t recognize herself in the picture, says: ‘You are right; I will vote for this woman. She seems reliable.’”

- (60) McCain convinced Palin to vote for herself

The problem is that on standard assumptions there is no grammatical LF of (60) that can be true in such a context. To see this, consider the following two candidates. (61a) is not true in the context given above because here the reflexive pronoun is co-referential with *PRO* which refers to Palin *de se*. (61b), on the other hand, would be true in such a context however it is ruled out by Condition A of binding theory as

²⁵The example in (59) is quite significant as it marks a departure from other, superficially similar examples in the literature. As noted in fn. 24, Schlenker (1999) provides evidence that in Amharic third person reference to the attitude holder is unacceptable *with the exception of cases where the situation being reported is explicitly non-de se*, as in the following example which Schlenker gives as OK.

- (i) a. John sees on TV a candidate he likes a great deal, and says, “I think he’s great!” Unbeknownst to John, he is talking about himself.
 b. John jigna nə-w alɛ
 John hero be.PRES-3smO say.PERF.3sm
 ‘John said that he is a hero.’

Again, there is a contrast between the Farsi case and the Amharic case.

the reflexive is not covalued with an NP in its local domain (marked here ‘LD’).

- (61) a. [McCain convinced Palin [1 [LD PRO₁ vote for herself₁]]]
 b. [McCain convinced Palin₂ [1 [LD PRO₁ vote for herself₂]]]

Sharvit’s solution to this puzzle is make sure that (61b) does *not* violate Condition A. The way she does this is by introducing a novel notion of covaluation that Condition A would rely on, what she calls type-II covaluation. The definition of this notion is rather complicated, and the reader is invited to consult Sharvit’s paper for the details. For my purposes, the important point is that what type-II covaluation does in (61b) is that it makes *PRO* and *herself* ‘synonymous’ even though the former refers to Palin *de se* while the latter does so *de re*. In other words, type-II covaluation is ‘blind’ to the *de se* / *de re* distinction; precisely what we need given the discussion above. On this basis, the proposed modification of BAI amounts to the following rule.

- (62) Let ϕ and ϕ' be two sentences such that the only difference between them is that a third person NP in ϕ (call it α) is replaced by an indexical pronoun in ϕ' (call it β). In any context in which α in ϕ is covalued with β in ϕ' (either in the standard sense, or in the sense of type-II covaluation), ϕ is unacceptable. (cf. (4) above)

Before I close this section, it should be noted that the data discussed in section seem to be subject to cross-linguistic variation. Specifically, Anand (2006) provides the following sentence from Zazaki (his example 342), which he claims is acceptable. This sentence involves a shifted second person pronoun and a third person pronoun which refers to the reported speaker. It should be clear that the abstract structure of this example is the same as that of (55) above.

- (63) Rojda Ali-ra va kε ae braye tiya pace kerda
 Rojda Ali-to said that she brother your kiss do-PERF
 ‘Rojda_i told Ali_k she_i kissed his_k sister.’

It is not clear to me how this rather peculiar cross-linguistic variation is to be accounted for (assuming that the reported judgments on all sides prove to be robust). It appears that Farsi and Zazaki are on the two extreme ends of an ill-understood typology pertaining to the referential possibilities of third person pronouns in attitudinal contexts. While in Farsi no third person pronoun is allowed to refer to the subject or indirect object of a context-shifting predicate, no such restriction is in place in Zazaki, and Amharic seems to occupy the middle ground (see fn. 24 and fn. 25). One possibility, if the discussion above is on the right track, is to put the blame on type-II covaluation. Recall that this notion practically amounts to blindness to *de re* vs. *de se* reference. This is very much in line with native speakers’ judgments, who, notoriously, have trouble with detecting *de re* readings that not *de se*. One might suspect that Farsi speakers have internalized this ‘blindness’ to the *de re* / *de se* distinction (or, rather, have not internalized the distinction) while Zazaki speakers are particularly attuned to it. One wonders whether the examples that motivated Sharvit’s type-II covaluation can be replicated in Zazaki with the same judgments; if Zazaki speakers refuse to ignore the *de re* / *de se* distinction, then they might also find the sentence corresponding to (60) false in the target context. Obviously much more work is needed in this area.

5.7 Comparing the operator-based theory to the binding-based theory

We have seen in sections 5.4 to 5.6 that the operator-based approach to indexical shift, coupled with BAI modified to rely on type-II covaluation instead of semantic equivalence at the clausal level, captures the relevant data points in Farsi while leaving some open cross-linguistic questions. My goal in this section is to investigate the binding-based approach (Schlenker 1999, 2003, von Stechow 2004; see also Schlenker 2004) in the context of the same paradigm. The question that I aim to answer is whether the present paradigm yields any information that can be used to evaluate the relative (de-)merits of these two proposals.

For the sake of concreteness, I will sketch a simple version of the binding-based approach to indexical shift. This toy model is designed to rely on a set of assumptions that are as similar as possible to the operator-based approach, as the latter was sketched in Section 5.3; in particular, the toy model follows the ‘double indexing’ guidelines by introducing both context and index variables / abstractors in the syntax (although the latter are ignored for convenience from a certain point on). The specific details of implementation do not matter, however, and as far as I can see the claims made below generalize to more sophisticated implementations of the binding-based approach.

The chief characteristic of the binding-based approach is the syntactic representation of context variables and abstractors at LF. Thus in this framework first and second person pronouns can be represented as consisting of (i) a context variable with (ii) first and second person features denoting functions from contexts to individuals. For example, the LF representation of the first person pronoun (abstracting away from gender and number) would be $[1st\ c_i]$ with the following semantics.

- (64) a. $\llbracket 1st \rrbracket^g = \lambda c.AUTH(c)$, $\llbracket 2nd \rrbracket^g = \lambda c.ADDR(c)$
 b. $\llbracket c_i \rrbracket^g = \#$ unless $g(i)$ is a context;²⁶ if $\neq \#$ then $\llbracket c_i \rrbracket^g = g(i)$
 c. $\llbracket [1st\ c_i] \rrbracket^g = AUTH(g(i))$, $\llbracket [2nd\ c_i] \rrbracket^g = ADDR(g(i))$

On the simplifying assumptions that context variables should always be bound and that predicates project index abstractors, the LF underlying (65a) is (65b) with the semantics provided in (65d) (as before, I abstract from tense and related issues).

- (65) a. I am hungry
 b. $[\lambda c_i \lambda i_k \llbracket [1st\ c_i] \text{ am } [hungry\ i_k] \rrbracket]$
 c. $\llbracket hungry \rrbracket^g = \lambda i \lambda x. x$ is hungry in $WORLD(i)$
 d. $\llbracket (65b) \rrbracket^g = \lambda c \lambda i. AUTH(c)$ is hungry in $WORLD(i)$

Adopting the truth convention given in (66), (65a) is predicted to be true in context c iff the speaker of c is hungry in the world coordinate of c (at the time coordinate of c).

²⁶A context is any tuple $\langle s, a, t, w \rangle$ where s and a are individuals, t is a time and w is a possible world such that (i) a and d exist in w at t and (ii) there is a speech-act in w at t the speaker of which is s and the addressee of which is a . For the present purposes the time coordinate can be ignored. For any context c , $AUTH(c)$ is the speaker coordinate of c , $ADDR(c)$ is the addressee coordinate of c and $WORLD(c)$ is the world coordinate of c .

- (66) **Truth Convention.** If ϕ is an LF of sentence S uttered in context c and if g_c is the assignment function that represents the ‘referential intentions’ of the speaker of c , ϕ is true relative to c iff $\llbracket \phi \rrbracket^{g_c}(c)(c) = 1$.

We assume that attitudes in general quantify over objects as fine-grained as contexts.

- (67) a. John believes that I am hungry.
 b. $\lambda c_i \lambda i_k \underbrace{[\text{John } [\text{believes } i_k] [\lambda i_{k'} \llbracket [1st\ c_i] \text{ am } [\text{hungry } i_{k'}] \rrbracket]]]}_{\phi}$
 c. $\llbracket [\text{believe } i_k] \phi \rrbracket^g = \lambda x. \forall i' \in \text{DOX}(x, g(k)) : \llbracket \phi \rrbracket^g(i') = 1$
 d. $\llbracket \phi \rrbracket^g = 1$ iff $\forall i' \in \text{DOX}(J., g(k)) : \text{AUTH}(g(i))$ is hungry in $\text{WORLD}(i')$
 e. $\llbracket (67b) \rrbracket^g = \lambda c \lambda i. \forall i' \in \text{DOX}(J., i) : \text{AUTH}(c)$ is hungry in $\text{WORLD}(i')$

Applying the Truth Convention, then, (67a) is predicted to be true in context c iff every center i compatible with what John believes in $\text{WORLD}(c)$ (at $\text{TIME}(c)$) is such that the speaker of c is hungry in $\text{WORLD}(i)$ (at $\text{TIME}(i)$).

Regarding attitudes that allow indexical shift, we simply assume that these introduce a context abstractor in their clausal complement on top of the index abstractor that is already projected by the embedded predicate. The presence of this new context abstractor then allows for indexical shift if the indexical expressions are bound by it. For example, on the pretense that *believe* is a context-shifter, the LF of (67a) is the one given below.

- (68) a. John believes that I am hungry.
 b. $\lambda c_i \lambda i_k \underbrace{[\text{John } [\text{believes } i_k] [\lambda c_{i'} \lambda i_{k'} \llbracket [1st\ c_i] \text{ am } [\text{hungry } i_{k'}] \rrbracket]]]}_{\phi}$
 c. $\lambda c_i \lambda i_k \underbrace{[\text{John } [\text{believes } i_k] [\lambda c_{i'} \lambda i_{k'} \llbracket [1st\ c_{i'}] \text{ am } [\text{hungry } i_{k'}] \rrbracket]]]}_{\phi}$
 d. $\llbracket [\text{believe } i_k] \phi \rrbracket^g = \lambda x. \forall i' \in \text{DOX}(x, g(k)) : \llbracket \phi \rrbracket^g(i')(i') = 1$
 e. $\llbracket (67b) \rrbracket^g = \lambda c \lambda i. \forall i' \in \text{DOX}(J., i) : \text{AUTH}(c)$ is hungry in $\text{WORLD}(i')$
 f. $\llbracket (68c) \rrbracket^g = \lambda c \lambda i. \forall i' \in \text{DOX}(J., i) : \text{AUTH}(i')$ is hungry in $\text{WORLD}(i')$

Now, going back to BAI-related data, the first thing to note is that, unlike the operator-based approach, the binding-based approach needs to be supplemented with auxiliary assumptions to make predictions to begin with. To see this consider example (69) taken from Section 5.4.

- (69) **[Sajjad to Qazal:]** Leilā be Minā goft Sajjād azat asabānie.
 L to M told S from-2sg angry.is-3rd
 ✓**Reading 1:** ‘Leila told Mina_j that Sajjad is angry at her_j’
 ✗**Reading 2:** ‘Leila told Mina that Sajjad is angry at Qazal.’

To account for how indexical shift bleeds BAI in (69) within the binding-based approach the proper noun needs to be replaced with the first person pronoun. But notice that this is not enough; we also need to decide which context-abstractor should bind the introduced first person pronoun. Put differently, the question is whether the LF of (69), sketched in (70), should be taken to compete with (70a) or (70b). (For simplicity, from now on index variables / abstractors are ignored.)

- (70) $[\lambda_{c_1} \text{ L told M } [\lambda_{c_2} \text{ S is angry at } [2\text{nd } c_2]]]$
 a. $[\lambda_{c_1} \text{ L told M } [\lambda_{c_2} [1\text{st } c_2] \text{ is angry at } [2\text{nd } c_2]]]$
 b. $[\lambda_{c_1} \text{ L told M } [\lambda_{c_2} [1\text{st } c_1] \text{ is angry at } [2\text{nd } c_2]]]$

Three potential decision-procedures come to mind. First, working our way backwards, notice that for (69) to be acceptable we need to make sure that it competes with (70a). In (70b) the introduced first person pronoun is bound by the matrix context-abstractor and, therefore, refers to the actual speaker; this LF is predicted to block (69) via BAI. Now, the difference between (70a) and (70b) is that in the former the first person indexical is bound *locally* while in (70b) it is bound by the matrix abstractor. So one idea is to encode a preference for local binding.

- (71) In computation of BAI-alternatives, introduced indexicals must be bound by the closest context-abstractor.

Notice that adopting (71) is tantamount to hard-wiring a preference for the shifty interpretation of the introduced indexicals. As a result of this, it should be clear that the cases where indexical shift bleeds BAI are now captured with no problem: a third person NP embedded under *goftan* can refer to the actual speaker / addressee with no problem because if it is replaced with an indexical pronoun, that pronoun will be shifted automatically and will fail to refer to the actual speaker / addressee, thereby generating a distinct reading. One immediate problem that (71) faces, however, is that it is too restrictive when it comes to cases where indexical shift might *feed* BAI. As pointed out in Section 5.6, indexical shift feeds BAI, generating unacceptability *only if it is made obligatory* (i.e., optional indexical shift does not feed BAI). To see the problem, consider the following example from in Section 5.6.

- (72) **[Sajjad to Qazal:]** Leilā_i be Minā goft $\text{pr}\theta_i$ asabānie.
 L to M told pro angry.is
 ‘Leila told Mina that she is hungry.’

This sentence, which is in fact impeccable, is now ruled out by (71) because the third person pronoun can be replaced with the first person and (71) would have the latter bound by the context variable introduced by the predicate, meaning that the introduced indexical is shifted to refer to the reported speaker. The resulting LF would then block the original LF via BAI on the assumption that BAI is blind to *de re / de se* distinction, as argued in the previous section.

The second option is to capitalize on the idea of charity: there is at least one way to disambiguate (70) which results in an LF that does *not* block (69) through BAI. We could stipulate that this is enough to make (69) felicitous.

- (73) Suppose ϕ competes with ψ and the latter can be disambiguated in different ways. If there is at least one grammatical LF of ψ which does not block ϕ through BAI then ϕ is acceptable.

Much like (71), this analysis can capture the facts pertaining to how indexical shift bleeds BAI, but it can also capture the acceptability of (72), as the reader can verify. Evidence against (73) comes from cases like the one below.

- (74) Leilā be Minā goft barāt ketāb xaride.
 L to M told for-2sg book bought-3sg

X Reading: ‘Leila_i told Mina_j that she_i bought her_j a book.’

The problem here is that the alternative constructed by replacing the third person pronoun with the first person and co-indexing the latter with the *matrix* context-variable results in a (grammatical) LF that does *not* block this sentence through BAI. Consequently, (73) would predict acceptability, contrary to fact.

Third, and finally, notice that in (70) the first person pronoun is clause-mate with a shifted indexical, namely the second person pronoun. As already discussed in Section 5.1 the major problem the binding-based approach is that it cannot account for SHIFT TOGETHER. If it could, then we would automatically predict (70a) to be the LF that we need to consider. So one way to approach the present problem is to reduce it to the latter: in effect, the idea here is that any modification to the binding-based approach that allows it to come to grips with SHIFT TOGETHER will allow it to capture the BAI related data as well.

(75) In the computation of alternatives, all introduced indexicals must obey SHIFT TOGETHER.

Similar to analysis (71), this analysis can capture the data in sections 5.4 and 5.6. However, in cases where there is no indexical in the sentence other than the first person introduced by BAI, we are back to square one: which context should the first person be co-indexed with? As there is no other indexical by assumption, (75) does not make any predictions. Now, one moral of the discussion in the previous sections was that while indexical shift can rescue LFs from BAI violations, it does not force such violations. Consequently, we can effectively combine (75) with (73) as follows.

(76) a. In the computation of alternatives, all introduced indexicals must obey SHIFT TOGETHER.
 b. If there is at least one grammatical LF of ψ which obeys SHIFT TOGETHER but does not block ϕ through BAI then ϕ is acceptable.

As the reader can verify, this analysis can capture all the data so far discussed in the paper within the binding-based approach. It supports the conclusion that any version of the binding-based approach which can capture SHIFT TOGETHER can capture the BAI-related data as well if it is coupled with a principle of charity along the lines of (73). I would like to finish this section by pointing out a final data point that casts doubt on this conclusion.

(77) a. *Context: it is common ground that Leila’s only brother is Ali.*
 b. # [Ali:] Leilā goft barādaresh goshnas.
 L said brother-3SG hungry-is
 Intended: ‘Leila said that her brother is hungry’

Here we have a third person possessive NP, *her brother*, which I will assume is referential, and which embeds a third person pronoun. The sentence is odd, a fact that follows from the operator-based approach and BAI immediately. To derive the unacceptability of this sentence we need to make sure that BAI rules out both the parse of (77), with the context-shifting operator and the parse of it without the context-shifting operator.

(78) a. Leila said her brother is hungry

- b. Leila said OP her brother is hungry

The latter is ruled out because of the embedded third person pronoun, effectively in the same manner as the data discussed in Section 5.6. The former is also ruled out: the complex NP itself refers to Ali by contextual assumptions (indeed, the sentence is quite acceptable if it is not part of the background assumptions that Leila's only brother is Ali), therefore the LF is contextually equivalent with the alternative that is derived by replacing the whole possessive NP with the first person pronoun.

How does the binding-based approach deal with this data point? To see the prediction, note that as there is no indexical other than the first person introduced by BAI, (76) effectively boils down to (73): if we can find at least one competing LF that does not block the sentence through BAI, the sentence is predicted, incorrectly, to be felicitous. In fact, we can find two.

- (79) a. $[\lambda c_1 L \text{ said } [\lambda c_2 [1\text{st } \text{pro}_2] \text{ is hungry}]]$
 b. $[\lambda c_1 L \text{ said } [\lambda c_2 [[1\text{st } \text{pro}_1]'s \text{ brother}] \text{ is hungry}]]$

In (79a) the whole NP *her brother* is replaced with a first person pronoun which is shifted to refer to Leila. Clearly this LF does not block (77) as the two have two entirely different meanings. The same applies to (79b) in which the third person pronoun is replaced with a first person pronoun which is not shifted. Again we get a meaning that is very different from that of (77).

I conclude that the data point in (77) in particular is a significant problem for the variable-based account, but as far as I can see the operator-based account accounts for all the data discussed in the paper.²⁷

5.8 Generalizing across sortal domains

In this section, I will briefly discuss the case of locative and temporal adverbials. The goal is to see whether the indexical expressions in these categories follow the same pattern as indexical pronouns discussed above.

To begin with, let us note that locative and temporal adverbials in Farsi are subject to SHIFT TOGETHER much like indexical pronouns. This is established by the following examples. In (80b), for instance, the indexical *diruz* (yesterday) must shift if the first

²⁷A potential response on the part of binding-based theorist would be post that third person features are indexical as sketched in fn. 6. If so, then the LF for (77b), assuming SHIFT TOGETHER, would be one of the following, with two third person features represented in the syntax, one for the embedded pronoun and one for the whole possessive.

- (i) a. $[\lambda c_1 \text{ Leila}_i \text{ said } \lambda c_2 [[3\text{rd } c_1] [[3\text{rd } c_1] \text{ pro}_i]'s \text{ brother}] \text{ is hungry}]$
 b. $[\lambda c_1 \text{ Leila}_i \text{ said } \lambda c_2 [[3\text{rd } c_2] [[3\text{rd } c_2] \text{ pro}_i]'s \text{ brother}] \text{ is hungry}]$

The third person feature on the possessive NP in the former LF triggers the presupposition that whoever the possessive refers to is not the actual speaker, which generates oddness in the target context. The third person feature on the embedded pronoun in the latter LF, in turn, triggers the presupposition that whoever that pronoun refers to, i.e., Leila, must not be the reported author (*modulo* the familiar *de re / de se* distinction), which, again, is falsified. Both LFs are, then, ruled out and the sentence is correctly predicted to be infelicitous by the binding-based approach. Many thanks to P. Schlenker for bringing this analysis to my attention. Note that this analysis has to cope with the problems mentioned in fn. 6.

person pronoun shifts (and vice versa). In (80a), the same point is made with the indexical *inja* (here).

- (80) a. Un ketābe ke, vaghti pāris boodim, Leilā goft az injā xaridam
 The book that when Paris were-1PL L said from here bought-1SG
 ine.
 this-is.
 ‘The book that, when we were in Paris_j, Leila_i told me she_i bought there_j
 is this one.’
- b. Un ketābe ke, dafeie pish ke didimesh, Leilā goft dirooz
 The book that time previous that saw-1PL-3SG, L said yesterday
 xaridam ine.
 bought-1SG this-is
 ‘The book that, last time that we saw Leila_i, she_i told us that she bought
 the day before is this one.’

Next, let us note that in Farsi, as well as in English, the term *Monday* (*doshanbe* in Farsi) cannot be used synonymously with *today* if the utterance is made on a Monday.

- (81) Context: *The conversation is taking place on a Monday.*
- a. Has Leila arrived already?
Leilā reside?
- b. Yes, she arrived {#Monday morning, today morning}.
Āre, {#doshanbe sobh, emruz sobh} resid.

Constructing similar examples with the indexical *here* is more complicated due to the fact that the meaning of this expression is usually underspecified; *here* could mean this room, this building, this city, this country, and so on. That said, the following discourse can be used to motivate the claim that a location-referring term cannot be used in a context in which it would be unquestionably synonymous with *here*.

- (82) Context: *Amir is having a sandwich in a less-than-fancy street-food joint in Tehran, called Zapata. Mina walks in . . .*
- a. **Mina:** Where is Leila?
Mina: *Leila kojāst?*
- b. **Amir:** She came {#to Zapata, (in) here} a few minutes ago to buy a sandwich.
Amir: *Hamin alan umad {#Zāpātā, injā} ie sāndevich begire.*²⁸

The next step is to verify that a non-indexical adverbial can be used to refer to the actual location or time of utterance if it is embedded in a context-shifted environment. The following examples show that this is indeed the case.

- (83) Context: *The conversation is taking place on a Monday.*
- a. Cherā bāhāsh vasate hafte qarār gozāshti?
 Why with-3SG middle week meeting put-2SG

²⁸Note that the *Zapata*-variant of the example becomes fully acceptable if Mina, instead of walking in, is somewhere else and calls Amir on the phone.

- ‘Why are you meeting her in the middle of the week?’²⁹
- b. Khodesh goft faqat doshanbe vaght dāram.
 SELF-3SG said only Monday time have-1SG
 ‘She herself told me that she only has time on Monday.’
- (84) *Context: The conversation is taking place between Amir and Mina in a less-than-fancy street-food join in Tehran, called Zapata.*
- a. **Mina:** Cherā bā Leilā injā qarār gozāshti?
 Why with Leila here meeting placed-2SG
 ‘Why are you meeting Leila here?’
- b. **Amir:** Khodesh goft mixām biyām Zāpātā!
 SELF-3SG said want-1SG come.INF-1SG Zapata
 ‘She told me herself that she wants to come here.’

The final step is to verify that a non-indexical form cannot be used to refer to the reported location or time of utterance, if it is embedded in a context-shifted environment. I have not been able to construct an adequate example with locatives; there always seems to be a potential confound, due ultimately to the under-specificity of *here*. But the following example provides evidence that the generalization holds at least for temporal adverbials.

- (85) a. *Context: The addressee has a regular appointment with his doctor on Mondays.*
- b. #hafteie pish, āghāie doktor behet goft doshanbe ba’d az zohr bā
 Week last, mr. doctor to-2SG said Monday after PREP noon with
 ki qarār dāram?
 who meeting have-1SG
 ‘Last week, who did the doctor tell you he has a meeting with on the Monday afternoon?’

I conclude that non-indexical adverbials are subject to the same constraints as third person NPs: (i) expressions in neither category can be used to refer to the coordinates of the actual speech context, unless embedded in a context-shifted environment, and (iii) expressions in neither category can be used to refer the coordinates of the reported speech context if embedded in a context-shifted environment.

5.9 Conclusion

In this paper, the predictions made by a *prima facie* plausible competition-based principle pertaining to the use of non-indexical noun phrases in conjunction with the phenomenon of indexical shift were made explicit and corroborated, building on data from Farsi. For the most part, no modification, be it on the part of the competition-based principle or the standard account of indexical shift, was required to capture data. However, certain problematic cases involving the ‘anti-logophoricity’ requirement of non-indexical NPs in context-shifted environments, motivated a modification to the competition-based account. As a result of this modification, the principle no longer encodes a bias in favor of indexical expression on the basis of the meaning of the alternative utterances at root; rather, it encodes a local requirement to use an

²⁹The reader should keep in mind that in Iran the week starts on Saturday.

indexical NP if it happens to be covalued with a non-indexical alternative, given the particular syntactic configuration. This move brings the principle one step closer to binding theory, understood as a set of constraints that regulate the distribution of referential expressions. Interestingly, it was also noted that the same pattern can be constructed on the basis of temporal and locative adverbials. This latter finding opens up several avenues for future research. Here, I will mention one.³⁰ Extrapolating on the basis of indexical pronouns and third person NPs, one might think that the nature of the competition is 'featural' in the sense that indexical pronouns are preferred to non-indexical NPs because first and second person features are marked while the third person is not. The finding that, whatever the nature of the competition may be, it applies to (non-)indexical adverbials strongly suggests that, *contra* the view just sketched, the competition is between indexical expressions as such and their non-indexical counterparts. Yet the question remains open, why is it that indexical expressions as such are preferred? I have to leave this question to future work.

³⁰A related but different line of work would involve Free Indirect Discourse. It is well known that in Free Indirect Discourse, there is a striking asymmetry between indexical (tense and) pronouns, on the one hand, and indexical adverbs on the other; specifically, in Free Indirect Discourse temporal and locative adverbials are interpreted with respect to the context of *reported* speech / thought, i.e., are shifted, while indexical (tense and) pronouns maintain their reliance on the actual context of utterance / thought, i.e., cannot be shifted. This raises two questions. First, do indexical and non-indexical expressions compete in Free Indirect Discourse in the same way that they do in normal discourse? Second, if so, does this competition apply to indexical pronouns and adverbials uniformly or not? This line of investigation, coupled with the findings of this paper, promises to provide some insight both into the semantics of indexical pronouns and adverbials and the nature of mechanisms that underly indexical shift on the one hand and Free Indirect Discourse on the other.

Appendices

.1 Exceptions to the Ban Against Illeism

In this appendix several apparent counter-examples to the generalization introduced in (3), repeated below for easy reference, are discussed. My conclusion based on the following (admittedly sketchy discussion) is that none of these cases threaten the substance of the discussion in the paper although some of these cases merit closer scrutiny in the connection to the claims made in the paper.

- (86) Using third person NPs to refer to the speaker or the addressee of the utterance is unacceptable. (= (3))

First, (3) can be obviated when the relevant material is quoted, compare (1b) and (2b) with the corresponding sentences in (87). It should be clear that (3) is meant as a constraint on how NPs are *used*, not *mentioned*.

- (87) a. **[Jimmy:]** Elaine said, “Jimmy is hungry”.
 b. **[Jimmy to Jerry:]** Elaine said, “Jerry is an idiot”.

Second, (3) does not apply to non-referential uses of noun phrases. This includes in particular the predicative uses of post-copular proper nouns, such as that in (88). As (3) only pertains to NPs that refer to individuals, felicity of this kind of example is not relevant to it.

- (88) **[Jimmy:]** My name is Jimmy.

Another type of example based on nominal appositives is given in (89).³¹ Putting aside the fact that such constructions are not naturally used in day to day discourse, I believe in the last analysis they can be plausibly assimilated either to examples like (87) or (88).

- (89) **[Jimmy to Elaine:]** I, Jimmy, take thee, Elaine, as my lawfully wedded wife.

Third, identity statements pose an interesting contrast. Cases like (90a), which amount to *nominal* identification, escape (3) while cases like (90b), which amount to *individual* identification, do not. It seems plausible that (90b) should be taken as baseline and (90a) be considered acceptable for orthogonal reasons, perhaps analogous to (88) (e.g., the felicitous interpretation of (90a) might amount to something like *I am the person whose name is Jimmy*).

- (90) a. **[Jimmy:]** I am Jimmy.
 b. # **[Jimmy:]** this _{π} is Jimmy.
 (π : pointing at one of several children in a photograph)

Fourth, and finally, *imposters* (Collins & Postal 2012) can be defined as a first approximation as those NPs that violate (3) *by design*; that is, those grammatically third person NPs that *specialize* in referring to the speaker / addressee of the utterance, as illustrated in (91a) and (91b).³² The felicitous uses of proper nouns to refer to the

³¹Thanks to Amy Rose Deal (p.c.) for bringing this type of example to my attention.

³²I am grateful to audiences at NYU Semantics Group for bringing the potential relevance of imposters to my attention.

actual speaker / addressee, such as (91c), presumably fall in the same category.³³

- (91) a. *Yours truly* practically screamed {her, *my} tonsils off.
 b. {Is, *are} *your highness* hungry?
 c. [**Nixon:**] You don't have Nixon to kick around anymore.

The potential relevance of imposters to the present discussion depends ultimately on one's analysis of their representation and interpretation (see Collins & Postal 2012 for a syntactic proposal and Podobryaev 2014 for a semantic one). For example, according to Collins & Postal's analysis, which is inspired by examples like (89) discussed above (referred to by Collins & Postal as 'imposter precursors'), the underlying (presumably LF) structure of (91c) already contains a first person pronoun. If so, then the gap between the surface form of these constructions (which lacks overt indexicals) and their underlying representation (which contains semantically interpreted indexical locutions) makes it non-trivial to see what prediction (3) makes in such cases. I have to leave this task to future research.

³³Collins & Postal explicitly categorize proper nouns as imposters but fail to discuss the fact that although proper nouns can be used felicitously as imposters in certain cases, it is in general infelicitous to use them as such, cf. (1a) and (91c) (see also fn. 3). A *prima facie* plausible assumption is that the 'imposterous' uses of proper nouns involve a particular mode of representation which, for unknown reasons, is not always available.

Chapter 6

Copy Theory of Movement and Indexical Shift in Farsi

6.1 Introduction

‘Indexical shifting’ is the phenomenon whereby an occurrence of an indexical expression in indirect discourse gets its value from the *reported*, not the *actual*, context of utterance (see Deal 2017, for a recent overview and references therein). For example, consider the Farsi sentence in (1).¹

- (1) Leilā be Minā goft kudum dāneshju-t-o dus-dāra-m?
L to M told pro which student-2sg-OM love-have-1sg

Reading 1: ‘Which of your students did Leila tell Mina that I love?’

Reading 2: ‘Which of Mina’s students did Leila_i tell Mina that she_i loves?’

In (1) there are two indexical pronouns in the scope of the speech-report predicate *goftan* (to say / tell): there is the second person pronoun in the restrictor of *kudum* (which), which surfaces as a clitic on the noun *daneshju* (student) giving rise to a possessive reading (your student), the *kudum*-phrase itself occupying the object position of the embedded complex predicate *dust dashtan* (lit. to have love), and there is the null first person pronoun which occupies the subject position of the embedded complex predicate triggering first person agreement on its light verb. As indicated, this string allows for at least two readings, depending on whether the embedded indexicals refer to the actual speaker / addressee (Reading 1) or to the reported speaker / addressee (Reading 2).² Importantly, ‘mixed’ readings are not allowed: either both indexicals ‘shift’ or neither do.

Overtly fronting the *kudum*-phrase in (1), as in (2) below, introduces a reading which is not available for (1). In the new reading the first person pronoun in the embedded clause shifts as before but the second person pronoun in the restrictor of *kudum* ‘unshifts’, i.e., refers to the actual addressee. From now on, this reading will

¹ ‘OM’ stands for the object marker *-ra* which in colloquial speech generally surfaces as *-o*.

² As *goftan* is a responsive predicate, (1) also has a reading in which *goftan* embeds the *kudum* question. Interestingly, this is only possible if the embedded indexicals do *not* shift (*Leila told Mina which of your students I hate*): if there is indexical shifting, the embedded question reading is impossible (**Leila_i told Mina_j which of her_j students she_i hates*). This issue is not relevant here as we will focus exclusively on matrix questions.

be referred to as the unshifted reading.³

- (2) kudum dāneshju-t-o Leilā be Minā goft __ dus-dāra-m?
 Which student-2sg-OM L to M told pro __ love-have-1sg
Unshifted reading: ‘Which of your students did Leila_i tell Mina that she_i loves?’

A proper analysis of (2) requires making specific assumptions about the syntax and semantics of movement and of indexical shifting. Building on data from Farsi, one contribution of this paper is the observation that a naïve combination of the copy theory of movement (Chomsky 1995 and much subsequent literature) with the so-called operator-based approach to indexical shifting (Anand & Nevins 2004; Anand 2006) leads to over-generation problems. In response to this observation, the second contribution of this paper is develop two techniques for solving these problems. On the one hand, I will show that if, broadly in line with certain proposals in the literature (Fox 1999; Merchant 2000), copy theory is coupled with a severely restricted deletion operation at the LF interface which can target expressions of *any size* (down to individual features) and if suitable assumptions are made to allow for the LF representation of indexical pronouns as variables in a manner compatible with indexical shifting, then one of the two over-generation problems can be solved in a principled manner but the other one remains open. On the other hand, I will argue that if the variable-based account of indexical shifting (Schlenker 1999, 2003; von Stechow 2004) is adopted instead of the operator-based account, a minor modification of an independently needed stipulation (namely, Anand’s (2006) SHIFT TOGETHER! constraint) allows us to solve both problems without making any other changes to copy theory.

This paper is organized as follows. Section 6.2 contains the relevant background on Farsi and indexical shifting. In Section 6.3, the problems that arise when copy theory and the operator-based account of indexical shifting are combined are discussed in detail. In Sections 6.4 and 6.5, two potential solutions to these problems are discussed, one based on the operator-based account of indexical shifting and one based on the variable-based account, and Section 6.6 concludes. In the rest of the present section a summary of the content of the paper is provided, leaving blanks to be filled in the following sections.

According to copy theory (hf. CT), a moved expression leaves behind, not a trace, but a full-fledged copy of itself. According to the operator-based account of indexical shift (hf. OP), the shifted reading of indexicals is due to a covert, ‘context-shifting’ operator that inhabits the ‘left periphery’ of the clausal argument of certain attitude verbs. The combination of these two theories, call it CT+OP, therefore, predicts that

³ Reading 1 of (1), i.e., the reading in which no indexical shifts, is, naturally, still possible with (2). Some but not all speakers can access the reading of (2) that would correspond to Reading 2 of (1), i.e., the reading in which both indexicals, including the one embedded in the restrictor of *kudum* shifts. As the latter reading presumably involves reconstruction, one can account for the inter-speaker variation on the assumption that there is some pressure in this language for the scope of operators to be ‘surface true’. Importantly, for those speakers who can access the reconstructed reading the following generalization holds: if the second person indexical in the restrictor of *kudum* shifts then the first person indexical in the embedded clause *must* shift as well, while the reverse does not hold. In other words, the output of reconstruction must obey SHIFT TOGETHER! as well. I will not be concerned here with the significance of this observation *vis-à-vis* the theory of reconstruction, although note that this observation is *prima facie* not compatible with a semantic account of reconstruction.

the schematic LF in (3) is a possible LF for (2).⁴

(3) [which ... you ...] L told M [OP I [which ... you ...] love]

The first problem with CT+OP is that (3) does not correspond to a reading that (2) intuitively can have. The core of the problem is that the two copies of the second person indexical are predicted to be interpreted with respect to different contexts: the higher copy is predicted to refer to the actual addressee while the lower copy, being in the scope of the context-shifting operator, is predicted to refer to the reported addressee. Let us refer to (3) as the ‘double-access’ LF.⁵ As the double-access LF does not denote a contradictory, or otherwise semantically defective, proposition (this will be discussed Section 6.3), the first problem for CT+OP is to somehow block it, whence the injunction in (4).

(4) Find a way to block the double-access LF (i.e., (3)) as a possible LF of (2)!

The second problem of CP+OP is that it incorrectly predicts an interaction between Condition C and indexical shift. To see this, consider the problem of finding a way to derive the unshifted reading of (2) from the double-access LF in (3). Evidently, the material in the restrictor of the based-generated copy of *kudum* needs to be neutralized: in order to derive the unshifted reading, (3) needs to be transformed into (5).

(5) [which ... you ...] L told M [OP I [which ~~you~~] love]

But if the only way to unshift the indexical in the restrictor of *kudum* is to delete the restrictor of its lower copy, then we predict that unshifting should have consequences for Condition C. For example, we predict that the matrix subject pronoun in (6) should be able to refer to Leila without violating Condition C, if the first person pronoun in the embedded clause is shifted and the second person pronoun in the restrictor of *kudum* is unshifted.

(6) *[*kudum* *dāneshju-ie-to-o-Leilā*]-ro be *Minā* goft ___ *dus-dāra-m?*
 Which student-EZ-2sg-and-L-OM pro to M told pro ___ love-have-1sg
Intended reading: ‘Which *x* is such that *x* is a student of you and Leila’s and Leila_{*i*} told Mina that she_{*i*} loves *x*?’

To see this, note that in order to derive the reading of (6) that corresponds to the unshifted reading of (2), the LF in (7a) must be transformed into (7b) via deletion of the restrictor of the lower copy, same as the (3) to (5) transformation. If so, then, on the target reading, the matrix subject pronoun (and, therefore, also the embedded first person pronoun which is shifted) should be able to refer to Leila without violating Condition C as the lower copy of the r-expression *Leila* is deleted and, hence, is not c-commanded by the matrix subject pronoun.

(7) a. [which ... you+Leila ...]

⁴ In order for LFs of this kind to be interpretable, some manipulation or other must be performed on the lower copy in order to generate a sensible operator-variable construction. The standard way to do this is via ‘trace conversion’ (Fox 2002).

⁵ The label is borrowed from the literature on tense. Its use in this paper does not rest on any substantive assumptions or claims.

- pro_x told M
 [OP I [which ... you+Leila ...] love]
- b. [which ... you+Leila ...]
 pro_x told M
 [OP I [which ~~... you+Leila ...~~] love]

This prediction is incorrect: the sentence is judged acceptable only if the null-pronoun subjects of the matrix and embedded predicates do not refer to Leila, a fact that can be explained on the assumption that the sentence generates a Condition C violation otherwise. The second problem of CP+OP can be summarized with the following injunction.

(8) Find a way to derive the unshifted reading without bleeding Condition C!

If the argument that a naïve combination of CT and OP generates empirical problems is on the right track then *a priori* there are several ways to proceed. The three most salient options are,

1. Keep CT and OP, but introduce auxiliary assumptions, or,
2. Replace OP but keep CT intact, or,
3. Replace CT but keep OP intact.

Starting with the last option, one could indeed simply discard CT and adopt trace theory in conjunction with OP. The problems mentioned above being direct consequences of having (\bar{A} -)movement leaving behind structurally complex objects, adopting trace theory does not so much solve them as make them disappear entirely. In other words, one may take the problems mentioned above as evidence against copy theory, in favor of trace theory, keeping the operator-based account of indexical shift constant.⁶ The switch to trace theory, however, veers the locus of discussion back to the arguments that motivated copy theory to begin with and this is a different enterprise altogether.⁷ As such, consideration of option 3 is beyond the scope of the present paper. I will instead dogmatically assume that copy theory is on the right track and explore the two alternative options 1 and 2 in sections 6.5 and 6.4 respectively.

In Section 6.5, neither CT nor OP are disposed of but auxiliary assumptions are introduced on both sides to solve the problems mentioned above. Working our way backwards, the core idea of the proposal is that the unshifted reading of (2) results from an LF that roughly looks like (9), where the pronoun in the *kudum*-restrictor of the lower copy is represented as a (feature-less) variable which is coreferential with the pronoun in the higher copy on which the second person feature is interpreted. Note that since the pronoun in the lower copy is simply a variable *without an* indexical feature its denotation is not controlled by the context-shifting operator which only affects the meaning of *indexical* expressions in its scope.⁸

(9) [which ... you_x ...] L told M [OP I [which ... x ...] love]

⁶ Thanks to Danny Fox (p.c.) for raising this issue and for helpful discussion.

⁷ See Takahashi 2010a,b for a review of the facts that motivate the claim that at least \bar{A} -movement obligatorily involves leaving copies in the pre-movement site, and references therein.

⁸ I am grateful to Amy Rose Deal (p.c.) who suggested the core idea of this analysis.

The challenge, of course, is deriving (9) from (3) in a principled manner. It will be proposed that two sets of assumptions suffice to accomplish this. First, on the OP side, certain assumptions need to be taken on board to allow indexical pronouns to be represented at LF as ‘variables’ with feature-bundles, as in (10a), as opposed to ‘constants’, as in (10b).⁹ As discussed in Section 6.5 (following up on relevant discussion in Schlenker 2003, 2004), if the entry in (10a) is adopted without further ado, a presupposition failure is incorrectly generated in certain cases where an indexical pronoun is c-commanded by the context-shifting operator, forcing the standard accounts of indexical shifting to stipulate that shifty indexicals are constants. As discussed there, this problem needs to be solved somehow for the account under discussion to work, and I will sketch a proposal based on Schlenker 2004.

- (10) a. $\llbracket [2nd\ pro_i] \rrbracket^{c,g} = \#$ unless $g(i) = ADDR(c)$; if defined then
 $\llbracket [2nd\ pro_i] \rrbracket^{c,g} = g(i)$.
 b. $\llbracket 2nd\text{-}pro \rrbracket^c = ADDR(c)$

The result of representing shifty indexicals as restricted variables is that the original LF of (2) now looks like this (cf. (3)). Notice that we are now one step closer to our target LF which was schematically represented in (9). If we incorporate a mechanism that can delete the indexical feature of the lower copy, the desired LF is generated.

- (11) [which ... [2nd pro_x] ...]
 L told M
 [OP [1sg pro_y] [which ... [2nd pro_x] ...] love]

This brings us to the second set of assumptions, this time on the CT side. It will be argued, following proposals in the literature, that CT must be coupled with a deletion operation at the LF interface which (a) can target expressions of any size (down to individual features), (b) is subject to an economy constraint that minimizes its application (in line with Fox (1999)); specifically, it can only target an expression if deleting that expression has semantic consequences of a particular kind, and (c) can only target the *smallest* possible expression that yields the designated interpretation (in line with Merchant (2000)). A deletion operation subject to these assumptions allows the grammar to delete the second person *feature* of the lower copy of the indexical pronoun, getting us from (11) to (12).¹⁰

⁹ In using terms ‘variable’ and ‘constant’ I have in mind the idea that the denotation a variable is assignment-dependent while the denotation of a constant is not. Since the denotation of an indexical (unlike, say, that of proper nouns) varies across contexts as a matter of definition, this terminology might be confusing. Furthermore, this terminology might be entirely inadequate as constants can be routinely re-analyzed as *prima facie* assignment-dependent expressions. Thus, instead of (ia) one could adopt (ib) on which the denotation of *John* is, strictly speaking, assignment-dependent.

- (i) a. $\llbracket John \rrbracket^{g,\dots} = \mathbf{John}$
 b. $\llbracket John_i \rrbracket^{g,\dots} = \#$ unless $g(i) = \mathbf{John}$; if defined, $\llbracket John_i \rrbracket^{g,\dots} = g(i)$

One could, however, adopt the following definition of assignment-dependency which side-steps this problem: the denotation of expression α is assignment-dependent iff there are assignments g_1, g_2 and context c and index i such that $\llbracket \alpha \rrbracket^{c,i,g_1}$ and $\llbracket \alpha \rrbracket^{c,i,g_2}$ are both defined but $\llbracket \alpha \rrbracket^{c,i,g_1} \neq \llbracket \alpha \rrbracket^{c,i,g_2}$. The reader can verify that on this definition *John* in (ib) is not assignment-dependent.

¹⁰There are various proposals in the literature that rely on ‘deletion under agreement’, the rough idea being that sometimes features on *bound* pronouns can remain semantically uninterpreted. The reader should bear in mind that in (12) the pronoun in the lower copy is *not* bound; therefore, the

- (12) [which ... [2nd pro_x] ...]
 L told M
 [OP [1sg pro_y] [which ... [~~2nd~~ pro_x] ...] love]

This LF then allows us to derive the unshifted reading by deleting only the indexical feature on the pronoun in the lower copy, which in turns allows us to make sure that Condition C need not be bled as no expression larger than the person feature can now be deleted. Thus the second problem (i.e., (8)) is solved. Unfortunately on this approach the double-access LF is not blocked. The first problem (i.e., (4) above), then, remains open. One possibility is that the feature deletion exemplified in (12) is in this case *obligatory* after all for some reason or other. The underlying generalization might be that ϕ -features on pronouns that occur in chains are always interpreted just once, perhaps only on the head of the chain. I am not aware of an independent motivation for this claim. An alternative, perhaps more conservative, line would be that the double-access LF is in fact generated by the grammar but the reading that it gives rise to, while semantically coherent, cannot be intuitively accessed either due to contextual / pragmatic factors or processing cost.

In Section 6.4, in the spirit of option 2, it will be argued that the variable-based account of indexical shifting (Schlenker 1999 and subsequent work), coupled with a minor modification of an independently needed stipulation, can be combined with the original version of copy theory without generating either of the two problems mentioned above. In a nutshell, according to the variable-based account of indexical shifting, LFs have context abstractors explicitly represented and shifty indexicals have a context variable that can be bound by either the matrix abstractor (unshifted reading) or an embedded abstractor introduced by certain predicates (shifted reading). Now, it is well-known that canonical implementations of the variable-based system suffer from an over-generation problem and, consequently, need to be constrained by explicit stipulations. One such stipulation is SHIFT TOGETHER! (Anand & Nevins 2004; Anand 2006) which says that indexical expressions in the same ‘minimal domain’ must be bound by the same context abstractor, where the minimal domain of an indexical can roughly be taken to be the smallest CP / TP that dominates it. We modify this latter definition in only one respect: the minimal domain of an indexical that occurs in a chain is the determined by (i.e., is the same as) the minimal domain of *its highest copy*. One side-effect of movement, then, is to extend the minimal domain of those indexicals that occur in the moved constituent. This modification solves both of the problems mentioned above in one move. The first problem (i.e., (4)) is solved because different copies of the same indexical are now assigned the same minimal domain and as such are forced by SHIFT TOGETHER! to be bound by the same context abstractor (‘vertical shift together’). The second problem (i.e., (8)) is solved because deriving the unshifted reading no longer *requires* deletion of any material in the lower copy and therefore, by economy, deletion is prohibited to apply in these cases, ensuring that Condition C is not bled. Indeed this account is compatible with a view of the grammar in which the grammar does not have access to a deletion operation at all.

It thus appears that the variable-based approach has an advantage over the operator-based approach as far as the data in this paper are concerned. It should be emphasized, however, that the variable-based solution comes at the high price of discarding the operator-based approach altogether, raising worries about the possi-

deletion operation involved in (12) cannot be subsumed under the same rubric.

bility of having thrown the baby out with the bath-water (for a recent review of the main arguments in favor of the operator-based approach see Deal 2017). It might be, then, that the first approach sketched above is in fact on the right track after all and the double-access problem has an independent solution. We leave the resolution of this dilemma to future work.

6.2 Farsi and indexical shift

Farsi (or Persian, Indo-Iranian) has SOV word order although clause-embedding predicates always precede their clausal complements. Furthermore, beyond the verbal domain Farsi is head-initial.

- (13) a. (Leilā) Ali ro zad.
Leila Ali OM hit
'Leila hit Ali.'
- b. (Minā) goft (Leilā) Ali ro zad.
Mina said Leila Ali OM hit
'Mina said that Leila hit Ali.'

Farsi is a *wh*-in situ language (but see Kahnemuyipour 2001) in which *wh*-phrases can be optionally fronted, sometimes with truth-conditional consequences (cf. (1) and (2) in the previous section).¹¹

- (14) a. Ki goft Leilā Ali ro zad?
who said Leila Ali OM hit
'Who said that Leila hit Ali?'
- b. Minā goft ki Ali ro zad?
Mina said who Ali OM hit
'Who did Mina say hit Ali?'
- c. Minā goft Leilā ki ro zad?
Mina said Leila who OM hit
'Who did Mina say Leila hit?'
- d. ?Ki Minā goft __ Ali ro zad?
Who Mina said __ Ali OM hit
'Who did Mina say hit Ali?'
- e. Ki ro Minā goft Leilā __ zad?
Who OM Mina said Leila __ hit
'Who did Mina say Leila hit?'

Farsi is a *pro*-drop language with obligatory subject-verb agreement. Throughout I will assume dropped subjects are represented at LF with a covert pronominal element which enters into agreement with the predicate and has interpretable ϕ -features.

- (15) a. (man) xāb-am
1sg slept-1sg
'I slept.'

¹¹Fronting is somewhat marked with *wh*-phrases that originate in the subject position, cf. (14d) and (14e).

- b. (mā) xāb-im
1pl slept-1pl
'We slept.'
- c. (to) xāb-i
2sg slept-2sg
'you slept.'
- d. (shomā) xāb-in (or, (shomā) xāb-id)
2pl slept-2pl
'You slept.'
- e. (Un) xāb-e
3sg slept-3sg
'He / She slept.'
- f. (unā) xāb-an
3sg slept-3sg
'They slept.'

The pronominal system has two sub-paradigms consisting of full pronouns (shown above) and enclitics. The enclitics can attach, e.g., to prepositions and to nouns, in the latter case generating possessive readings.¹² Either way, only person and number are grammatically marked.

- (16)
- a. Ketāb-am
book-1sg
'My book'
 - b. ketāb-emun
book-1pl
'Our book'
 - c. ketāb-et
book-2sg
'Your book'
 - d. ketāb-etun
book-2pl
'Your book'
 - e. ketāb-esh
book-3sg
'His / Her book'
 - f. ketāb-eshun
book-3pl
'Their book'

There is no object-verb agreement *per se* but the enclitics can attach to inflected verbs, a process which is obligatory if the object is dropped but optional otherwise.¹³

- (17)
- a. Man Ali ro zad-am-(?esh)
1sg Ali OM hit-1sg-3sg
'I hit Ali.'

¹² I assume the 'ezafe' clitic *-e* must be syntactically present between the enclitic, in which case, e.g., *ketaab-emun* must be parsed as *keetab-e-mun*. I abstract away from this as it is not pertinent.

¹³ Enclitics that double the object are in general somewhat marked.

- b. Man zad-am-*(esh)
 1sg hit-1sg-3sg
 ‘I hit him / her.’

Turning to indexical (i.e., first and second person) pronouns, as pointed out in the previous section these can appear as full pronouns (18a), as enclitics (18b), or as null pronouns (18c).

- (18) a. Man Ali (r)o dus-daram.
 ‘I like Ali’
 b. Xāharam Ali (r)o dus-dare.
 ‘My sister likes Ali’
 c. \emptyset Ali (r)o dus-daram.
 ‘I like Ali’

All three forms give rise to an ambiguity when embedded under the speech-report predicate *goftan*, depending on whether they take their value from the actual or the reported context of speech.¹⁴

- (19) a. Mina goft man Ali ro dus-daram.
Reading 1: ‘Mina said that I like Ali’
Reading 2: ‘Mina_i said that she_i likes Ali’
 b. Mina goft xāharam Ali ro dus-dare.
Reading 1: ‘Mina said that my sister like Ali’
Reading 2: ‘Mina_i said that her_i sister likes Ali’
 c. Mina goft \emptyset Ali ro dus-daram.
Reading 1: ‘Mina said that I like Ali’
Reading 2: ‘Mina_i said that she_i likes Ali’

When two indexicals occur in the complement of *goftan* either both shift (21b) or neither does (21a). In other words, Farsi respects Anand’s (2006) SHIFT TOGETHER! constraint which is informally given in (23).

- (20) Leilā be Minā goft duset daram.
 L to M said love-2sg have-1sg
- (21) Possible readings for (20):
 a. Leila told Mina that I love you
 b. Leila_i told Mina_j that she_i loves her_j
- (22) Impossible readings for (20):
 a. Leila told Mina_j that I love her_j
 b. Leila_i told Mina that she_i loves you

¹⁴Some consultants report that the shifted reading with full pronouns, e.g., Reading 2 of (19a), is somewhat marked particularly in comparison with (19c). The important point, for my purposes, is that full pronouns do shift. Evidence for this comes from constructions in which the null form is barred and only the full form can be used such as cases in which the pronoun is stressed. The shifted reading can be accessed in such cases very easily.

- (i) Minā goft faghat man_F Ali ro dus-dāram.
 Mina said only 1sg Ali \emptyset M love-have
Possible Reading: ‘Mina_i said that only she_i likes Ali.’

- (23) a. SHIFT TOGETHER! Indexical expressions in the same minimal domain must get their value from the same context.
 b. The minimal domain of a particular occurrence of an indexical expression is the smallest CP / TP that contains it.

Note that extraction from the complement of *goftan* is possible even when there are shifted indexicals in the embedded clause. (Here too SHIFT TOGETHER! must be respected.) This suggests that the shifted reading of indexicals in Farsi is not due to quotation or any other process that generates an environment opaque to grammatical transformations.

- (24) a. Leilā be Minā goft p̄r̄ ketāb-e ki ro barā-t xaridam?
 L to M told pro book-EZ who OM for-2sg bought.1sg
 ‘Whose book did Leila_i tell Mina_j she_i bought for her_j?’
 b. Ketāb-e ki ro Leilā be Minā goft p̄r̄ __ barā-t xaridam?
 book-EX who OM L to M told pro __ for-2sg bought.1sg?
 ‘Whose book did Leila_i tell Mina_j she_i bought for her_j?’

Finally, shifted indexicals are obligatorily interpreted *de se*. Thus (25a) can be felicitously and truthfully used to report the situation given in (26), but the sentence in *shifted sentence*, on its shifted reading, cannot.

- (25) a. Doktor be parastār goft hale-sh xeili bad-e.
 Doctor to nurse told condition-3sg very bad-is
 Possible: ‘The doctor_i told the nurse that he_i is in a bad condition.’
 b. Doktor be parastār goft hala-m xeili bad-e.
 Doctor to nurse told condition-1sg very bad-is
 Not possible: ‘The doctor_i told the nurse that he_i is in a bad condition.’
 (26) The doctor has been reviewing patients’ files, which are anonymous. Unbeknownst to him, his own file is among the files that he is reviewing. After studying his file the doctor turns to the nurse and says, “this patient is in a very bad condition.”

The standard analysis of indexical shifting, due to Anand & Nevins 2004; Anand 2006, rests on the following assumptions (various complications are ignored for the sake of exposition). The interpretation function is parametrized to a context *c*, an index (of evaluation) *i*, and an assignment function *g*, $\llbracket \cdot \rrbracket^{c,i,g}$. Contexts are taken to be tuples with various coordinates. For our purposes, we abstract from the temporal and locative information and assume that contexts consist of an individual author / speaker AUTH(*c*), an individual addressee (possibly plural) ADDR(*c*), a possible world WORLD(*c*). It is assumed that indices are ‘homologous’ to contexts, thus an index *i* also has a speaker AUTH(*i*), an addressee ADDR(*i*), a possible world WORLD(*i*). Within this framework a speech-report predicate like *goftan* (to tell / say) is interpreted as in (28) and indexical pronouns are assigned the lexical entries in (27).¹⁵ Note that, unlike third person pronouns, indexical pronouns are assignment-independent. We will revisit this difference in Section 6.5.

- (27) a. $\llbracket I \rrbracket^{c,i,g} = \text{AUTH}(c)$
 b. $\llbracket \text{you} \rrbracket^{c,i,g} = \text{ADDR}(c)$

¹⁵The entries in (27) will be modified in Section 6.5.

- c. $\llbracket \text{she}_x \rrbracket^{c,i,g} = \#$ unless $g(x)$ is female in $\text{WORLD}(c)$; if $\neq \#$ then $\llbracket \text{she}_x \rrbracket^{c,i,g} = g(x)$.
- (28) a. $\llbracket x \text{ tell } y \text{ that } \phi \rrbracket^{c,i,g} = 1$ iff $\forall i' \in \text{SAY}_{x,y}^i : \llbracket \phi \rrbracket^{c,i',g} = 1$
 b. $i' \in \text{SAY}_{x,y}^i$ iff i' is compatible with what x told y in $\text{WORLD}(i)$.

The final ingredient is the context-shifting operator, OP , which over-writes the context parameter with the index parameter (this is why the assumption is made that contexts and indices are homologous). A syncategorematic analysis of this operator is given in (29).

$$(29) \quad \llbracket \text{OP } \phi \rrbracket^{c,i,g} = \llbracket \phi \rrbracket^{i,i,g}$$

The assumptions enumerated above coupled with the assumption that speech-report predicates in certain languages (such as Farsi) license the context-shifting operator in their scope allow us to derive not only the shifted reading of indexicals but also the fact that shifted indexicals must be interpreted *de se* and the **SHIFT TOGETHER!** constraint.

- (30) a. $\llbracket x \text{ tell } y \text{ [I hate you]} \rrbracket^{c,i,g} = 1$ iff
 $\forall i' \in \text{SAY}_{x,y}^i : \llbracket \text{I hate you} \rrbracket^{c,i',g} = 1$ iff
 For all i' that is compatible with what x told y in $\text{WORLD}(i)$, $\text{AUTH}(c)$ hates $\text{ADDR}(c)$ in $\text{WORLD}(i')$.
- b. $\llbracket x \text{ tell } y \text{ [OP I hate you]} \rrbracket^{c,i,g} = 1$ iff
 $\forall i' \in \text{SAY}_{x,y}^i : \llbracket \text{OP I hate you} \rrbracket^{c,i',g} = 1$ iff
 $\forall i' \in \text{SAY}_{x,y}^i : \llbracket \text{I hate you} \rrbracket^{i',i',g} = 1$ iff
 For all i' that is compatible with what x told y in $\text{WORLD}(i)$, $\text{AUTH}(i')$ hates $\text{ADDR}(i')$ in $\text{WORLD}(i')$.

6.3 Two problems with copy theory and the operator-based approach

In this section, I will elaborate on the two problems mentioned in Section 6.1. First, I will describe the problem raised by the double-access LF. I will describe the truth conditions that the double-access LF generates, given standard semantic assumptions, showing that the resulting truth conditions are not contradictory or otherwise semantically deficient. Second, I will argue that indexical shifting and Condition C do not interact in Farsi: the unshifted reading mentioned in Section 6.1 does not derive from an LF in which movement bleeds Condition C.

Consider again the sentences in (31a) and (31b). In both sentences a *which*-phrase contains a second person pronoun in its restrictor and the embedded clause contains a first person pronoun. The only structural difference between the two sentences is that in (31b) the *which*-phrase is overtly fronted while in (31a) it remains *in situ*.

- (31) a. Leila_i be Mina_j goft [kudum daneshju-t-o] dus-dara-m?
 L to M told which student-2sg-OM love-have-1sg
 ‘Which of Mina_j’s students did Leila_i tell her_j she_i loves?’
 *‘Which of your students did Leila_i tell Mina she_i loves?’

- b. [kudum daneshju-t-o] Leila be Mina goft e dus-dara-m?
 which student-2sg-OM L to M told e love-have-1sg
 ‘Which of Mina_j’s students did Leila_i tell her_j she_i loves?’
 ‘Which of your students did Leila_i tell Mina she_i loves?’

Interestingly, overt movement of the *which*-phrase in (31b) introduces a semantic ambiguity which otherwise does not exist. In (31a) if the first person pronoun is shifted then the second person pronoun in the restrictor of *which* must shift as well (as predicted by the SHIFT TOGETHER! generalization). Not so in (31b); here, the indexical in the restrictor of *which* may or may not shift (in fact the latter is the more salient option) if the first person pronoun in the embedded clause is shifted.¹⁶

The findings above are summarized in (32).

- (32) a. In a sentence of the form

x told y that [... [which ... indx₁ ...] ... indx₂ ...]

indx₂ shifts if and only if indx₁ shifts as well.

- b. In a sentence of the form

[which ... indx₁ ...] [x told y that [... e ... indx₂ ...]]

if indx₂ shifts, indx₁ may or may not shift but (i) if indx₁ shifts then indx₂ must shift as well and (ii) if indx₂ is not shifted then indx₁ cannot shift either.

From now on, I will focus exclusively on the cases like (31b) where the *which*-phrase moves overtly.

Assuming copy theory of movement, the LFs underlying (31b) would be (33a). To yield an interpretable, operator-variable structure, this LF must undergo trace conversion (Fox 1999) with the result in (33b). Note that the presence of *OP* is dictated by the assumption that the embedded first person (which is represented by *1st-pro*) is shifted.

- (33) a. [which student-of-2sg]
 L to M told *OP*
 [1st-pro [which student-of-2sg] love]
- b. [which student-of-2sg]
 1 [L to M told *OP*
 [1st-pro [the [student-of-2sg =1]] love]]

Even a superficial examination of (33b) immediately reveals two issues. First, the LF in (33b) generates a reading according to which the lower copy of the second person pronoun shifts while the higher copy does not. Second, neither of the two readings in (31b) correspond to the LF in (33b).

Regarding the first issue, the reading that (33b) generates can be called the double-access reading for indexicals: the copy at the tail (being in the scope of *OP*) will be shifted to refer to the reported addressee, Mina, while the copy at the head (being outside of the scope of *OP*) remains unshifted and refers to the actual addressee. I

¹⁶If the first person pronoun in the embedded clause is *not* shifted, then the second person pronoun in the restrictor of *kudum* cannot shift.

will now argue that the double-access reading is not semantically incoherent. To see the truth conditions of the LF in (33b), it suffices that we focus on the constituent that *which* abstracts over.

- (34) 1 [L to M told OP
[1st-pro [the [student-of-2sg =1]] love]]

(34) denotes a set of individuals. Each of these individuals x have the following two properties. First, Leila thinks of x (on some description or other) as a student of the individual that she identifies as Mina. Second, Leila has told Mina that she, Leila, loves x . The first of these properties is the restrult of the presupposition triggered by the definite description [*the [student-of-2sg =1]*] and the second is simply the assertive content of expression. Now, what happens when the set of individuals denoted by (34) is fed to [*which student-of-2sg*] in (33b)? The first task is to determine how the presupposition noted above will project. Let us assume that this projection is universal.¹⁷ The result is the matrix presupposition that every individual in the restrictor of *which* satisfies the presupposition of (34); i.e., every individual x who is one of the actual speaker's students is such that Leila thinks of x (on some description) as a student of Mina's. Moving on to the assertive side, thinking of *which* as an existential quantifier for simplification, we predict the sentence to be true iff there is at least one individual who in fact is one of actual speaker's students, who Leila thinks of as one of Mina's students, and of whom Leila has spoken (very) positively to Mina. Notice that the result of the double-access reading is simply the inference that the relevant domain of quantification consists of individuals who are in fact the speaker's students and are also thought by Leila to be Mina's students. As pointed out to be by Philippe Schlenker (p.c.), the double-access reading can be roughly paraphrased in English as in (35).

- (35) Which student of yours is such that Leila_i told Mina_j that she_i likes this student of hers_j?

As far as I can see, although the truth conditions are certainly rather involved there is no contradiction here. I conclude that the reading that the double-access LF gives rise to is not semantically incoherent. Given that this reading is not intuitively perceived (cf. (35)), either the double-access LF is not generated by the grammar at all or, perhaps, the relevant LF is ruled out for contextual / processing reasons. As the nature of the putative contextual / processing constraints remains unclear to me, I will tentatively assume that the grammar simply does not generate double-access LFs (see also Section 6.4 for further discussion).

Let us now move on to the second issue. Appropriate LFs for the two readings of (31b) can be derived from (33a) via deletion. In (36a) the higher copy is deleted. The result is total reconstruction of the *which*-phrase. Assuming a treatment of *which*-phrases which allows *in situ* interpretation,¹⁸ this LF derives the reading in (31b)

¹⁷My goal in this discussion is to establish that the double-access reading is not semantically defective. Whether the target presupposition projects universally or in a weaker fashion is somewhat beside the point. Note however that if the meaning that results from universal projection is not contradictory, then the meaning that results from any form of weaker projection is not contradictory either.

¹⁸Alternatively, deletion can only target the restrictor of the higher copy. This, coupled with trace conversion, can generate a structure that is compatible with a semantics of *which*-phrases that requires them to be interpreted *ex situ*. The net-result for my purposes is the same.

where both first and second person pronouns are shifted as both pronouns are now c-commanded by *OP*. In (36b) the restrictor of the lower copy is targeted by deletion. As the second person pronoun is now not c-commanded by *OP* it is not shifted while the first person of course is. This corresponds to the other reading of (31b) (i.e., the unshifted reading).

- (36) a. {which student-of-2sg}
 L to M told *OP*
 [1st-pro [which student-of-2sg] loves
 b. [which student-of-2sg]
 1 [L to M told *OP*
 [1st-pro [the [student-of-2sg =1]] loves]]

It would appear, then, that we can derive the desired LFs, specifically for the unshifted reading, if we allow the grammar to have access to a deletion operation that can target material in the lower copy. Evidence against this hypothesis comes from the fact that unshifting does not bleed Condition C, as illustrated in (37). The restrictor of the *which*-phrase in (37) contains the r-expression *Leila* conjoined with the first person pronoun. Crucially, the subject pronoun cannot be co-referential with *Leila*.

- (37) [kudum ketab-e-man-o-Leila_i-ro] *prø_i* be Mina goft [... ____ ...]?
 which book-of-1sg-and-L-OM pro to M told ...

This is puzzling for the approach sketched above as here deleting the lower copy would force the unshifted reading of the first person pronoun, much like (36b) above. But if deletion was possible, the sentence would have been acceptable with the subject *pro* referring to *Leila*, contrary to fact.

6.4 The operator-based approach, deletion and indexical pronouns as restricted variables

The idea that I will pursue in this section is that the LF of a sentence like (38) (= (2) from Section 6.1) has the abstract shape in (39) (= (9) from Section 6.1). The key characteristic of this LF is that the two ‘copies’ of the second person indexical embedded in the restrictor of the moved *which*-phrase have different forms; the one at the tail of the chain is ‘structurally impoverished’ in the sense that it is represented as a feature-less, unrestricted, variable while the one at the head of the chain is represented as a variable with a bundle of features, including in particular the second person feature (throughout, I ignore number and gender). Furthermore, the two copies are semantically related in that the lower variable is (‘accidentally’) co-referential with the higher one.

- (38) kudum daneshju-t-o Leila be Mina goft ____ dus-dara-m?
 Which student-2sg-OM L to M told pro ____ love-have-1sg
 Unshifted reading: ‘Which of your students did Leila_i tell Mina that she_i loves?’
 (39) [which ... you_x ...] L told M [*OP* I [which ... x ...] love]

The crucial question, of course, is how the grammar can generate (39) for (38). One option is to hardwire certain assumptions in the mechanics of the copying / movement operation. For example, one could stipulate that copying a *which*-phrase (or, more generally, any quantificational noun phrase) involves replacing every pronoun that occurs in the ‘original’ copy with a feature-less variable that is co-referential with the corresponding pronoun (with features and all) in the ‘new’, higher copy. Call this the brute force approach. The brute force approach immediately yields the following generalization which, in a sense, is the core intuition behind the analysis explored in this section.

- (40) ϕ -features on pronouns that occur in chains are always interpreted (i) just once, (ii) at the head of the chain.

There is an obvious sense in which the brute force approach, true to name, is completely stipulative as it involves an entirely *ad hoc* manipulation of already constructed structures. Therefore, my aim in the rest of this section is to design an alternative that is less outrageous. Nevertheless it should be stated at the outset that the brute force approach has two desirable consequences. First, the brute force approach solves *both of the two problems* that were discussed in Section 6.3. It predicts that unshifting does not bleed Condition C and it predicts that what I have referred to as double-access LFs are not generated by the grammar at all. Unshifting is predicted not to bleed Condition C because no deletion operation is called upon to remove material in the lower copy. Double-access LFs are not generated because, as pointed out above, indexical features are interpreted *only once*, at the head of the chain, *outside of the scope of the context-shifting operator*. Second, as the brute force approach does not rely on a ‘free standing’ operation of deletion, it is highly restricted in this respect: there is no danger of over-generation by misapplication of deletion and, obviously, no stipulations need to be made regarding the workings of deletion. Neither of these two desirable consequences are easy to capture in less stipulative alternatives to the brute force approach, as far as I can see. In particular, the analysis that I will sketch below runs into an over-generation problem because it cannot block double-access LFs and it relies on a particular deletion operation which must be subjected to very specific constraints at the pain of over-generation. Nevertheless, as discussed at the end of this section, these issues are probably not as problematic as they might appear.

The proposal that I would like to consider as an alternative to the brute force approach relies on two independent sets of assumptions. The core ideas are (i) that shifty indexicals can be represented as variable with feature-bundles and (ii) that there is a deletion operation that can, in certain cases, remove the person features of pronouns in the lower copies. The LF of (38), then, looks more like (41) (= (12) from Section 6.1) than (39) above but it should be clear that the two will end up being semantically equivalent.

- (41) [which ... [2nd pro_x] ...]
 L told M
 [OP [1sg pro_y] [which ... [~~2nd~~ pro_x] ...] love]

Let us discuss the assumptions needed to cash out (i) and (ii), beginning with the former which might appear as rather innocuous. As noted in Section 6.2, the standard assumption in the literature on shifty indexicals is that, at LF, these expressions are represented as individual-denoting constants (see fn. 9) as in (42).

$$(42) \quad \llbracket I \rrbracket^{c,i,g} = \text{AUTH}(c)$$

Why should *shifty* indexicals have the representation in (42), instead of the more familiar representation in (43c)?

$$(43) \quad \begin{array}{l} \text{a. } \llbracket 1st \rrbracket^{c,i,g} = \lambda x : \text{AUTH}(c) = x. x \\ \text{b. } \llbracket pro_x \rrbracket^{c,i,g} = g(x) \\ \text{c. } \llbracket 1st pro_x \rrbracket^{c,i,g} = \# \text{ unless } g(x) = \text{AUTH}(c); \\ \quad \text{if defined then } \llbracket 1st pro_x \rrbracket^{c,i,g} = g(x). \end{array}$$

Notice the fact that indexicals can have bound readings (Heim 1991c), as shown in (44), is strong evidence that the representation in (43c) is at least available. The claim that *shifty* indexicals should be represented as in (42), then, amounts to the claim that some choice is involved in the LF representation of indexicals.

$$(44) \quad \begin{array}{l} \text{Only I did my homework.} \\ \rightsquigarrow \text{Nobody else did his or her homework} \end{array}$$

The question is rarely discussed explicitly in the literature, with the notable exception of Schlenker (2003, 2004) (who attributes the observation to Arnim von Stechow (p.c.)). The problem in a nutshell is that the assignment-dependent representation in (43c) incorrectly predicts presupposition failure in certain cases when the indexical is embedded in a context-shifted environment. To see why, consider the schematic example in (45) in which I pretend that the predicate *believe* allows indexical shifting.¹⁹ In what follows, the reader is well-advised to keep in mind that we are working in a framework in which an individual may exist in several possible worlds ('trans-world individuals').²⁰

$$(45) \quad \text{Leila believes } [\phi \text{ OP } [1st pro_x] \text{ am hungry}]$$

Suppose we want to interpret (45) relative to context c , index i and assignment function g . It is a general fact that assignment functions (once made salient by the contexts) map variables to individuals *independently from the context parameter* (and independently from the index parameter, for that matter). This means that quantification over the context parameter does not influence the denotation of x in (45), because even if the context parameter is quantified over, the assignment function remains constant. Now, in order to compute the denotation of (45) we need to evaluate the embedded clause ϕ with respect to indices that are compatible with what Leila believes in i , i.e., with respect to every i' in $\text{DOX}_{\text{Leila}}^i$. Let i' and i'' be two members of $\text{DOX}_{\text{Leila}}^i$. Crucially, there is no guarantee that $\text{AUTH}(i')$ and $\text{AUTH}(i'')$ are the same individual: just as Leila may be uncertain regarding the world in which she lives, she may be uncertain regarding who she is.²¹ If so, then $\text{AUTH}(i')$ and

¹⁹The choice to use *believe* instead of *tell* is made only for expository purposes as it allows side-stepping certain complications that arise with the latter. Nothing hinges on this and the point made in the text applies to *tell* just as it does to *believe*.

²⁰One might suspect that if we drop this assumption, i.e., if we adopt a framework of 'world-bound' individuals this problem disappears. This, unfortunately, is not the case; once relevant assumptions are made which determine which world-bound individuals are to be considered 'identical', the same problem re-surfaces albeit in a more complicated fashion.

²¹Examples of this in the philosophical literature include Lewis's (1979a) two gods who are omniscient as far as propositional knowledge is concerned (i.e., can locate themselves in the logical space exactly) but suffer from the unfortunate condition of not knowing who they are, and Perry's (1979)

AUTH(i'') could be different individuals (just as uncertainty about the world would mean that WORLD(i') and WORLD(i'') could be different worlds). If this happens to be the case, then the assignment function g fails to assign a value to $[1st\ pro_x]$. Note that due to the context-shifting operator, the context parameter is over-written with the index parameter: *believe's* quantification over the index parameter is turned into quantification over both the index *and the context parameter*. The interpretation of (45), then, involves, in part, computing the denotation of $[1st\ pro_x]$ with respect to i' and i'' . If the author coordinates of i' and i'' end up being distinct, then there is no individual that g can assign x to *while satisfying the presupposition triggered by the first person feature* according to which $g(x)$ must be identical to both AUTH(i') and AUTH(i''). The inevitable result in that case would be presupposition failure due to transitivity of identity and (45) is predicted to be undefined.

The LF in (45), then, triggers a presupposition failure as soon as Leila has some uncertainty regarding who she is. This, however, is incorrect: we want this LF to denote a proposition that comes out as true as soon as Leila has a first-personal thought of the form, "I am hungry", and this could happen even if Leila harbors some uncertainty regarding her identity. One might react to this problem by dismissing it; perhaps uncertainty about who one is is rare enough that it can be safely ignored. This may be true regarding uncertainty about who one is, but it certainly is not true regarding uncertainty about who one is talking to. Consider the context in (46), nothing about which is particularly fantastic.

- (46) *Context: Doctor Rezai calls his patient's house with some bad news. The person who picks up the phone is a young woman. The doctor knows that his patient has two daughters (call them Zahra and Zohreh), so he infers that the person on the other side is one of the daughters; however, he does not know which one. He says, "your father needs antibiotics which are hard to find".*

If I know that, as a matter of fact, the doctor was talking to Zahra (because Zohreh is out of town), I can report the doctor's utterance as follows.

- (47) Doktor be Zahra goft babat be ye dāruye nāyāb ehtiyāj dāre.
 Doctor to Zahra told father-2SG to one medicine rare need has-3SG
 'The doctor told Zahra that her father needs a rare medicine.'

Suppose we represent the shifted second person pronoun in (47) as a restricted variable, as in (48). What individual should the variable x refer to? Note in particular that x cannot be mapped to Zahra; as far as the doctor is concerned his addressee may well have been Zohreh, the other sister. Thus, while every index / context i' compatible with the doctor's utterance is such that ADDR(i') is either Zahra or Zohreh, as the doctor is uncertain about who his addressee is, there are both indices compatible with his utterance in which ADDR(i') is Zohreh and indices compatible with his utterance in which ADDR(i') is Zahra. As Zohreh and Zahra are different individuals, the assignment-function cannot assign a value to x without incurring presupposition failure.

- (48) The doctor told Zahra [OP [2ND pro_x]'s father needs a rare medicine]

The lesson that Schlenker (2003, 2004) draws from this problem reflects what I

amnesiac in the Stanford library.

take to be the common wisdom, namely that the LF representation of indexical pronouns (perhaps pronouns in general) is systematically ambiguous between a structure that contains a *free* individual-denoting variable (paralleling (43c)) and one that does not (paralleling (42)), and, importantly for us, that shifty indexicals are typically (perhaps necessarily) represented in the latter form. Here I will adopt a variant of Schlenker’s (2004) particular implementation which is easier to embed in the general framework assumed here. In this variant, every pronoun has the abstract form of a definite description with the descriptive content provided by ϕ -features and an optional variable which is type-shifted with the identity operator, as in (49).²² (Material in angle brackets are optional.)

- (49) [THE [[ϕ -features] $\langle = x \rangle$]]
 a. $\llbracket \text{THE} \rrbracket^{c,i,g} = \lambda P_{et} : \exists ! x P(x). \iota x P(x)$
 b. $\llbracket x \rrbracket^{c,i,g} = g(x)$
 c. $\llbracket = \rrbracket^{c,i,g} = \lambda x_e. \lambda y_e. x = y$

Crucially, (49) also necessitates systematic re-analysis of the meaning of ϕ -features; these are now assumed to be of type *et*. The relevant entry for first and second person are given below.²³

- (50) a. $\llbracket \text{1ST} \rrbracket^{c,i,g} = \lambda x_e. x = \text{AUTH}(c)$
 b. $\llbracket \text{2ND} \rrbracket^{c,i,g} = \lambda x_e. x = \text{ADDR}(c)$

The representation for the second person pronoun, for example, would be (51a) with the denotation given in (51b).

- (51) a. $\text{you}_{\langle x \rangle} \longrightarrow [\text{THE} [[\dots \text{2ND}] \langle = x \rangle]]$
 b. $\llbracket \text{THE} [[\text{2ND}] [= x]] \rrbracket^{c,i,g} = \#$ unless $g(x) = \text{ADDR}(c)$; if defined, $= g(x)$.
 c. $\llbracket \text{THE} [[\text{2ND}]] \rrbracket^{c,i,g} = \text{ADDR}(c)$

Note that the variable-full representation in (51a) is synonymous with (43c) and the variable-free representation in (51c) is synonymous with (42). The advantage of the former representation is that by deleting the ϕ -features, a pronoun can now be reduced to an unrestricted variable.

- (52) $\llbracket \text{THE} [[\text{2ND}] [= x]] \rrbracket^{c,i,g} = \llbracket x \rrbracket^{c,i,g} = g(i)$

²²An alternative route would be to analyze shifty indexicals as *de re* pronouns with a *de se* layer ‘on top’ contributed by the person feature. Adopting Percus & Sauerland’s (2003a) account of *de re* ascription for simplicity, the idea is that the LF of (45) looks something like (i) where the ‘concept-generator’ variable G intervenes between the variable *x* and the first person feature. The reader can verify that the assumptions proposed in Percus & Sauerland (2003a) allow the assignment function to map *x* to Leila if it maps the variable G to Leila’s ‘self’-concept (approximately, the function that maps every index compatible with Leila’s beliefs to the author coordinate of that index, i.e., the individual in that world with whom Leila self-identifies, and, say, is undefined for other indices except possibly for the index of evaluation in which the function simply returns Leila).

(i) Leila believes λG [OP [1st [G pro_x]] am hungry]

²³Appropriate entries for number and gender can be similarly constructed. For familiar reasons having to do with the interpretation of *plural* indexical pronouns, the entries in (50) must be re-stated in terms of *overlap* as opposed to *identity*. For example, $\llbracket \text{2ND} \rrbracket^{c,i,g} = \lambda x_e. x \circ \text{ADDR}(c)$ (where ‘ \circ ’ is the overlap operator: $x \circ y \Leftrightarrow \exists z(z \leq x \wedge z \leq y)$).

This brings us to the second set of assumptions needed to make this analysis work. At this point, the following LF can be generated for our running example (38). What we need in order to generate the unshifted reading is to remove the second person feature in the (underlined) lower copy (as the indexical feature on the lower pronoun will be removed, the problem raised above does not arise; uncertainty regarding the addressee will be possible once the feature on the lower copy is removed and the one on the higher copy is interpreted).

- (53) [which ... [THE [[2ND] [= x]]] ...]
 L told M
 [OP I [which ... [THE [[2ND] [= x]]] ...] love]

There are more than one ways to design a system in which the feature on the lower copy of the pronoun is not semantically interpreted. As far as I can see all of these collapse for the particular case of examples discussed here.²⁴ The most salient way to do this, I believe, is to assume that the grammar has access to a deletion operation at the LF interface. This is not enough however, we also need to assume that deletion can target expressions of any size *down to individual features*. (This is broadly in line with Merchant's (2000) 'partial deletion' operation, augmented with the claim that individual features also qualify as candidates for 'partial deletion'.) If so, then in principle deletion can apply to the person feature of the lower copy of the indexical in (53) generating (54) as the result, as desired.

- (54) [which ... [2nd pro_x] ...]
 L told M
 [OP [1sg pro_y] [which ... [~~2nd~~ pro_x] ...] love]

If deletion is an operation that the grammar has access to then there *must* be constraints on its application. To begin with, a freely available deletion operation threatens to make copy theory completely vacuous as lower copies can always be freely turned into impoverished structures. Furthermore, free deletion makes a host of undesirable empirical predictions. For example, note that if deletion of features was a freely available option then a sentence like *everybody did my homework* would allow for the reading, everybody *x* is such that *x* did *x*'s homework (cf. *everybody did his homework*). This reading can be easily generated if *my* is bound by *everybody* and its features are removed by deletion. The most salient solution, building on some remarks in Fox 1999, is to posit that deletion is subject to an economy constraint that severely limits its application. The intuition is that deletion can only apply if the resulting LF has an interpretation that is 'different' from the original, pre-deletion LF. Furthermore, assuming that the 'cost' of deletion is positively correlated with the size of the deleted constituent, we also posit that deletion can only target the smallest constituent possible to yield an LF with a particular interpretation that is different

²⁴One possibility, with interesting empirical consequences is this. One could postulate that ϕ -features on lower copies in chains are not interpreted because ϕ -features on unpronounced pronouns are in general not interpreted. For example, the first person feature on the bound pronoun in *only I did my homework* and *I did my homework, and you did too* must remain uninterpreted to derive the target readings. One could imagine that this is because in both cases covert instances of the bound pronoun are the culprit (the occurrence of the bound pronoun in the ellipsis site and in the computation of focus-alternatives). An approach along these lines unifies all these cases under one rubric. (Many thanks to Philippe Schlenker for pointing this out.)

from the original LF's. These two requirements are codified below.

- (55) Deletion can target the constituent α of ϕ only if (i) $\phi[\alpha]$, i.e., the result of deleting α in ϕ , is not strawson equivalent with ϕ and (ii) there is no α' such that α' is a proper part of α and $\phi[\alpha']$ is strawson equivalent with $\phi[\alpha]$.

To cash out what it means for two LFs to have different interpretations within a trivalent framework (in which a third truth-value is introduced to capture presupposition failure), (55) relies on the notion of strawson equivalence instead of the more straightforward notion of strict equivalence. According to the latter, two propositions are distinct if and only if it is possible for them to have distinct truth-values. According to the former, weaker notion, two propositions are distinct if and only if it is possible for one of them to be true and the other one be false.

- (56) a. ϕ and ψ are strictly equivalent relative to assignment function g iff for any context c and indices i and i' , $\llbracket \phi \rrbracket^{c,i,g} = \llbracket \psi \rrbracket^{c,i',g}$.
b. ϕ and ψ are strawson equivalent relative to assignment function g iff for any context c and indices i and i' , if $\llbracket \phi \rrbracket^{c,i,g} \neq \llbracket \psi \rrbracket^{c,i',g}$ then either $\llbracket \phi \rrbracket^{c,i,g} = \#$ or $\llbracket \psi \rrbracket^{c,i',g} = \#$.

To see the justification for using strawson equivalence consider again the LF $[_\phi \textit{ everybody } \lambda x [x \textit{ did } [1\textit{st pro}_x]\textit{'s homework}]]$. We want to prevent deletion from targeting the first person feature of the bound pronoun, thereby generating $[_\phi \textit{ everybody } \lambda x [x \textit{ did } [\cancel{1\textit{st}} \textit{ pro}_x]\textit{'s homework}]]$. Importantly, ϕ and ψ are *not* strictly equivalent. On the standard assumption that presuppositions triggered in the scope of the universal quantifier project universally to root, as soon as the domain of quantification contains an individual distinct from the speaker, ϕ is undefined (has the third truth-value) while ψ is either true or false depending on the facts; therefore it is indeed possible for the two to have distinct truth-values. What is impossible, however, is to find a world in which *both LFs are defined* but have distinct truth-values. If ϕ is defined then the domain of quantification contains only the speaker; if so, then either both sentences are true or both are false depending on whether the speaker has done her homework or not. If we want to prevent deletion from generating ψ from ϕ , then, we need to assume that what counts as 'having different meanings' is 'not being strawson equivalent'. Let us note that the LFs in (53) and (54) are indeed non-strawson equivalent and, therefore, deletion can be legitimately applied to the former to derive the latter. The reason in a nutshell is that whether or not the indexical feature in the lower copy of (53) is interpreted has consequences for the *denotation* of the pronoun in the lower copy: if the indexical feature is present the pronoun refers to the attitude-holder (simplifying somewhat) and if it is not present the pronoun refers to the actual speaker. The choice between the two has truth-conditional repercussions beyond the presuppositions making the two sentences non-strawson equivalent.

Finally, to see the justification for the second clause of (55) (beyond *a priori* considerations pertaining to economy) note that we need to make sure that deletion is not allowed to completely remove the material in the lower copy (thereby generating a structure which is very similar to the one generated by trace theory). Doing that would indeed generate a new meaning, much as before (because it involve deleting the indexical feature *inter alia*), but it would also mean that any r-expression in the restrictor of *which* gets removed with the consequence that movement is now predicted to bleed Condition C; i.e., the second problem mentioned in the previous

section would remain unsolved. The current formulation of (55), however, makes sure that only the *smallest* possible constituent can get removed, i.e., the indexical feature, leaving the rest of the material in the lower copy (including potential r-expressions) intact.

Let us recap. The goal of this section was to design a method whereby the two problems mentioned in the previous section can be addressed without discarding either copy theory of movement or the operator-based account of indexical shifting. We noted that there is a way of doing this which is entirely stipulative, i.e., the brute force approach. To achieve a more principled solution we introduced certain assumptions regarding the LF representation of shifty indexicals and deletion at the LF interface. As explained above, the economy condition on deletion makes sure that this approach solves one of our problems: unshifting is not predicted to bleed Condition C. Unfortunately, however, this approach does not solve the other problem: although we have a way to derive the desired LF in (54) which yields the target truth conditions, we have no way of preventing the double-access LF derived by using the non-assignment-dependent entry of the shifty indexical (which, recall, we need in order to solve the problem of shifty indexicals under attitudes that involve uncertainty regarding the identity of one of the coordinates). Consequently in this framework we predict that the double-access LF is indeed generated by the grammar. As briefly pointed out in Section 6.3, however, this problem might not be so bad after all. Perhaps there is an independent reason to block double-access LFs. The nature of such a constraint, then, remains to be explicated in future work.

Furthermore, as I made explicit above, the approach outlined in this section relies crucially on a deletion operation at the LF interface. As far as I know, there is no independent evidence for such an operation.²⁵ As we will see in the section, if one is willing to adopt the variable-based account of indexical shifting, a single modification of the SHIFT TOGETHER! constraint (which itself is needed to avoid over-generation in that system) solves both of the problems, including the problem of double-access LFs, without relying on an operation of deletion at all.

6.5 Copy theory and the variable-based account of indexical shifting

The core claim of the operator-based approach to indexical shifting, as sketched in Section 6.1, is that the shifted readings of indexicals in indirect discourse result from the presence of a covert context-shifting operator which occupies the left periphery of the embedded clause and which is licensed by only certain attitude predicates as a matter of, *inter alia*, their lexically determined selectional requirements. In contrast, the core claim of the variable-based analysis, which is the focus of this subsection, is that the shifted reading of an indexical results from it being bound by a local context abstractor.

Various implementations of the variable-based account are conceivable with subtle

²⁵There are other questions that one might ask about the proposal in this section. For example, one might worry about an over-generation problem caused by deletion. What happens if, in (53), deletion targets the indexical feature of the *higher* copy and not the lower one? What if it targets both? These questions are potentially problematic for this approach. I leave a proper discussion of these questions to future work.

empirical and conceptual differences (Schlenker 1999, 2003; see also von Stechow 2004 for an attempt in the same spirit). For our purposes here the following “toy” implementation will do just fine: (i) LFs allow for explicit representation of context variables and abstractors; in particular, every LF must be closed off by a context abstractor λc_0 (the ‘matrix’ abstractor, which as a matter of principle is always fed the actual context of utterance), (ii) (at least some) attitudinal predicates quantify over, not possible worlds (or world-time pairs), but contexts and therefore may introduce context abstractors in their scopes, and (iii) indexical expressions involve context variables lexically equipped with features that determine their binding possibilities: in languages that do not allow indexical shifting, every indexical is lexically restricted to be bound by λc_0 while in other languages an indexical may be free to be bound either by λc_0 or a local binder.²⁶ Here, it suffices to view indexical pronouns as context variables with person features which determine which coordinate of the context they refer to, as in (57).²⁷

- (57) a. $\llbracket \text{1st-}c_i \rrbracket^g = \text{AUTH}(g(i))$, $\llbracket \text{2nd-}c_i \rrbracket^g = \text{ADDR}(g(i))$ ²⁸
 b. $\lambda c_0 \dots A \text{ told } B [\lambda c_2 \dots \text{1st-}c_i \dots] \dots$ ²⁹
 If $i = 0$: not shifted, reference to the actual speaker.
 If $i = 2$: shifted, reference to the reported speaker, A.

It is often claimed (Anand & Nevins 2004; Anand 2006; Deal 2017) that the operator-based approach to indexical shift is superior to the veteran, variable-based approach because the former yields a principled explanation of SHIFT TOGETHER! (and, more generally, *No Intervening Binder*) while the latter needs to stipulate this. In other words, in general nothing prevents the variable-based system to generate the LFs in (58) which violate SHIFT TOGETHER!.³⁰

²⁶Note that something also needs to be said, in the variable-based framework, to explain the fact that even in languages that do allow indexical shifting, only certain predicates license indexical shifting. In Farsi, for example, speech-report predicates like *goftan* allow indexical shifting while predicates of belief do not. One way or the other, then, assumption (ii) in the text needs to be modified. One possibility is that only predicates that allow indexical shifting introduce context abstractors in their scope. Another possibility is that every predicate (that quantifies over contexts) may in principle introduce context abstractors but the semantics of only some predicates is such that binding to the introduced context abstractor amounts to indexical shifting. Other possibilities are conceivable, none particularly elegant. The issue is orthogonal to our purses here and is therefore ignored in the body of the text.

²⁷In other words, we need not be explicit about those features that determine whether the indexical must be matrix-bound or not, we assume that indexicals under discussion here, i.e., the first and second person pronouns in Farsi, do not involve any restriction of this sort (because they may or may not shift). We furthermore abstract away from gender (which is not encoded in Farsi to being with) and number features assuming singular marking throughout.

²⁸To avoid clutter I am suppressing type / sortal restrictions, specifically the requirement that the index i , in both entires, needs to be mapped by g to a context.

²⁹Schematic LFs of this kind are adopted to avoid getting into the question of whether, in a variable-based system, indices (traditionally, world / situation variables) need to be represented at LF on top of contexts or not.

³⁰There are variants of the variable-based system which do derive SHIFT TOGETHER!. For example, if one assumes that there is only one context variable, c_0 and that certain predicates can optionally introduce context-abtractors, then, with some technical details straighten out, one arrives at a system which is essentially notational variant of the simplest operator-based system (in which there is only one operator that shifts all coordinates of the context) and, from the present perspective, suffers from the same short comings (to my knowledge, this possibility was suggested to P. Schlenker by E. Zimmermann (p.c.), see also fn. 24 of Deal 2017).

- (58) a. $\lambda c_0 \dots$ A told B [$\lambda c_2 \dots$ 1st- $c_2 \dots$ 2nd- $c_0 \dots$] \dots
b. $\lambda c_0 \dots$ A told B [$\lambda c_2 \dots$ 1st- $c_0 \dots$ 2nd- $c_2 \dots$] \dots

Whether **SHIFT TOGETHER!** can be used as an argument in favor of the operator-based approach depends ultimately on whether the formulation of **SHIFT TOGETHER!** which the operator-based approach derives as a descriptive generalization is in fact empirically adequate. If the analysis provided in this section is correct then **SHIFT TOGETHER!** is more flexible than the operator-based account allows, and this particular argument in favor of the operator-based account is in effect turned into an argument against it.

In (59) an explicit formulation of **SHIFT TOGETHER!** as a stipulation in the variable-based system is provided. I take (59) to be merely a precisification of **SHIFT TOGETHER!** as described in Anand 2006. In other words, (59) is more or less what **SHIFT TOGETHER!** should amount to, according to the operator-based account.³¹ Within the variable-based system, a constraint like (59) can be thought of as an aspect of a binding theory of context variables (cf. Percus 2000).³²

- (59) a. **SHIFT TOGETHER!** Indexical expressions that occur in the same minimal domain must be bound by the same context abstractor.
b. Let α be an occurrence of an indexical in LF ϕ . The minimal domain of α in ϕ , $\text{MD}(\alpha, \phi)$, is the smallest CP / TP of ϕ that dominates α .
c. Two occurrences of indexicals α_1 and α_2 in LF ϕ are in the same minimal domain iff $\text{MD}(\alpha_1) = \text{MD}(\alpha_2)$. (to be revised)

Now, what if an indexical occurs in a chain? In that case, (59) predicts that two different copies of the same indexical may *have to* be bound in some cases by different abstractors. To see this, consider the schematic LF in (60), where the binding possibilities of the two copies of the second person indexical are to be determined.

- (60) [$\alpha \lambda c_1 \dots$ [which \dots 2nd- $c_i \dots$] \dots
A told B
[$\beta \lambda c_2 \dots$ 1st- $c_2 \dots$ [which \dots 2nd- $c_j \dots$] \dots] \dots]

³¹This is not exactly correct. Consider a sentence like *Leila told Mina [ϕ I think [ψ you are smart]]*. Here the MD of *I* is ϕ while the MD of *you* is the smaller ψ . **SHIFT TOGETHER!** as formulated in the text predicts that it should be possible for *you* to shift even if *I* does not, contrary to fact (based on Farsi judgments). Evidently in determining minimal domains, not every CP / TP is taken into account. The (ugly) solution is to introduce the set of predicates that allow indexical shifting, call it *S*, and modify (59b) along these lines: the minimal domain of α in ϕ is the smallest CP / TP of ϕ which (i) dominates α , and (ii) is either the matrix CP / TP or is immediately embedded by a predicate that belongs to *S*. According to this modification the MD of *you* in the above example is ϕ ; ψ is ruled out because it is neither the matrix CP / TP nor immediately embedded by a context-shifting predicate.

³²The more general **No Intervening Binder** can be formulated as follows. The reader can verify that **No Intervening Binder** yields **SHIFT TOGETHER!** as a special case when $\text{MD}(\beta) = \text{MD}(\alpha)$ (i.e., $\text{MD}(\beta) \sqsubseteq \text{MD}(\alpha)$ and $\text{MD}(\alpha) \sqsubseteq \text{MD}(\beta)$).

- (i) a. **No Intervening Binder.** Let α and β be two occurrences of indexicals in LF ϕ . If $\text{MD}(\beta) \sqsubseteq \text{MD}(\alpha)$ then $\text{MD}(c_\beta) \not\sqsubseteq \text{MD}(c_\alpha)$.
b. Let α be a particular occurrence of an indexical pronoun or context abstractor in the LF ϕ . The minimal domain of α , notated $\text{MD}(\alpha)$, is the smallest CP / TP of ϕ that contains α .
c. For any indexical expression α in ϕ , c_α is the context abstractor in ϕ that binds α .
d. Let C_1 and C_2 be two subconstituents of the LF ϕ . Then $C_1 \sqsubset C_2$ iff C_1 is a (proper) subconstituent of C_2 and $C_1 \sqsubseteq C_2$ iff either $C_1 = C_2$ or $C_1 \sqsubset C_2$.

First, we need to determine the MD's of the three occurrences of the indexicals in (60): (a) the smallest CP / TP that contains the first person pronoun is β , (b) the smallest CP / TP that contains the higher copy of the second person pronoun is α , and, finally, (c) the smallest CP / TP that contains the lower copy of the second person pronoun is β . From (a) and (c) it follows by (59) that the lower copy of the second person indexical must be bound by the same abstractor that binds the first person pronoun, i.e., $j = 2$. The MD of the higher copy of the second person pronoun, however, is different (that is, it is larger) therefore there is no need for the higher copy to be co-indexed with λc_2 . It must be, then, that the higher copy is bound by the matrix abstractor, i.e., $i = 1$. The only licit resolution of i and j , then, leads to (61) which is precisely the double-access LF that we need to block.

- (61) $[\alpha \lambda c_1 \dots [\text{which} \dots \text{2nd-}c_1 \dots] \dots]$
 A told B
 $[\beta \lambda c_2 \dots \text{1st-}c_2 \dots [\text{which} \dots \text{2nd-}c_2 \dots] \dots] \dots]$

The core problem here is that the MD of the lower copy is determined to be the embedded clause by definition (59b).

I would like to suggest an *ad hoc* modification of (59) that changes this. Specifically, I'd like to propose that the *head* of the chain dictates the minimal domain of the lower copies: if an indexical occurs in a chain then its MD is the same as that of the highest copy. This modification can be captured by a slightly different formulation of definition of minimal domains given in (59b).

- (62) Let α be a particular occurrence of an indexical in LF ϕ . The minimal domain of α in ϕ , $\text{MD}(\alpha, \phi)$, is the smallest CP / TP of ϕ that dominates *the highest copy of α in ϕ* .³³

Let us see how the variable-based analysis, supplemented with (59) as modified in (62) solves both of the problems mentioned above when combined with the standard formulation of copy theory. Consider the LFs in (63). These four LFs exhaust the logically possible options *vis-à-vis* the binding of the indexical pronouns. (The embedded first person pronoun is shifted by assumption.) Note that the minimal domains of the indexicals are the same in all four cases, as their structural location is constant. They only differ in binding relations.

- (63) a. $[\alpha \lambda c_1 \dots [\text{which} \dots \text{2nd-}c_1 \dots] \dots]$
 A told B
 $[\beta \lambda c_2 \dots \text{1st-}c_2 \dots [\text{which} \dots \text{2nd-}c_1 \dots] \dots] \dots]$
 b. $[\alpha \lambda c_1 \dots [\text{which} \dots \text{2nd-}c_1 \dots] \dots]$
 A told B
 $[\beta \lambda c_2 \dots \text{1st-}c_2 \dots [\text{which} \dots \text{2nd-}c_2 \dots] \dots] \dots]$
 c. $[\alpha \lambda c_1 \dots [\text{which} \dots \text{2nd-}c_2 \dots] \dots]$
 A told B
 $[\beta \lambda c_2 \dots \text{1st-}c_2 \dots [\text{which} \dots \text{2nd-}c_2 \dots] \dots] \dots]$
 d. $[\alpha \lambda c_1 \dots [\text{which} \dots \text{2nd-}c_2 \dots] \dots]$
 A told B
 $[\beta \lambda c_2 \dots \text{1st-}c_2 \dots [\text{which} \dots \text{2nd-}c_1 \dots] \dots] \dots]$

³³If α does not occur in a chain, then I assume it is its own 'highest copy' in a trivial sense.

LFs (63c) and (63d) are immediately ruled out as the indexical in the higher copy in both cases is co-indexed with the embedded context abstractor which does not c-command them. What the exact reason is to rule out (63c) and (63d) is theory-dependent. Within the toy model that I have sketched, it suffices to assume that context variables must be bound: contexts cannot be referred to via deixis.³⁴

This leaves us with (63a) and (63b). As to the latter, which amounts to the problematic double-access LF, notice that by (62) the minimal domain of the second person pronoun in both copies is effectively the matrix CP. Consequently, (63b) now violates SHIFT TOGETHER!, a prediction which can be thought of “vertical shift together”. The only LF in (63), therefore, which survives is (63a) which represents exactly the reading of (31b) in which the second person pronoun in the restrictor of *kudum* is unshifted. Finally, note that no application of deletion is required to generate the reading denoted by (63a) (assuming that the grammar has access to such an operation to begin with). It follows from economy, then, that deletion is barred from applying here and Condition C is not predicted to be violated as the material in the lower copy, including possible r-expressions, remain intact.

One theory-internal point is worth discussion at this point. Consider again the LF in (63b), repeated below. As pointed out above, (62) makes sure that this LF is ruled out by SHIFT TOGETHER!. But notice that in this LF the two copies of the context variable carry two distinct indices, i.e., are two different variables. But if these two are *copies* then perhaps this should not be allowed on independent grounds. Specifically, one might think that we can drop (62) and simply keep in mind that in copy theory movement is a *copying* operation, and copies are identical by definition.

- (64) $[\alpha \lambda c_1 \dots [\text{which} \dots \text{2nd-}c_1 \dots]] \dots$
 A told B
 $[\beta \lambda c_2 \dots \text{1st-}c_2 \dots [\text{which} \dots \text{2nd-}c_2 \dots]] \dots \dots]$

Two remarks are in order. First, although copies are by definition identical there might be processes that apply *after* copies are generated that manipulate one copy but not the other one, causing divergence. Indeed we have already discussed two such ‘post-syntactic’ processes: trace conversion and deletion. The relevant characterization, then, is this:

- (65) Elements of a chain are identical *modulo* post-syntactic operations such as trace conversion, deletion and so on.

The crucial question, then, is whether indexation of pronouns, including ‘context pronouns’ is a post-syntactic operation or not. If it is not, then indeed (63b) is ruled out by (65). If it is, then whether (63b) is ruled out or not depends on constraints on the indexation process.

Second, and more to the point, even if indexation is not a post-syntactic process (62) is still necessary. To see this, suppose indexation is not a post-syntactic process. On this assumption, we can only generate two LFs for our running example one of which is (63a), repeated below (the other LF, which I put aside, is one in which no indexical shifts). The problem is that if we adopt the standard version of SHIFT TOGETHER! given in (59) this LF is ruled out. Note that the MD of the lower copy of second person pronoun in this LF, according to the definition in (59b), is *the embedded*

³⁴As P. Schlenker points out (p.c.), this constraint can only be stated as a *global* condition.

clause. As a result SHIFT TOGETHER! blocks this LF and we are left with only one possible LF, one in which no indexical shifts.

- (66) $[\alpha \lambda c_1 \dots [\text{which} \dots \text{2nd-}c_1 \dots] \dots]$
 A told B
 $[\beta \lambda c_2 \dots \text{1st-}c_2 \dots [\text{which} \dots \text{2nd-}c_1 \dots] \dots] \dots]$

To summarize, then, (62) is necessary independently from whether indexation of pronouns is a post-syntactic process or not; i.e., independently from whether different copies of a pronoun are allowed carry difference indices.

What we have seen in this subsection is that if we are willing to adopt the variable-based account of indexical shifting, then the problems described in the previous section receive a uniform solution by a simple modification of the SHIFT TOGETHER! constraint which needs to be stipulated in this framework on independent grounds.

Before closing this section, I would like to point out that there might be independent motivation for the way in which SHIFT TOGETHER! was modified on this analysis. Specifically, we think of SHIFT TOGETHER! as a principle of binding theory for context variables. From this perspective, SHIFT TOGETHER! is rather similar to Conditions A and B of (standard) binding theory, as all three conditions are domain-dependent. I would now like to sketch an argument to the effect that the same modification that was made to SHIFT TOGETHER! (i.e., that the minimal domain of indexicals that occur in chains is determined by the head of the chain) might be needed for Condition A as well.

Consider the (English) example in (67). This sentence is acceptable, and in particular, it does not violate Condition A.³⁵

- (67) John asked [which picture of himself] Mary saw.

Now, the LF of this sentence under copy theory would be (68).

- (68) John asked [which picture of himself] Mary saw [which picture of himself].

On a standard formulation of Condition A, given in (69), the lower copy of the anaphor violates Condition A, generating the prediction that (67) should be unacceptable, contrary to fact.

- (69) **Condition A.** Anaphors must be bound by (or be co-referential with) some element in their local domain, where the local domain of an anaphor is the smallest XP that has a subject and contains it.

One might attempt to solve this problem by deleting the lower copy, as in (70).

- (70) John asked [which picture of himself] Mary saw [which ~~picture of himself~~].

Here, too, Condition C can be used to test whether the material in the lower copy is deleted or not. Consider the paradigm in (71).

- (71) a. *John asked [which picture of himself and Mary_i] she_i saw.
 b. John asked [which picture of himself and Mary] Bill saw.
 c. John asked [which picture of himself and Mary_i] her_i father saw.

³⁵A standard formulation of Condition A suffices here: This would be the embedded TP in (67).

The judgments are subtle, but if the contrast between (71a), on the one hand, and (71b) and (71c), on the other, proves robust, then the paradigm in (71) furnishes a strong piece of evidence in against of the analysis in (70). But notice that if we perform the same move for Condition A that was proposed above for SHIFT TOGETHER!, cf. (72), this problem disappears. On the formulation of Condition A in (72), the lower copy of the anaphor in (68) does not violate Condition A because its local domain is extended by the movement, and is the same as the local domain of the higher copy (i.e., the local domain in both cases is the matrix TP).

- (72) **Condition A. (revised)** Anaphors must be bound by (or be co-referential with) some element in their local domain, where the local domain of an anaphor is the smallest XP that has a subject and contains the occurrence of the anaphor at the head of the chain.

6.6 Conclusion

I pointed out two problems that arise when copy theory of movement is naively combined with the operator-based account of indexical shift. I discussed two possible ways to tackle these problems. In the first approach auxiliary assumptions are introduced to fix these problems keeping copy theory and the operator-based account constant. As pointed out, this approach faces a number of challenges. In the second approach the operator-based account is dismissed in favor for of the variable-based analysis of indexical shift. I argued that once this move is made, a fairly natural modification of an independently needed stipulation, i.e., SHIFT TOGETHER!, provides a unified solution to both problems.

Part III

Part Three: Context *qua* Paralinguistic Signals

Chapter 7

On Co-nominal Pointing

7.1 Introduction

On an abstract level of analysis, the two sentences in (1) below both tend to convey the information, namely, that some but not all students failed.

- (1) The exam was very difficult . . .
- a. . . and some students failed.
 - b. . . yet, not all students failed.

It is, therefore, surprising that if a pointing gesture toward an individual, say, John, accompanies the quantificational noun phrase of (1a), as in (2a) below, the result is an inference that is partially different from the inference generated by having the same gesture accompany the quantificational noun phrase of (1b), as in (2b) below.¹ On the one hand, both utterances in (2) trigger the inference that the ‘pointing target’, i.e., John, is one of the students. On the other hand, (1a) triggers the inference that John failed the exam while (2b) triggers the opposite inference.

- (2) The exam was very difficult . . .
- a. . . and $\overline{\text{IX}}$ some students failed. (IX to John)
 \rightsquigarrow *John is one of the students and he failed the exam*
 - b. . . yet, $\overline{\text{IX}}$ not all students failed. (IX to John)
 \rightsquigarrow *John is one of the students and he did not fail the exam*

This paper is about the interpretive process that underlies such cases of ‘co-nominal pointing’. Specifically, I will focus on the interpretation of those pointing gestures that are temporally aligned with (i) *quantificational* (ii) *noun phrases*, as opposed to referential noun phrases or predicates. In such configurations, pointing generates a

¹Throughout this paper, the symbol ‘IX’ (standing for ‘index’) is used to represent an occurrence of a pointing gesture. Overlining is used to give an upper bound on the time slot in which pointing can felicitously occur with the intended interpretation. (Thanks to Jeremy Kuhn for suggesting this notation.) These notational devices are both adopted from sign language research, yet while my use of the former follows convention, the latter might be slightly misleading since in sign language literature overlines are typically associated with non-manual signs, such as eye-gaze. Regardless of any potential similarity between eye-gaze in sign language and pointing in speech, no substantial *assumption* is intended by this notation here (although the hint for possible future research very much is). In particular, I will consistently ignore the potential role of eye-gaze in this study.

rich inferential pattern that depends, crucially, on the determiner, as seen already in the contrast between (2a) and (2b), and also the logical structure of environment that surrounds the gesturally modified noun phrase, as we will see later on. Examining this pattern forms the empirical side of this paper.²

On the theoretical side, the question pertaining to the interpretation of co-nominal pointing can be split into two parts. The first task is to analyze the *content* of the co-nominal pointing inference in abstraction from how this inference projects, i.e., interacts with logical operators. Second, the projection facts need to be brought into the picture and accounted for. As to the former, I will propose that the function of co-nominal pointing is to ‘exemplify’ the set of individuals that is made salient by the quantifier for subsequent anaphoric pick-up (‘plural anaphora’, Nouwen 2003 a.o.). More specifically,

- (3) **The exemplification hypothesis. (to be slightly revised below)** If a pointing gesture IX toward individual T_{IX} accompanies a quantificational noun phrase α , it triggers the inference that $T_{IX} \in S_\alpha$, where S_α is a set of individuals (a ‘discourse referent’) that is made salient by α for anaphoric uptake.

The exemplification hypothesis immediately predicts that there should be a parallelism, on a case by case basis, between the co-nominal pointing inferences that are triggered if pointing modifies the quantificational noun phrase α and the anaphoric potential of α in discourse, as exposed by the referential possibilities of pronouns that are anteceded by α . Formulated in strongest possible sense (possibly too strong), the prediction is that,

- (4) If a pointing gesture IX toward individual T_{IX} accompanies a quantificational noun phrase α then the inference is triggered that $T_{IX} \in S$ if and only if the set S can be referred back to by a (plural) pronoun that is referentially dependent on α .

That there is some *prima facie* plausibility to this prediction is established by the following pair of examples.

- (5) a. $\overline{\text{IX}}$
Some students failed the exam.
 $\rightsquigarrow T_{IX}$ is one of **the students who failed the exam**
- b. Someⁱ students failed the exam. They_i are upset.
 \rightsquigarrow **The students who failed the exam are upset**

²One case that, abstractly speaking, falls in the same category but is not discussed in this paper is the effect of having pointing gestures with *wh*-words. In particular, Alexandre Cremers (p.c.) points out the following data point (the inference, particularly as it pertains to John’s beliefs *vis-à-vis* whether x is a student or not, is not entirely clear to me; more work is required to determine the content of this inference adequately).

- (i) John knows $\overline{\text{IX}}$ which students failed the exam.
 \rightsquigarrow *The person pointed to x is one of the students and John correctly believes that x failed the exam*

Notice that, as is well known, *wh*-phrases license anaphoric dependencies much like run-of-mill quantificational indefinites (Haida 2007, a.o.). As the exemplification hypothesis, introduced just below, accounts for the interpretation of co-nominal pointing on the basis of their anaphoric potentials, it seems reasonable to assume that it applies to these cases as well. I have to leave a detailed discussion to future work.

Indeed, the exemplification analysis builds on the intuition that the interpretive effect of pointing in, e.g., (2a) can be quite transparently paraphrased as in (6).

(6) Someⁱ students failed, and John is one of them_i.

In Section 7.2, the exemplification hypothesis, and the prediction in (4) in particular, are discussed in detail in connection to the anaphoric potential of various determiners. Interestingly, there is a *prima facie* problem for the exemplification hypothesis when it comes to the universal determiners *every* and *no* (and their variants). These determiners seem to trigger inferences of a *scalar* nature (cf. the paraphrase in (7b)) when modified by co-nominal pointing.

- (7) a. $\overline{\text{IX}}$ Every student passed the exam.^{3,4} (IX to John)
 \rightsquigarrow *John is one of the students and he was unlikely to pass the exam*
 b. Every student passed the exam, even John!

In Section 7.3, the relevant facts pertaining to *every* and *no* are briefly summarized for subsequent reference. Putting these problematic cases temporarily aside, in Section 7.4, I turn to the projection problem of co-nominal pointing inferences. I will argue that these inferences are *presuppositional*, in a standard sense, but this claim requires a stipulative modification of the exemplification hypothesis. According to this modification, pointing exemplifies the discourse referent made salient by the determiner *on the condition that this set is not empty*.

- (8) **The exemplification hypothesis. (final version)** If a pointing gesture IX toward individual T_{IX} accompanies a quantificational noun phrase α , it triggers the *presupposition* that $S_\alpha \neq \emptyset \Rightarrow T_{IX} \in S_\alpha$.

In the last analysis, then, co-nominal pointing inferences are conditional presuppositions. This makes the present analysis look very similar to Schlenker’s (2018) ‘cosuppositional’ analysis of *iconic* gestures that modify predicates; according to Schlenker’s analysis, the iconic gesture SLAP in an utterance of *John [punished his son]_{SLAP}* triggers the presupposition that *if John punished his son, he did so by slapping him*. This parallelism is discussed in Section 7.6, where it is also noted that co-nominal pointing inferences share the same behavior as their iconic brethren in non-monotonic environments and environments involving ellipsis and focus-sensitive operators. Before that, however, in Section 7.5, I will have a second look at the problem raised by *every* and *no*. I will argue that once we are equipped with the final version of the exemplification hypothesis in (8), the problem with universal determiners is automatically solved *if* we are open to make certain assumptions regarding the anaphoric potential of these determiners that are somewhat non-orthodox, in the sense that they cannot be directly verified using pronouns; specifically, one assumption would have to be that *no* triggers a discourse referent for the set that consists of the intersection of the set denoted by its restrictor and scope (i.e., the REFERENCE SET, see Section 7.2 for the terminology) and, more surprisingly, *every* triggers a discourse referent for the set that consists of the intersection of the set denoted by its restrictor and the *complement*

³Some consultants report that this utterance improves significantly with a sharp focus on the determiner. It is my impression that this intonation is not essential, however.

⁴Some consultants prefer to paraphrase the inference in less than polite terms, as in *the person pointed at is the most stupid of the relevant students (or is plainly stupid)*.

of its scope (i.e., the infamous COMPLEMENT SET). Some remarks in support of these assumptions are provided in that section as well. Section 7.7 concludes the paper by, *inter alia*, sketching certain facts pertaining to co-nominal *iconic* gestures.

Some preliminary remarks pertaining to the limitations of this study are best stated at the outset. First, the main language studied in this paper is English. Every judgment reported has been checked against the intuitions of at least three native English speakers, including two non-linguists. Whenever comparable constructions existed, the judgments have been replicated in Farsi (based on my own intuitions and one non-linguist consultant) and French (one linguist). As of yet, I have not found any cross-linguistic variation.

Second, the timing of the gesture matters. For my consultants, (9) below was either unacceptable or it triggered the inference that the person pointed to is a (heavy?) smoker. A possible paraphrase would be, *not every student here is a smoker like he is!*, with no indication of whether *he* refers to a student in the mentioned department. This is in clear distinction with the effect of pointing in (2b) above.

- (9) Not every student in this department is a ^{IX}smoker.

I have to leave the question of how *co-nominal* pointing, as in (2a), relates to *co-predicative* pointing, as in (9), to future research. In the domain of co-nominal pointing, I am not aware of any effect depending on whether the gesture co-occurs with the determiner (e.g., *not every* in (2b)) or its argument (e.g., *students* in (2b)), therefore throughout I have over-lined the entire quantificational noun phrase. That said, the following data point suggests that the precise timing of the gesture *vis-à-vis* the internal make-up of the quantificational phrase can be quite important.

- (10) a. ^{IX}Some but not all students smoke.
 \rightsquigarrow *The person pointed to smokes.*
- b. Some but not ^{IX}all students smoke.
 \rightsquigarrow *The person pointed to does not smoke.*
- c. ??Some but not all ^{IX}students smoke.⁵

Third, the inferences studied in this paper are sometimes difficult to detect, let alone articulate or paraphrase. My response to this problem has been to focus primarily on cases where judgments are relatively clear and robust. As it happens, the paradigm in (2) is already rich enough to allow us explore and evaluate several interesting hypotheses. Needless to say, this line of inquiry can benefit greatly from controlled, quantitative studies. My hope is that the hypothesis advanced here and the generalizations upon which it is built will be of some value to prospective experimental works of this kind.

Fourth, there are two major omissions in this paper. One pertains to ‘plural pointing’, i.e., pointing to several individuals, and one pertains to the shape of the pointing gesture. Regarding the former, I can offer no apology. (The reader can consult Anvari 2016 for some very sketchy relevant discussion.) I hope future work

⁵This utterance improves if pointing is directed at multiple individuals, in which case the (possible?) inference is triggered that the plurality pointed to consists of students, with no implications pertaining to their smoking habits.

will address this omission. Regarding the latter, I *can* offer an apology although, perhaps, not a very convincing one. Essentially, the link between the pointing shape and the inferences triggered in the relevant constructions *appears* to be fairly indirect. For example, as far as I can see, the inferences reported in (2) remain stable if instead of pointing with extended index finger another pointing shape is used, such as pointing palm up with all fingers extended, side-ways pointing with the thumb, head-tilt or even a simple tap on the back. Evidently, all that matters is that a certain individual is made salient in an appropriate time-frame. But this issue needs to be investigated more thoroughly. See [Kendon & Versante 2003](#) for some very interesting observations.

Fifth, even a superficial examination of the examples in (2), repeated below, strongly suggests an analysis which slightly differs from the exemplification hypothesis.

- (11) a. ... and $\overline{\text{some students}}^{\text{IX}}$ failed.
 \rightsquigarrow *The person pointed at is one of the students who failed*
- b. ... yet, $\overline{\text{not all students}}^{\text{IX}}$ failed.
 \rightsquigarrow *The person pointed at is one of the students who did not fail*

Specifically, suppose we are equipped with an inferential engine that can ‘catch’ those inferences that are licensed by an utterance that are existential in nature. On the basis of this system, we can then postulate that all pointing does is that it ‘witnesses’ some of these existential statements, where a witness of an existential inference is defined roughly as follows.

- (12) An individual *a* witnesses the existential inference $\exists x\phi(x)$ if and only if $\phi(x)$ is true relative to an assignment that maps *x* to *a*.

For example, consider (11b) (it is obvious how this would apply to (11a)). The asserted sentence (minus the gesture) triggers the existential inference that there is at least one student who did not fail. The idea is that pointing simply provides a witness to this existential statement: ‘I guy I am pointing at’ is a student who did not fail. Although this line of thinking is intuitively very appealing, I think it should be resisted. First of all, note that (as implied at the very outset), both sentences (11a) and (11b) normally trigger two existential inferences: that at least one student fail and that at least one student did not. The former inference is asserted by (11a) but is an (indirect) scalar implicature of (11b). The latter inference is a scalar implicature of (11a) but is asserted by (11b). Now, why is it that pointing witnesses only one of these existential inferences? The obvious answer would seem to be that pointing only witness inferences that are part of the ‘at-issue’ content of accompanying sentences. But, then, consider (13).

- (13) $\overline{\text{Only some students}}^{\text{IX}}$ failed.
 \rightsquigarrow *The person pointed at is one of the students and he failed the exam*

Interestingly *only some* patterns with *some*, rather than *not every*, despite the fact that the relevant existential inference in the case of (13), i.e., that there is at least one student who failed, is in fact presupposed by (13), i.e., is not at-issue. Note, on the other hand, that the exemplification hypothesis makes the correct prediction for (13):

- (14) Only some students failed. . . they are French.
 \rightsquigarrow *The students who failed are French*
 $\not\rightsquigarrow$ *The students who did not fail are French*

The should also note that the exemplification hypothesis, because it relies on plural anaphora, does come very close to the analysis just sketched; it is just different enough from it to make correct predictions in cases like (13).

Sixth, and finally, the reader should note that in advancing the exemplification hypothesis I am not claiming that exemplification is the only function that co-nominal pointing can perform; there may very well be some ambiguity in how co-nominal pointing is interpreted. Indeed, I am not even claiming that exemplification is all there is to the particular interpretation of co-nominal pointing that I will focus on in this paper. For instance, I suspect that the exemplification inference must be strengthened in certain cases so as to imply, not only that the pointing target belongs to the discourse referent that the determiner makes salient, but also that the fact that pointing target belong to this set yields (some form of) support for the assertive content of the sentence. I can only hope that future work will address these issues.

With the preliminary remarks out of the way, we can now move on to the main content.

7.2 The exemplification hypothesis

Consider again the first version of the exemplification hypothesis, i.e., (3), from the previous section, and the prediction it makes *vis-à-vis* anaphora, both repeated below.

- (15) **The exemplification hypothesis. (to be revised)** If a pointing gesture ix toward individual T_{ix} accompanies a quantificational noun phrase α , it triggers the inference that $T_{\text{ix}} \in S_{\alpha}$, where S_{α} is a set of individuals (a ‘discourse referent’) that is made salient by α for anaphoric uptake.
- (16) If a pointing gesture ix toward individual T_{ix} accompanies a quantificational noun phrase α then the inference is triggered that $T_{\text{ix}} \in S$ if and only if the set S can be referred back to by a (plural) pronoun that is referentially dependent on α .

This section is an extended attempt in verifying this prediction. I will go about this task by systematically comparing the anaphoric potential of various determiners with the inference that is generated when these inferences are modified by a pointing gesture to a single individual. Before I do so, it is helpful to lay down some terminology and a descriptive generalization pertaining to plural anaphora.

I will follow the literature in assuming that three sets in particular are the candidates for anaphoric salience and that different determiners differ in which of these sets they make available, both as a matter of their lexically determiner logical meaning and contextual factors. These three sets are defined in (17)^{6,7} and the two crucial generalizations pertaining to their availability is given in (18) (from Nouwen 2003,

⁶I follow Nouwen’s (2003) terminology. The reader should note that different researchers use these terms differently.

⁷In (17) I have ignored the form / denotation distinction for simplicity.

whose discussion I take to be the state of the art as far as the empirical picture is concerned).

- (17) [[DETERMINER N] M]
- | | | |
|----|-----------------------|------------------|
| a. | $N \cap M$ | = REFERENCE SET |
| b. | $N \cap \overline{M}$ | = COMPLEMENT SET |
| c. | N | = MAXIMAL SET |

- (18) Nouwen's (2003) generalizations:
- a. Every determiner makes the REFERENCE SET available.
 - b. (i) The COMPLEMENT SET is only made available if its non-emptiness can be inferred. (ii) Even if the COMPLEMENT SET is made available, it cannot be referred to unless there is no other set available with higher salience which, if chosen as the antecedent, leads to a coherent interpretation.

To begin with, consider the upward-monotonic determiners *some*, (cardinal) *many* and *a few*. As shown in (19), all three determiners license anaphoric reference to the REFERENCE SET. As shown in (20), none of the three determiners license anaphoric reference to the COMPLEMENT SET, even if the property that is predicated of the members of that set requires this resolution at the pain of contextual incoherence. For example, to the extent that (20a) is felicitous at all, it can only mean either that, for some reason or other, the students who failed are happy (REFERENCE SET) or that the students are all happy (MAXIMAL SET), cf. (21).

- (19) a. Someⁱ students failed the exam, they_i are upset.
 \rightsquigarrow *The students who failed are upset*
- b. Manyⁱ students failed the exam, they_i are upset.
 \rightsquigarrow *The students who failed are upset*
- c. A fewⁱ students failed the exam, they_i are upset.
 \rightsquigarrow *The students who failed are upset*
- (20) a. ??Someⁱ students failed the exam, they_i are very happy.
 $\not\rightsquigarrow$ *The students who did not fail the exam are very happy*
- b. ??Manyⁱ students failed the exam, they_i are very happy.
 $\not\rightsquigarrow$ *The students who did not fail the exam are very happy*
- c. ??A fewⁱ students failed the exam, they_i are very happy.
 $\not\rightsquigarrow$ *The students who did not fail the exam are very happy*
- (21) Someⁱ students failed the exam, the rest_i are very happy.
 \rightsquigarrow *The students who did not fail the exam are very happy*

Co-nominal pointing with these determiners generates a reading which essentially boils down to the exemplification of the REFERENCE SET. ('T_{IX}' is used as a shorthand for 'pointing target'.)

- (22) a. $\overline{\text{IX}}$ Some students failed the exam.
 \rightsquigarrow *T_{IX} is one of the students who failed the exam*
 $\not\rightsquigarrow$ *T_{IX} is one of the students who did not failed the exam*
- b. $\overline{\text{IX}}$ Many students failed the exam.
 \rightsquigarrow *T_{IX} is one of the students who failed the exam*

- $\not\rightsquigarrow T_{IX}$ is one of the students who did not failed the exam
 $\frac{IX}{}$
- c. A few students failed the exam.
 $\rightsquigarrow T_{IX}$ is one of the students who failed the exam
 $\not\rightsquigarrow T_{IX}$ is one of the students who did not failed the exam

Next, consider the downward-monotonic determiners *few* and *less than n*. The first thing to note is that judgments in these cases are slightly less robust. Nevertheless, as shown in (23), these determiners allow reference to the REFERENCE SET and, as established by the contrast between (24) and (25), these determiners allow reference to the COMPLEMENT SET as well but, importantly, only if the meaning that results from the REFERENCE SET is contextually incoherent.

- (23) a. Fewⁱ students failed the exam, they_i are upset.
 \rightsquigarrow The students who failed are upset
- b. Less than threeⁱ students failed the exam, they_i are upset.
 \rightsquigarrow The students who failed are upset
- (24) a. Fewⁱ students failed the exam, they_i are French.⁸
 \rightsquigarrow The students who failed are French (possible reading)
 $\overset{?}{\rightsquigarrow}$ The students who did not fail are French
(not easily accessible, if possible at all)
- b. Less than threeⁱ students failed the exam, they_i are French.
 \rightsquigarrow The students who failed are French (possible reading)
 $\overset{?}{\rightsquigarrow}$ The students who did not fail are French
(not easily accessible, if possible at all)
- (25) a. Fewⁱ students failed the exam, they_i are happy.
 \rightsquigarrow The students who did not fail are happy
- b. Less than threeⁱ students failed the exam, they_i are happy.
 \rightsquigarrow The students who did not fail are happy

Co-nominal pointing with *few* and *less than n* can certainly license exemplification of the REFERENCE SET; however, it is not at all clear whether the COMPLEMENT SET can be exemplified as well. Of the five individuals that I have consulted, none was able to access the COMPLEMENT SET reading easily, but two did not rule it out either.

- (26) a. $\frac{IX}{}$ Few students failed the exam.
 $\rightsquigarrow T_{IX}$ is one of the students who failed the exam
 $\overset{?}{\rightsquigarrow} T_{IX}$ is one of the students who did not failed the exam
(probably impossible)
- b. $\frac{IX}{}$ Less than three students failed the exam.
 $\rightsquigarrow T_{IX}$ is one of the students who failed the exam
 $\overset{?}{\rightsquigarrow} T_{IX}$ is one of the students who did not failed the exam
(probably impossible)

Notice that, even with pronominal anaphora, *few* and *less than n* do not allow uncon-

⁸Reference to the MAXIMAL SET is also available for these cases, but I put these aside.

ditional access to the COMPLEMENT SET, as pointed out above: explicit linguistic cues are required to make the REFERENCE SET incoherent, leaving the COMPLEMENT SET as the only possibility.⁹ Thus, it could be argued that since such linguistic cues are not available with pointing, REFERENCE SET can hardly be ruled out and, therefore, blocks COMPLEMENT SET. That said, arguably contextual factors that perform the same function as linguistic cues. At least some of my consultants found the COMPLEMENT SET reading easier to access in the following context; note the contrast between *few* and *some* (which, undeniably, never allows access to COMPLEMENT SET).

- (27) Context: *The conversation is taking place in a math teacher's office. At the moment, the teacher is tutoring several students. These students have struggled with the material during the whole semester. The dean walks in the room . . .*
- a. **The dean:** You have to include some Algebra in the final exam.
 - b. **The teacher [shocked]:** But, then, $\overline{\text{IX}}$ (very) few students will pass the exam!
 $\rightsquigarrow T_{IX}$ is one of the students who will not pass the exam
 - c. **#The teacher [shocked]:** But, then, $\overline{\text{IX}}$ some students will pass the exam!
 $\rightsquigarrow T_{IX}$ is one of the students who will pass the exam
 (the only available reading)

I conclude that not only is the parallelism between anaphora and pointing not threatened by these cases, but that in fact the downward-entailing determiners *few* and *less than n* provide support for it.

Next, consider the determiner *not all* (or *not every*, the distinction will not matter here), which is also downward-entailing but slightly differs from *few* and *less than n* in its anaphoric potential. On the one hand, it appears that anaphoric access to the COMPLEMENT SET is possible with *not all* just as it is with *few*. To see this, compare (28) with (25) above; in both cases, the pronoun can access the COMPLEMENT SET (i.e., the students who did *not* fail) with the help of linguistic cues (i.e., *are happy*). On the other hand, it appears that anaphoric access to the REFERENCE SET of *not all*, while possible, is not quite as salient as anaphoric access to the REFERENCE SET of *few*. To see this, compare (29) with (23). As shown in (29), referring to the REFERENCE SET of *not all* is marked regardless of whether linguistic cues are present to introduce a bias for it (*are upset*) or not (*are French*). Indeed, several of my consultants could not access the REFERENCE SET for the examples in (29) at all. This is contrast with *few* because, as shown in (23), referring to the REFERENCE SET of *few* appears to be the unmarked option.

- (28) a. Not allⁱ students failed the exam, they_i are happy.
 \rightsquigarrow *The students who did not fail the exam are happy*
- b. John didn't fail allⁱ of his students, they_i are happy.

⁹Nouwen suggests that the MAXIMAL SET is made available by any determiner that presupposes the non-emptiness of its domain of quantification (i.e., its restrictor). These are the 'strong' determiners in the sense of Milsark (1974). Importantly, however, as Nouwen points out, specific constructions (such as existential-*there* sentences) aside, many weak determiners also have strong readings which are generally available. The prediction, then, is that reference to the MAXIMAL SET should be generally available unless the determiner is weak and occurs in a construction which does not allow the strong interpretation. I will have a few words to say regarding the MAXIMAL SET in Section 7.7, but I will ignore it until then.

- (29) a. \rightsquigarrow *The students who did not fail the exam are happy*
 Not allⁱ students failed the exam, they_i are upset / French.
 \rightsquigarrow [?] *The students who failed the exam are upset / French*
 b. John didn't fail allⁱ of his students, they_i are upset / French.
 \rightsquigarrow [?] *The students who failed the exam are upset / French*

As expected, the bias that *not all* introduces in favor of the COMPLEMENT SET is mimicked by the co-nominal pointing facts. Indeed, as far as I can tell, co-nominal pointing with *not all* can only exemplify the COMPLEMENT SET.

- (30) a. $\overline{\text{IX}}$
 Not all students failed the exam.
 \rightsquigarrow T_{IX} is one of the students who did not fail the exam
 $\not\rightsquigarrow$ T_{IX} is one of the students who failed the exam
 b. John didn't fail all of his students.
 \rightsquigarrow T_{IX} is one of the students who did not fail the exam
 $\not\rightsquigarrow$ T_{IX} is one of the students who failed the exam
 c. It is not the case that John failed all of his students.
 \rightsquigarrow T_{IX} is one of the students who did not fail the exam
 $\not\rightsquigarrow$ T_{IX} is one of the students who failed the exam

As pointed out at the beginning of this section, the exemplification hypothesis predicts that the interpretation of both co-nominal pointing and plural pronouns with quantificational antecedents rely on the anaphoric potential of the determiner in question. As discussed above, there is a fair amount of evidence in favor of this prediction.

7.3 The problem with *every* and *no*

Universal determiners *every* and *no* generate interesting inference when modified by pointing gestures. As pointed out in Section 7.1, co-nominal pointing with *every* triggers a scalar inference which can be paraphrased with an appositive *even*.

- (31) a. $\overline{\text{IX}}$
 Every student passed the exam.
 \rightsquigarrow T_{IX} is one of the students and he was unlikely to pass the exam
 b. Every student failed the exam, even $\overline{\text{IX}}$ him
 \rightsquigarrow T_{IX} is one of the students and he was unlikely to pass the exam

The case of *no* is entirely symmetric to that of *every*: the inference is again scalar, but it is 'flipped'.

- (32) $\overline{\text{IX}}$
 No student passed the exam.
 \rightsquigarrow T_{IX} is one of the students and he was likely to pass the exam

Even without looking at the details, it is clear that the exemplification hypothesis falls short here as on the basis the exemplification hypothesis, the part of these two inferences pertaining to likelihood is entirely unaccounted for. Indeed, there is

prima facie good reason to think that the likelihood inference might be pragmatic in nature. Consider the case of *every* for example. Evidently, this determiner makes the REFERENCE SET available for anaphoric uptake (I will re-analyze this example in Section 7.5 as involving anaphoric reference to the MAXIMAL SET).

- (33) Every^{*i*} student failed the exam, they_{*i*} are upset.
 ~> *The students who failed the exam are upset*

On this basis, the exemplification hypothesis predicts that if a quantificational noun phrase headed by *every* is modified by pointing, the inference should be triggered that the pointing target belongs to the REFERENCE SET. For example (31a), this amounts to the inference that the pointing target is a student who passed. Importantly, note that the assertive content of the sentence entails that for any *x*, if *x* is a student then *x* passed. Thus, the pointing inference, to the extent that it is informative at all, can only convey the information that the pointing target is one of the students; the rest follows from the meaning of the sentence. One might then suspect that this is the reason why the likelihood inference is generated. Intuitively, the hearer reasons that presumably there is a reason why the speaker took the trouble of singling out that particular individual. The information that that individual is a student is not relevant; rather the question under discussion (presumably) pertains to the proportion of students who passed the exam (e.g., none, some, most, all, etc). A possible conclusion, then, could be that the speaker took the trouble of making that individual salient because he or she finds its remarkable that that particular student also passed the exam. (Analogous reasoning should presumably be applied to (32).)

Some form of pragmatic reasoning based on relevance might indeed account for the scalar inferences in (31a) and (32). However, as I will discuss in Section 7.5, once a certain modification is made to the exemplification hypothesis to make it suitable to be coupled with a presuppositional analysis for projection, the puzzling scalar inferences of *every* and *no* follow immediately, *on certain assumptions about the anaphoric potential of these determiners* which can be to some degree motivated on conceptual and empirical grounds.

7.4 A projection recipe for the exemplification hypothesis

My claim in this section is that, to account for co-nominal pointing inferences, no new mechanism of ‘projection’ is required. More specifically, the claim is that co-nominal pointing inferences interact with logical operators in their environment *in the same way that presuppositions do*. To establish this claim, I will rely on one particularly strong piece of evidence. To setup the background, consider the factive predicate *know* which triggers the presupposition that its complement clause is true.

- (34) John knows that Mary failed the exam.
 ~> *Mary failed the exam*

It is well-known that presuppositions (i) project from polar questions, (35a), and (ii) can give raise to conditional inferences when embedded in the consequent of conditionals, (35b). Indeed, these two tests can be combined. Thus in (35c), the

factive presupposition first projects from the embedded polar question and then is filtered through the antecedent, giving rise to the conditional presupposition at root.

- (35) a. Does John know that Mary failed the exam?
 \rightsquigarrow *Mary failed the exam*
 b. If the exam involved math, then John knows that Mary failed.
 \rightsquigarrow *If the exam involved math, then Mary failed*
 c. If the exam involved math, does John know that Mary failed?
 \rightsquigarrow *If the exam involved math, then Mary failed*

With this in mind, let us consider a case of co-nominal pointing with the determiner *many* as a case study (I believe the same points can be made with other determiners discussed in Section 7.2 *modulo* confounding factors such as the wide-scope reading of indefinites and so on).

- (36) *Context: The person pointed at is John, we know that John is one of the students.*
 $\overline{\text{IX}}$
 Many students failed the exam.
 \rightsquigarrow *John is one of the students who failed the exam*

In (37), (36) is embedded in a polar question in the consequent of a conditional, analogous to (35c).

- (37) If the exam involved math, did $\overline{\text{IX}}$ many students fail the exam?

If we supplement the exemplification hypothesis, as given in Section 7.2, with the assumption that the exemplification inference is presuppositional, then, on analogy with (35c), we expect the inference to be generated at root that if the exam involved math, then John is a student who failed. This prediction is almost correct, intuitively, but it is a bit too strong. Specifically, none of my consultants inferred from (37) that if the exam involves math then at least one student failed; that is, in their judgment the utterance allows for every student passing the exam, even if the exam involved math. But our prediction is that precisely this inference should be triggered: if every world compatible with the exam having involved math is one in which John is a student who failed, then it follows that if the exam involved math then at least one student failed.

My suggestion is to weaken the exemplification hypothesis. Specifically, the idea is that exemplification, as involved in co-nominal pointing, is predicated on the non-emptiness of the set being exemplified: the inference is that *if that set is not empty* then the pointing target belongs to it. This, coupled with the claim that the pointing inference is presuppositional, is made explicit below (repeated from (8) in Section 7.1).

- (38) **The exemplification hypothesis. (final version)** If a pointing gesture IX toward individual T_{IX} accompanies a quantificational noun phrase α , it triggers the *presupposition* that $S_{\alpha} \neq \emptyset \Rightarrow T_{\text{IX}} \in S_{\alpha}$.

The effect of this modification is precisely that it removes the extra piece of information pointed out above. What we now predict for the case of (37) is the inference that if the exam involved math, then either no student failed or John is one of the students

who failed. This inference seems to match the intuitive judgments pertaining to (37). Indeed, with this in place, we can look at the simpler cases below.

- (39) a. Did $\overline{\text{many students}}$ ^{IX} fail the exam?
 b. If the exam involved math, then $\overline{\text{many students}}$ ^{IX} failed the exam.

The inference predicted for (39a), on the basis of (38), is that either no student failed the exam, or John is a student who failed, and the inference predicted for (39b) is the same as that of (37) above. Both of these predictions seem to be on the right track, according to the judgments that I have been able to obtain, with one proviso. Intuitively, the inference that John is a student seem to project to root without getting ‘conditionalized’ by the antecedent of the conditions in (37) and (39b). I will return to this point in Section 7.7.

7.5 Another look at *every* and *no*

If the discussion in the previous section is on the right track, then the predicted inference for (40a) is (40b). That this predicted inference should surface as the attested inference in (40c) is not particularly mysterious, as the assertive content of the sentence entails that at least one student failed.

- (40) a. $\overline{\text{Many students}}$ ^{IX} failed the exam.
 b. *Either no student failed or John is a student who failed*
 c. *John is a student who failed*

But notice what the predicted inference (40b) amounts to in isolation. One way to read this inference in common sensical terms, on the assumption that John is one of the students, is that John is very likely to fail. This is because this inference rules out the possibility of there being some students who failed but John having passed: if there is student who failed, it is John. Indeed, some of my consultants’ reaction to one of the crucial examples of the previous section, repeated below, was precisely to report the inference that John is likely to have failed the exam, and that is why he is singled out by the speaker via a pointing gesture.

- (41) Did $\overline{\text{many students}}$ ^{IX} fail the exam? (IX toward John)
 \rightsquigarrow *John is one of the students and he is considered likely to fail*

I submit that the puzzling scalar inference that was discussed for the universal determiners *all* and *no* can, in principle, be accounted for using the same reasoning. Specifically, consider the case of *all*.

- (42) $\overline{\text{All students}}$ ^{IX} failed the exam. (IX toward John)
 \rightsquigarrow *John is one of the students and he was considered unlikely to fail*

If this scalar inference is to be accounted for based on the discussion above about (41), then it must be the case that *all* makes the COMPLEMENT SET available for anaphoric reference; i.e., it must be the case that the set that is exemplified by pointing is the COMPLEMENT SET.

- (43) If there is any student who did *not* fail the exam, then John is one of the students who did not fail the exam. [In common sensical terms, John was very likely *pass* or very *unlikely* to fail.]

But notice that the COMPLEMENT SET of *all* cannot be referred to using a pronoun. This is entirely predicted, however, on the assumption that plural pronouns carry the presupposition that they refer to a non-empty set and given that the assertive content of *all* entails that the COMPLEMENT SET is empty.

There is also indirect evidence for the claim that *all* makes the COMPLEMENT SET salient, coming from negation. As noted in Section 7.2, *not all* allows access to the COMPLEMENT SET. This claim is supported both by facts pertaining to anaphora and facts pertaining to co-nominal pointing, as discussed there. The situation with *all* embedded under sentential negation is exactly the same as *not all*. There are two possibilities pertaining to how negation ‘projects’ the discourse referents that are triggered in its scope (if this process is compositional at all, which I assume it is). Either negation prevents such discourse referents from projecting (this seems to be the common assumption) or, similar to the case of presuppositions, it allows them to project unmodified. There two problems with the former possibility. First, if negation completely blocks discourse referents, then double negation should do the same (i.e., the second occurrence of negation would not be able to ‘recover’ the discourse referents). Evidence against this comes from examples of the following kind. (See Schlenker 2011 for evidence from ASL and LSF that point in the same direction.)

- (44) a. It is not true that John failed none^{*i*} of his students, they_{*i*} are waiting outside of his office.
b. It is not true that Sam doesn’t have an^{*i*} umbrella, it_{*i*} is upstairs in his room.

Furthermore, and more to the point, if merely blocks discourse referents, it is entirely puzzling why *not all* should allow reference to the COMPLEMENT SET. If, on the other hand, negation is thought of as a ‘hole’ for anaphora (much like it is a ‘hole’ for presuppositions), the facts start making more sense: *all* makes the COMPLEMENT SET available for reference, but pronouns cannot refer to this discourse referent due to their existential presuppositions. Co-nominal pointing, on the other hand, does not carry an existential presupposition. Indeed, according to the modification proposed in the last section, exemplification is designed to deal with the emptiness of a given set.

7.6 Co-nominal pointing gestures and co-predicative iconic gestures

Consider the following example from Schlenker 2018.

- (45) John [punished his son]_{SLAP}.
↔ *John punished his son by slapping him*

This example involves (i) an iconic, slapping gesture, which (ii) is temporally aligned with the predicate of the sentence. Consequently, I will refer to such cases as iconic

co-predicative gestures. Schlenker’s (2018) analysis of iconic co-predicative gestures is summarized below.

- (46) If an iconic gesture G modifies the predicate α of a simplex sentence of the form $[x \alpha]$, where x is a referential expression, the presupposition is triggered that if x satisfies the property denoted by α then x satisfies the property denoted by α as conjunctively modified by G .

To be concrete, then, on Schlenker’s analysis, (45) asserts that John punished his son and presupposes that if John punished his son he punished his son by slapping him. Note that, on this analysis, the attested inference of (45) is derived in the sense that the utterance in (45) is predicted to be *true* if and only if John punished his son (assertion) and if he punished his son, he did so by slapping (presupposition), which amounts to John punished his son by slapping him. According to Schlenker, this conditional inference becomes visible when (45) is embedded under negation (see Tieu et al. 2017, 2018 for two experimental investigations on this point).

- (47) It is not the case that John [punished his son]_{SLAP}.
 \rightsquigarrow *If John had punished his son he would have done so by slapping him*

It should be clear at this point that Schlenker’s ‘cosuppositional’ analysis of iconic co-predicative gestures bears a resemblance to the final version of the exemplification-based analysis of co-nominal pointing. More specifically, on both analyses the gestural inference is presuppositional and on both analyses the gestural inference is conditional. The crucial difference between the analyses pertains to the antecedent of the putative inferences. Specifically, according to Schlenker, the inference generated by iconic co-predicative inferences is ‘conditionalized’ on the assertive content of the sentence. In contrast, according to the exemplification-based analysis of co-nominal pointing, the inference generated by these gestures is conditionalized on the non-emptiness of the discourse referent that is being exemplified.

This immediately raises the question of whether one of these analyses can be reduced to the other. I have to leave the possibility of assimilating Schlenker’s cosuppositional analysis to the exemplification-based analysis to future work. Here I would like to point out that a reduction in the opposite direction is probably not desirable. Suppose we make the exemplification hypothesis cosuppositional as follows.

- (48) **The exemplification hypothesis, cosuppositionalized.** If a pointing gesture IX toward individual T_{IX} accompanies a quantificational noun phrase α the scope of which is β , it triggers the presupposition that if $\alpha(\beta)$ is true then $T_{\text{IX}} \in S_{\alpha}$.

To be concrete, we now predict the *many* example to trigger the inference in (49b), cf. the inference predicted by the non-cosuppositional version of the exemplification hypothesis in (49c).

- (49) a. $\overline{\text{IX}}$ Many students failed the exam. (IX to John)
 b. *If there are **many** students who failed then John is one of the students who failed the exam*
 c. *If there are **some** students who failed then John is one of the students who failed the exam*

The difference between the two inferences cannot be detected in the case of (49a), given the assertive content of the sentence. The two come apart, however, when (49a) is embedded in a non-veridical environment, as in (50).

- (50) If the exam involved math then $\overline{\text{many students}}^{\text{IX}}$ failed. (IX to John)
- a. Cosuppositional prediction: *If the exam involved math and if there are **many** students who failed then John is one of the students who failed the exam*
 - b. Original prediction: *If the exam involved math and if there are **some** students who failed then John is one of the students who failed the exam*

I believe that an inspection of intuitions regarding (50) reveals that the non-cosuppositional inference (50b) is more adequate. The cosuppositional version predicts that the utterance in (50) is true if the exam involved math, a few students failed but John did not. My consultants all agreed that (50) would be false in such a situation. Of course, the judgment is delicate and merits further scrutiny. I tentatively conclude, then, that *if* the cosuppositional analysis and the exemplification-based one are to be assimilated, at least one of the two options (i.e., making the exemplification hypothesis cosuppositional) is not the right approach.

On the one hand, above I pointed out that a unified approach to co-nominal pointing and iconic co-predicative gestures on the *theoretical* side is not straightforward. I would like to close this section by pointing out that, on the other hand, the two cases share substantial peculiarities. Specifically, both types of gestural enrichment show the same behavior in non-monotonic and focus-sensitive constructions.

Regarding the non-monotonic case, consider the predicate *unaware*. This predicate has a positive presupposition, (51a), and a negative assertion, (51b), making it non-monotonic in its clausal argument.

- (51) Mary is unaware that some students failed.
- a. *Presupposition*: Some students failed.
 - b. *Assertion*: It is not the case that Mary believes that some students failed.

As pointed out by Schlenker, embedding iconic gestures in the clausal complement of *unaware* leads to an interesting pattern: the gestural inference enriches the presupposition but not the assertive content of the sentence.

- (52) Mary is unaware that John [punished his son]_{SLAP}.
- a. *Presupposition*: John punished his son by slapping him
 - b. *Assertion*: It is not the case that Mary believes that John punished his son (*in any way*)

The same applies to co-nominal pointing.

- (53) Mary is unaware that Bill failed $\overline{\text{many of his students}}^{\text{IX}}$. (IX to John)
- a. *Presupposition*: Bill failed many of his students and John is one of them
 - b. *Assertion*: It is not the case that Mary believes that Bill failed some of his students [i.e., for all Mary knows, Bill did not fail any of his students]

Regarding the focus-sensitivity constructions, as again pointed out by Schlenker,

iconic inferences ‘disappear’ (or at least *can* disappear) under ellipsis, (54), and in the computation of focus-alternatives, (55).

- (54) a. John [punished his son]_{SLAP}, and Bill did too.
 ~> *Bill punished his son, in a way that is possibly different from slapping*
 b. John [punished his son]_{SLAP}, but Bill did not.
 ~> *Bill did not punish his son in any way*
- (55) (Between John and Bill,) only John [punished his son]_{SLAP}.
 ~> *Bill did not punish his son in any way*

The same applies to co-nominal pointing.

- (56) a. Mary failed many of his students^{IX}, and Bill did too. (IX to John)
 ~> *Bill failed many of his students, but John may not have been one of them [indeed, it could be that John is not even one of Bill’s students]*
- b. Mary failed many of his students^{IX}, but Bill did not. (IX to John)
 ~> *Bill did not fail many of his students, no inference pertaining to John [indeed, it could be that John is not even one of Bill’s students]*
- (57) (Between Mary and Bill,) only Mary failed many of his students^{IX}. (IX to John)
 ~> *Bill did not fail many of his students, no inference pertaining to John [indeed, it could be that John is not even one of Bill’s students]*

Does this empirical similarity between pointing and iconic gestures lend support to the idea that the same mechanism underlies both cases? Perhaps, but not necessarily. Note that at least the data pertaining to focus-sensitive constructions can be replicated based on the interpretation of ϕ -features on (bound) pronouns (Heim 2008; Kratzer 2009; Sauerland 2013, a.o.).

- (58) a. Mary_i did her_i homework, and John did too.
 ~> *John did **his** homework*
 b. (Between Mary and John,) [only Mary]_i did her_i homework.
 ~> *John did not do **his** homework*

In this context, the case of *unaware* becomes interesting. Can we replicate (52) and (53) above with pronouns? Clearly, if a pronoun is deictic (i.e., free) this is not possible; in (59), for example, the gender feature on the pronoun cannot be ignored; the inference is inescapable that Mary knows that whoever *she* refers to is female.

- (59) Mary is unaware that she is French. (*she* refers to a female passing by)

Things become more interesting when we look at data pertaining to bound pronouns. As shown below, *regardless of whether the relevant pronoun is bound by a focus-sensitive operator or not*, there is a possible reading in which the gender feature on the pronoun only enriches the presupposition.

- (60) a. *Context: My father threw a big party recently. He knows that one (and, let’s say, only one) friend of mine was at the party, but he does not know whether that friend of mine is male or female.*

My father is unaware that (among the people at the party) [only my friend]_i brought her_i partner.

↗ (i) *My friend is, in fact, female.* (ii) *My father may or may not know this.*

- b. *Context: My father is a famous writer, and he is coming to my class today to give a guest lecture. My students are big fans of his.*

My father is unaware that [every student of mine]_i has brought her_i copy of his book to sign.

↗ (i) *All my students are, in fact, female.* (ii) *My father may or may not know that I have no male students.*

In the first example (60a), the gender feature on the pronoun clearly triggers the inference at root that the friend in question is female. This inference, however, need not be something that the speaker's father believes for the sentence to be true; for all he knows, the friend in question is male. Similarly, in (60b) the gender feature on the pronoun projects universally from the scope of the universal quantifier and is felt clearly at root; the truth of the sentence as a whole by no means requires the speaker's father to have any particular beliefs regarding this inference. As (60b) in particular makes clear, presuppositions triggered by (at least) the gender feature on bound pronouns, embedded under *unaware*, behave in a way that is quite similar to gestural enrichments.

Let me close this section by pointing out that, as of yet, the facts pertaining to gestures embedded in focus-sensitive and non-monotonic environments are not accounted for. Schlenker is quite explicit about this (I refer the reader to his paper), and I may as well point out that the exemplification hypothesis is no better. Finally, the facts in (60) are also not accounted for by any extant analysis of gender features in English, as far as I know. The similarity between these phenomena, then, remains a descriptive generalization at this point.

7.7 Conclusion

In the preliminary remarks of Section 7.1, I made several limitations of this study explicit. However, even the narrow class of examples that were discussed in this paper exhibit at least one potential shortcoming of the exemplification hypothesis. Consider example (61a), from Section 7.4. In this example, co-nominal pointing triggers two inferences, only one of which is accounted for. Specifically, while the conditional inference is accounted for, the inference that the pointing target is one of the students is not. One might suspect that this inference comes about via exemplification of the MAXIMAL SET (as defined in Section 7.2) *on top of the REFERENCE SET*.¹⁰ But notice that if this were the case, we would predict the inference in (61b).

¹⁰In this brief discussion, I will tentatively assume that co-nominal pointing can simultaneously exemplify two distinct sets. Thus, the idea is that in (61a), co-nominal pointing exemplifies both the MAXIMAL SET and the REFERENCE SET. However, given the fact that the REFERENCE SET (and the COMPLEMENT SET) is a subset of the MAXIMAL SET, this assumption might prove unnecessary; that is, it could be that in (61a), the only available inference is the one that results from exemplifying the REFERENCE SET, but certain processes of enrichment (perhaps along the lines sketched below) apply to this inference to strengthen that part of it which pertains to the student-hood of John.

- (61) a. If the exam involved math, did $\overline{\text{many students}}^{\text{IX}}$ fail the exam?
 \rightsquigarrow *John is one of the students*
 \rightsquigarrow *If the exam involved math, and if there is any student who failed, it is John*
- b. *If the exam involved math, and if there are any students, then John is one of the students* (predicted inference)

I will leave this problem to future work; however, I would like to point out that two independently motivated facts might conspire to generate the attested inference on the basis of the predicted inference in (61b). First, it is well-known that presuppositions triggered in the consequent of conditionals are sometimes strengthened. This is known as the proviso problem. Thus in (62) the attested presupposition at root is that John has a sister, not that if John goes to Toronto, he has a sister.

- (62) *If John goes to Toronto, his sister will pick him up from the airport.*
 \rightsquigarrow *John has a sister*

One intuition about this strengthening process pertains to whether the truth or falsehood of the antecedent is relevant to that of the presupposition. Indeed, in (62), it is reasonable to think that John's having a sister is entirely independent from his going to Toronto. Going back to (61a), it seems to me equally reasonable to assume that John's being a student is entirely independent from whether or not the exam in question involved math. Thus, whatever mechanism underlies the strengthening involved in (62) might also apply to the prediction inference in (61b), generating (63). (Note, on the other hand, that we would not expect the same to apply to the other inference: whether or not the exam involved math seems to be very much relevant to whether or not John is one of the students who failed, if any.)

- (63) *If there are any students, then John is one of the students*

The inference in (63) is still too weak. But this time the same reasoning may or may not be possible, depending on the details of the strengthening process. Note that in worlds in which there are no students, John cannot possibly be a student either. Regardless, I think it is reasonable to assume that *many students* triggers the inference / presupposition that there are students. If this latter inference projects to root (as intuitively it does), then the conjunction of this inference with (63) allows us to derive the inference in question.

Finally, let me note that there are obvious ways in which the present paradigm can be systematically extended in a potentially insightful way. At the moment, we have an analysis of co-nominal pointing gestures, on the one hand, and an analysis of co-predicative iconic gestures on the other. What about co-nominal iconic gesture? (Or, co-predicative pointing gestures, for the matter—I have nothing to say about this latter here.) Consider the following example (here I report my own judgments).

- (64) a. [Some of my students]_{TALL} are in the basketball team.
 \rightsquigarrow *The students of mine who are in the basketball team are tall*
- b. [Not all of my students]_{SHORT} are in the basketball team.
 \rightsquigarrow *The students of mine who are not in the basketball team are short*

To begin with, note that the inferences generated by the gestures can be paraphrased

with plural anaphora, much like co-nominal pointing inferences. This is done for (64a) in (65a), and for (64b) in (65b).

- (65) a. Someⁱ of my students are in the basketball team, they_i are tall.
 b. Not allⁱ of my students are in the basketball team, they_i are short.

Furthermore, note that as in the case of both co-nominal pointing and plural anaphora, *some* does not allow the iconic gesture that co-occurs with it to access the COMPLEMENT SET, even at the pain of contextual incoherence. A but more surprisingly, neither does *not all* allow the iconic gesture to predicate something of the REFERENCE SET.

- (66) a. ??[Some of my students]_{SHORT} are in the basketball team.
 b. ??[Not all of my students]_{TALL} are in the basketball team.

These observations, as sketchy as they are, strongly suggest that co-nominal pointing and iconic gestures rely on the same underlying rule. Both predicate something of the discourse referents that the determiner makes salient; iconic gestures such as TALL and SHORT do this by predicating that the set consists of tall / short individuals, while pointing does this by predicating that the set contains the pointing target.

Chapter 8

Dislocated Cosuppositions

8.1 Introduction

As a preliminary illustration of the problem this paper is concerned with, consider the sentence in (1). [On notation: a speech accompanying (or, co-speech) gesture is notated as a subscript in small capitals after the expression it co-occurs with. The modified expression is put between square brackets if it contains several words.]

- (1) a. John [punished his son]_{SLAP}.¹
 ↗ *John punished his son by slapping him*
 b. John [took the elevator]_{UP}.²
 ↗ *John took the elevator to go up*

In each case, the co-occurring gesture enriches the basic meaning of the sentence in a manner that is clearly keyed to its iconic shape. As with any other form of enrichment, there are three questions that one may ask about gestural enrichments: (i) what is the *content* of the gestural enrichments?,³ (ii) what is the *projection profile* of gestural enrichments? (i.e., how do gestural enrichments interact with logical operators), and (iii) what is the *epistemic status* of gestural enrichments? (e.g., must gestural enrichments provide new or old information, relative to background assumptions). Building on Schlenker 2018, this paper aims at contributing to each of these questions. Beginning with question (ii), made more explicit in (2). In order to investigate this question, we need to embed gesturally modified expressions in the scope of logical operators and inquire about the fate of the gestural inference as it projects through these operators. The salient case of negation is given in (3) (see Tieu et al. (2017, 2018) for experimental investigation of the judgment reported here).

- (2) **The projection problem for co-speech gestures.** How are the enrichments of expressions modified by co-speech gestures inherited by complex sentences? (from Schlenker 2018, see also the pioneering work of Ebert & Ebert 2014)
- (3) a. John did not [punish his son]_{SLAP}.
 ↗ *If John had punished his son, he would have done so by slapping*

¹SLAP' stands for a slapping gesture in neutral position, i.e., near the torso.

²UP' stands for an upward movement of the arms.

³For example, the content of the scalar implicature associated with a sentence of the form *some As B* is *not all As B*, the form of the homogeneity inference associated with a sentence of the form *the As B* might be taken to be *either all As B or all As not B*, and so on.

- b. John did not [take the elevator]_{UP}.
 \rightsquigarrow *If John had taken the elevator, he would have done so to go up*

The judgments reported in (3) become sharper once appropriate context is provided. For example, compare an utterance of (3b) “out of the blue” versus in the context specified in (4).

- (4) *Context: the building has ten floors. Mary’s office is on the 5th. We do not know where John’s office is. John does not know where Mary’s office is. He has been looking for her.*
A: Did John manage to find Mary’s office?
B: No . . . he got lost on the 5th floor . . .
A: How did *that* happen? Her office is right in front of the elevator!
B: Well, he didn’t [take the elevator]_{UP}, he used the stairs instead.

The inference suggested in (3b) is quite sharply felt in (4): John’s office (or at least his starting point before he went looking for Mary) is on a floor below the 5th: he did not use the elevator, but if he had done so, he would have gone up. Thus, it appears that even though the intuitively perceived gestural inference in (1a) is simply that the way that John punished his son was by slapping, once the sentence is negated the inference reveals itself directly as a conditional inference: if John punished his son, he did so by slapping. In the case of (1a) the conjunction of this conditional inference with the assertive content of the sentence yields the perceived inference, but in (3a), given that the sentence is embedded under negation, the gestural inference cannot be strengthened by the assertive content of the sentence.

8.2 The cosuppositional analysis

The starting point of this paper is Schlenker’s (2018) “cosuppositional” analysis of gestural enrichments. This analysis takes the form of the judgments provided in (3) quite seriously. With a good deal of simplification, it can be summarized as follows.

- (5) **The Cosuppositional Approach. (hf. CA)** If a predicate α embedded in a sentence ϕ uttered in context C is accompanied by a gesture \mathbf{G} , the local context of α in ϕ relative to context C must entail $\alpha \Rightarrow \mathbf{G}$:⁴

$$\vDash_{lc(\alpha)} \alpha \Rightarrow \mathbf{G}.$$

In words, the cosuppositional analysis requires that the gesturally modified expression must entail the content of the accompanying gesture in its local context. The intuition behind this requirement is that iconic, co-predicative gestures ‘illustrate’ the local meaning of the expressions they modify. ‘Local meaning’ is here understood as semantic denotation relative to a given local context, where the latter is formulated on the basis of Schlenker’s (2009b) theory of local contexts. Thus, if α is some predicate and $lc(\alpha)$ is its local context, the local meaning of α boils down to $lc(\alpha) \wedge \alpha$ (where \wedge is generalized conjunction). ‘Illustration’ is quite simply cashed out as entailment.

⁴Here and throughout, for any expression α , $\alpha = \llbracket \alpha \rrbracket$. For any iconic gesture \mathbf{G} , \mathbf{G} is the model-theoretic object that \mathbf{G} ‘denotes’. I will have nothing to say about how gestures end up with the denotations assumed here.

The requirement, therefore, can be formalized as $\models_{lc(\alpha)\wedge\alpha} \mathbf{G}$ which is equivalent to $\models_{lc(\alpha)} \alpha \Rightarrow \mathbf{G}$. The analysis, thus, is tantamount to saying that a predicate / gesture complex ' α_G ' triggers the presupposition that $\alpha \Rightarrow \mathbf{G}$.

CA answers the three questions posed at the beginning of this paper as follows: (i) gestural enrichments are pieces of information that are conditionalized on the assertive content of the expressions they modify, (ii) gestural enrichments project like presuppositions do in general, and (iii) gestural enrichments receive the same epistemic treatment as root as presuppositions, namely they must be entailed by the common ground (for the utterance to be acceptable).⁵

CA accounts for the judgments reported in (3) immediately. Presuppositions project from under negation, therefore, e.g., (3a) is predicted to put the following requirement on the context C : any world w in C is such that either John did not punished his son in w , or John punished his son by slapping him in w (i.e., C entails that John did not punished his son by any mean other than slapping). Exactly the same prediction is made for the unembedded case, (1a). The predicted net-effect is of course correct: if C entails that John did not punished his son without slapping him, adding the information that John did punish his son will contextually convey that John punished his son by slapping him.⁶

CA also makes welcome predictions for the cases of embedding gesturally modified expressions in the scope of the universal quantifiers *every* and *no*, as shown in (6). As is well-known, presuppositions project universally from the scope of universal quantifiers. The cosupposition associated to the predicate ' $\lambda x. x$ punished_{SLAP} x 's son' is the property $[\lambda x. \mathbf{punished}(x, x\text{'s son}) \Rightarrow \mathbf{slapped}(x, x\text{'s son})]$. Once this presupposition is projected universally to root, one gets the predicted inferences in (6) which line up nicely with the attested inferences.

- (6) a. Each of these ten guys punished_{SLAP} his son.
 \rightsquigarrow *Each of the guys punished his son by slapping him* (attested)
 $\rightsquigarrow \forall x \in \mathbf{guys} : \mathbf{punished}(x, x\text{'s son}) \Rightarrow \mathbf{slapped}(x, x\text{'s son})$ (predicted)
- b. None of these ten guys punished_{SLAP} his son.
 \rightsquigarrow *Each of the guys would have slapped his son, had he punished him* (attested)
 $\rightsquigarrow \forall x \in \mathbf{guys} : \mathbf{punished}(x, x\text{'s son}) \Rightarrow \mathbf{slapped}(x, x\text{'s son})$ (predicted)

However, as Schlenker points out, the predictions made by CA are in some cases *too strong*. This is in particular the case for non-monotonic environments.

- (7) a. Mary is unaware that John punished_{SLAP} his son.
 \rightsquigarrow *John punished his son by slapping him* (attested)
 $\rightsquigarrow \mathbf{punished} \wedge (\mathbf{punished} \Rightarrow \mathbf{slapped}) \wedge B_M(\mathbf{punished} \Rightarrow \mathbf{slapped})$ ⁷ (predicted)
- b. Some but not all of these ten guys punished_{SLAP} their son.

⁵The common ground is the conjunction of all propositions that the interlocutors take for granted at a particular point of a conversation, and the context set is the set of all possible worlds that are compatible with the common ground. Note that just presuppositions can be informative (Stalnaker 2002; von Stechow 2008; Schlenker 2012a), so can gestural enrichments.

⁶Just why the conditional 'force' of the gestural inference is not felt for the unembedded cases in (1) is a question that I will follow Schlenker by ignoring.

⁷Here ' $\mathbf{punished}$ ' is short for ' $\mathbf{punished}(\mathbf{John}, \mathbf{John's son})$ '. Same with ' $\mathbf{slapped}$ '. For any P, ' $B_M(P)$ ' stands for 'Mary believes that P'.

- \rightsquigarrow *Some of the guys punished their son by slapping, the rest did not punished their sons in any way* (attested)
 $\rightsquigarrow \forall x \in \mathbf{guys} : \mathbf{punished}(x, x's\ son) \Rightarrow \mathbf{slapped}(x, x's\ son)$ (predicted)

Consider (7a). It is reasonable to analyze a sentence of the form ‘S is unaware that P’ as presupposing that P and asserting that it is not the case that S believes that P, $\neg B_S$. Therefore, regarding presuppositions triggered in the subordinate clause, we predict that, first, these must project to root (8a) and, second, these must be entailed by the beliefs of the attitude holder (8b).

- (8) Mary is unaware that John has stopped smoking
 a. *John used to smoke but no longer does*
 b. *Mary believes that John used to smoke*

The problem raised by (7a) is that an utterance of (7a) can easily be understood such that only the first of these prediction is born out. The sentence itself presupposes that John punished his son, since the cosupposition that if John punished, he slapped also projects to root, we predict the overall presupposition that John punished his son by slapping him. But, the second prediction (namely, that Mary should believe that John did not punish his son without slapping him), if available at all, is not easily accessible.

The problem raised by (7b) is similar: presuppositions triggered in the scope of the complex determiner ‘some but not all’ project universally to root, (9); consequently, CA predicts, not only that some guys punished their son by slapping, but also that for each of the guys who did not punished their son, if they had done so, they would have slapped. This latter inference is at least not easily accessible (but see the discussion in section 8.4); (7b) can very naturally be understood to imply that those guys who did in fact punish their son did so by slapping, without making any implication about the punishing habits of the other guys.

- (9) Some but not all students have stopped smoking.
 \rightsquigarrow *Every student used to smoke*

In the next section I will discuss a solution to the problems raised in (7) which is formulated by Schlenker himself. Once the limits of that solution are made explicit, I will turn to my own proposal in section 8.4.

8.3 The *supervaluationist* analysis

Let us go back to the problem raised by ‘unaware’ in (7a) repeated below.

- (10) Mary is unaware that John punished_{SLAP} his son.
 \rightsquigarrow *John punished his son by slapping him* (attested)
 $\rightsquigarrow \mathbf{punished} \wedge (\mathbf{punished} \Rightarrow \mathbf{slapped}) \wedge B_M(\mathbf{punished} \Rightarrow \mathbf{slapped})$ (predicted by CA)

Consider the following line of attack. What happens when a gesture modifies an expression, as in (10), is that two propositions are made salient for the audience to choose from. In the case of (10) these could be (11a) and (11b).

- (11) a. That Mary is unaware that John punished his son. $\underline{P} \wedge \neg_{\text{BM}}(P)$ ⁸
 b. That Mary is unaware that John punished his son by slapping him.
 $(\underline{P \wedge S}) \wedge \neg_{\text{BM}}(P \wedge S)$

What would the audience do, when they are faced with such a choice? One possible answer is that the audience are ultra-conservative: they ‘focus attention’ only to those situations in which both propositions in (11) are simultaneously true (/ false). In other words, they assume the speaker would not make an utterance like (10) if he believes that the two propositions in (11) have distinct truth values. The prediction, then, is that an utterance of (10) is true (/ false) iff both propositions (11a) and (11b) are true (/ false), and is undefined otherwise. Interestingly, this prediction is *weaker* than the one made by CA. Since $P \wedge S$ is stronger than P while $\neg_{\text{BM}}(P)$ is stronger than $\neg_{\text{BM}}(P \wedge S)$, (10) is predicted to be true if and only if $(P \wedge S) \wedge \neg_{\text{BM}}(P)$. No problematic inference is predicted pertaining to Mary’s beliefs, as desired.

The general principle underlying the reasoning spelled out in the previous paragraph can be summarized as follows.

- (12) The supervaluationist Analysis. (hf. SA)⁹ Let ϕ be a sentence that contains the predicate α , $\phi = \phi[\alpha]$. An utterance of $\phi[\alpha_G]$ is judged true (false) iff both $\phi[\alpha]$ and $\phi[\alpha \wedge G]$ are true (resp. false).

Here is another example that is adequately dealt with by SA.

- (13) Exactly one of these ten guys punished_{SLAP} his son.

\rightsquigarrow *Exactly one of the guys punished his son by slapping, the rest did not punish their sons in any way*

Since CA is built on the Transparency Theory as its projection engine,¹⁰ it predicts the cosupposition triggered in the scope of *exactly one* in (13) to project universally to root, quite the same as the case of *some but not all*. The result is the correct prediction that one guy punished his son by slapping and the rest did not punish their son and the *incorrect* prediction that for each of the guys who did not punish their son, if they had done so, they would have slapped. Here again, the prediction made by SA is adequately weak; as the reader can easily verify, if an utterance of (13) is true iff both (14a) and (14b) is true, then an utterance of (13) is true iff one guy punished his son by slapping and the rest did not punish their son in any way. No inference is predicted regarding the guys who did not punish their son, as desired.

- (14) a. Exactly one of these ten guys punished his son.
 b. Exactly one of these ten guys punished his son by slapping him.

Unfortunately, SA has problems of its own (which Schlenker also points out). Specifically, the predictions made by SA are sometimes *too weak*, in some cases to the point of triviality. For example, the prediction made for (7b), repeated below, is that it is true iff some guys punished their son by slapping and some guys did not punish

⁸Underlining marks presuppositionality.

⁹This principle is *reminiscent* of the type of reasoning that supervaluationist logics are known for, whence the title.

¹⁰Transparency Theory predicts in general presuppositions triggered in the scope of quantifiers projects universally to root.

their son in any way; this is too weak, as it allows for there being guys who punished their son in some way other than by slapping.

- (15) Some but not all of these ten guys punished_{SLAP} their son.
- a. Some but not all of these ten guys punished their son.
 - b. Some but not all of these ten guys punished their son by slapping him.

Further, when a gesturally modified expression is embedded in a downward-entailing environment, SA predicts *no enrichment* to the *truth*-conditions of the the sentence. For example, (6b), repeated below, is predicted to be true iff none of the guys punished their son in any way. The reason being that since (16a) entails (16b), the requirement that both be true boils down to the requirement that (16a) be true.

- (16) None of these ten guys punished_{SLAP} his son.
- a. None of these ten guys punished his son.
 - b. None of these ten guys punished his son by slapping him.

To recap, the predictions made by CA are sometimes too strong while those made by SA are sometimes too weak. One might wonder whether the two should be put together. There are two main obstacles to this idea. First, SA and CA seem two entirely distinct mechanisms, a marriage between the two (regardless of the exact details) seems hopelessly disjunctive (“conceptually odd” in Schlenker’s words). Second, it is not entirely clear just how the two analyses must be ‘linked’ together anyways. To see this, consider Schlenker’s own suggestion.

- (17) A co-speech gesture is treated in terms of SA (= (12)) unless this fails to strengthen the meaning, in which case it is treated in terms of CA (= (5)).

This way of linking CA and SA immediately runs into a problem with (15): in that case, as I have noted, SA *does* strengthen the meaning, but it does not do so sufficiently. In the next section I will formulate a proposal that solves these two problems (i.e., the linking problem and the problem of conceptual oddity) in one stroke. I will then show that this new principle coupled with a new bridge principle to link the predicted inferences with the background assumptions yields empirically adequate predictions.

8.4 Dislocated Cosuppositions

To spell out my proposal, I need to define several auxiliary notions. Let α be a predicate, and ϕ a sentence that contains (an occurrence of) α . We can construct a sequence β_i of property- or proposition-denoting constituents of ϕ with the following properties: (i) $\beta_0 = \alpha$, (ii) $\beta_n = \phi$, and (iii) for each $i \in \{0, \dots, n - 1\}$, $\beta_i \sqsubseteq \beta_{i+1}$ (β_i is contained in β_{i+1}). Let me call this the formation sequence of ϕ relative to α . Further, given a context set C , we can annotate each β_i with its local context, $lc(\beta_i)$, given Schlenker’s (2009b) algorithm.¹¹ Finally, I need a notion of logical / contextual entailment which applies to property- and proposition-denoting expressions.

- (18) Let β and β' be two expressions of a type that ‘ends in t’ which can take n arguments. let C (the “context”) be a model-theoretic object of the same type.

¹¹The more accurate notation is $lc(C, \beta_i, \phi[\cdot])$.

Then,

- a. $\beta \vDash \beta'$ iff for all objects x_1, \dots, x_n of appropriate types, if $\llbracket \beta \rrbracket(x_1) \dots (x_n) = 1$, then $\llbracket \beta' \rrbracket(x_1) \dots (x_n) = 1$.
 - (i) $\beta \vDash_C \beta'$ iff for all objects x_1, \dots, x_n of appropriate types, if $C(x_1) \dots (x_n) = 1$ and $\llbracket \beta \rrbracket(x_1) \dots (x_n) = 1$, then $\llbracket \beta' \rrbracket(x_1) \dots (x_n) = 1$.

My proposal can now be formulated as follows.

- (19) The Dislocated Cosuppositions Analysis. (hf. DC) Let ϕ be a sentence that contains the predicate α , and let $\langle \beta_0 = \alpha, \dots, \beta_n = \phi \rangle$ be the formation sequence of α relative to ϕ , and let G be some gesture. An utterance of $\phi[\alpha_G]$ is admitted by a context C only if there is some $i \in \{0, \dots, n\}$ such that (i) $\beta_i[\alpha] \not\vDash \beta_i[\alpha \wedge G]$ but (ii) $\beta_i[\alpha] \vDash_{lc(\beta_i)} \beta_i[\alpha \wedge G]$. If felicitous in C , $\phi[\alpha_G]$ is interpreted as $\phi[\alpha]$.

The reasoning that is compressed in (19) can be unpacked as follows. Consider an utterance of $\phi[\alpha_G]$, where α is a predicate and G is a co-occurring gesture. For each constituent β of ϕ that contains α , a “gestural alternative” can be constructed by conjoining the “meaning” of G with α , $\beta[\alpha \wedge G]$.¹² Among these constituents, one can identify those that do not semantically entail their gestural alternatives. Then, the utterance is acceptable in C as soon as one of these constituents *contextually* entails its gestural alternative (in its local context).

I would like to make three remarks immediately. First, it is always the case that the inference generated by $\beta_0 = \alpha$ is identical with the cosupposition predicted by CA. Second, for every example discussed in this paper, the inference generated by $\beta_n = \phi$ is identical with the one generated by SA. This, indeed, is the sense in which CA and SA can be viewed as the outcomes of the same algorithm applied locally and globally. Third, (19) as it stands predicts “intermediate” inferences. I have not been able to construct good examples to establish whether this is a good or a bad prediction, but it should be clear that in case there are no such intermediate inferences (19) can be reformulated to make reference only to the most global and the most local constituents. This issue will not be relevant in the rest of this paper.

I will now work through the examples discussed above to evaluate the predictions of (19). Let us begin with the case of the universal quantifier ‘every’.

- (20) Every one of these ten guys punished_{SLAP} his son.
- a. $\beta_0 = \lambda x. x$ punished x 's son
 $lc(\beta_0) = \lambda w. \lambda x. w \in C \wedge x$ is one of the guys in w
 - b. $\beta_1 =$ [every guy] [$\lambda x. x$ punished x 's son]
 $lc(\beta_1) = \lambda w. w \in C$

In the case of (20) since the local context of the scope (viewed extensionally) is simply the set of all guys, the inference triggered by both (20a) and (20b) boils down to the same; (20a) predicts the inference that for each guy g , if g punished his son, he slapped him and (20b) predicts the inference that if every guy punished his son, then every guy punished his son by slapping. This is of course the same prediction that CA makes, which in conjunction with what the sentence (20) (without the gesture) asserts, yields the attested inference that every guy punished his son by slapping

¹²I am conflating meta- and object-languages here. This is merely to avoid clutter.

him. Next, consider the case of the negative quantifier ‘no’ (which, remember, was problematic for SA).

- (21) None of these ten guys punished_{SLAP} his son.
- a. $\beta_0 = \lambda x. x$ punished x 's son
 $lc(\beta_0) = \lambda w. \lambda x. w \in C \wedge x$ is one of the guys in w
 - b. $\beta_1 =$ [no guy] [$\lambda x. x$ punished x 's son]
 $lc(\beta_1) = \lambda w. w \in C$

Here, no inference is predicted to arise by (21b) because β_1 logically entails $\beta_1[\alpha \wedge G]$ (= [no guy] [$\lambda x. x$ punished x 's son by slapping]), violating the condition (i) of (19). The only option, therefore, is for (21a) to trigger an inference, which, as with (20a), boils down to the presupposition that for each guy g , if g punished his son, he slapped him. This is again the same (correct) prediction that CA makes.

Let me now move on to the case of ‘unaware’ (which was problematic for CA).

- (22) Mary is unaware that John punished_{SLAP} his son.
- a. $\beta_0 =$ John punished his son
 $lc(\beta_0) = \lambda w. \lambda w'. w \in C \wedge w' \in (\text{DOX}_{\text{M}}^w \cup \{w\})$ ¹³
 - b. $\beta_1 =$ Mary is unaware that John punished his son
 $lc(\beta_1) = \lambda w. w \in C$

DC predicts two possible inferences for (22). One option is (22a), which will generate the same the prediction as the one made by CA. The second option is (22b), which will generate the same the prediction as the one made by SA. Before elaborating on this ambiguity, let me also mentioned another example, involving ‘exactly one’.

- (23) Exactly one of these ten guys punished_{SLAP} his son.
- a. $\beta_0 = \lambda x. x$ punished x 's son
 $lc(\beta_0) = \lambda w. \lambda x. w \in C \wedge x$ is one of the guys in w
 - b. $\beta_1 =$ [exactly one guy] [$\lambda x. x$ punished x 's son]
 $lc(\beta_1) = \lambda w. w \in C$

Here again, the inference predicted by (23a) is the same as CA, while it can easily be verified that the inference predicted by (23b) is that of SA. Now, is the ambiguity predicted by DC regarding, e.g., (22) and (23) undesirable? Not necessarily. Although the facts are at the moment rather unclear, ? find that ‘exactly one’ at least sometimes gives rise to universal inferences. The important point, for my purposes was to construct a system which can derive the inferences that Schlenker’s CA could not. But the resulting system predicts systematic ambiguity. The evaluation of this prediction needs to be postponed until the facts are cleared up.

Finally, let me point out that one problem still remains, having to do with ‘some but not all’ (the same point can be made with ‘between n and m ’, ‘an odd number of’, etc.).

- (24) Some but not all of these ten guys punished_{SLAP} his son.
- a. $\beta_0 = \lambda x. x$ punished x 's son

¹³For a proof that the local context of the clause that is embedded under ‘unaware’ is the one given here, see Schlenker (to appear). $w' \in \text{DOX}_{\text{M}}^w$ iff w' is compatible with what Mary believes in w .

- $$lc(\beta_0) = \lambda w. \lambda x. w \in C \wedge x \text{ is one of the guys in } w$$
- b. $\beta_1 = [\text{some but not all guy}] [\lambda x. x \text{ punished } x\text{'s son}]$
 $lc(\beta_1) = \lambda w. w \in C$

The problem is that since the predictions made by DC match those made by CA and SA, DC *cannot* account for (24); the prediction made on the basis of (24a) is too strong while the one made on the basis of (24b) is too weak. This is indeed the same problem that Schlenker’s proposal (17) was faced with. To solve this problem, I’d like to submit that inferences triggered by DC do not receive the same epistemic treatment as root as presuppositions. It is a common assumption, following Stalnaker, that, at root, presuppositions are epistemically interpreted as in (25).

- (25) Stalnaker’s Bridge Principle. If a sentence ϕ presupposes that p , it can be felicitously used in context C only if C entails p .

I would like to propose that DC-triggered inferences are epistemically ambiguous in the following sense. Intuitively, for a sentence ϕ to be acceptable in context C , (25) requires that the presupposition of ϕ be true at every world of C . I would like to claim that DC-triggered inferences come with the following requirement: either every world of C makes the DC-triggered inferences true or every world of C in which the assertive content of the sentence (without the gesture) is true makes the DC-triggered inferences true. Let me implement this idea. Let W be the set of all possible worlds, and $\phi[\alpha_G]$ a sentence that contains a predicate-accompanying gesture. Construct the set C^* such that (i) C^* admits $\phi[\alpha_G]$ and (ii) no super-set of C^* admits $\phi[\alpha_G]$. Then, $\phi[\alpha_G]$ can be felicitously used in a context C only if either $C \subseteq C^*$ or $(C \cap \{w : \llbracket \phi[\alpha] \rrbracket^w\}) \subseteq C^*$.

Let me briefly show why this move solves the problems of (24). Regarding the inference generated by (24a) in the scope of ‘some but not all’, we now have two options as to its epistemic treatment. Option one is that we impose the universal inference (that for each of the guys g , if g punished his son, he did so by slapping him) on the common ground, as we have been doing all along. This of course generates undesirable inferences regarding the guys who did not punish their son. Option two is to require the following: every world in the Context Set which makes the sentence ‘some but not all of these ten guys punished his son’ true, must make the inference that for each of the guys g , if g punished his son, he did so by slapping him true as well. This second option is a weaker imposition on the common ground than the first; for example, it is allowed that there be a world in the context set in which all guys punished their son by pulling his ear. What is required is that if some but not all guys punished their son, then all of them did so by slapping him, which is of course the target inference.

8.5 Conclusion

Co-speech gestures have only recently been studied by formal semanticists. Ebert (& Ebert) and Schlenker take a healthy attitude towards this freshly noticed phenomena: they try to assimilate them to better known phenomena (appositives in the case of Ebert (& Ebert), presuppositions in the case of Schlenker) and study how they diverge. The attitude taken in this paper was to build on the disciplined approach of Schlenker in particular and ask the following question: what is the minimum amount of change

that the cosuppositional analysis must go through, to make it empirically adequate? The resulting system is certainly rather baroque. My hope is that its empirical force can be used a basis to build a conceptually more elegant system.

Bibliography

- Anand, Pranav. 2006. *De De Se*. Cambridge, MA, USA: Massachusetts Institute of Technology dissertation. <http://hdl.handle.net/1721.1/37418>.
- Anand, Pranav & Andrew Nevins. 2004. Shifty operators in changing contexts. In Robert B. Young (ed.), *Proceedings of the 14th Semantics and Linguistic Theory (SALT) Conference*, 20–37. doi:10.3765/salt.v14i0.2913.
- Anvari, Amir. 2015. Scalar implicatures and anti-presuppositions: a unified approach. Term Paper (supervisor: B. Spector).
- Anvari, Amir. 2016. *Co-nominal Pointing: toward a formal semantic analysis*: École Normale Supérieure MA thesis.
- Anvari, Amir. 2018. Logical integrity: from *Maximize Presupposition!* to *Mismatching Implicatures*. Ms. Insititute Jean Nicod. Paris.
- Bartsch, Renate. 1973. ‘Negative transportation’ gibt es nicht. *Linguistische Berichte* 27. 1–7.
- Bogard, S. 2006. El clítico se. Valores y evolución. In Concepción Company (ed.), *Sintaxis histórica de la lengua española. primera parte: La frase verbal*, 755–870. México: Fondo de Cultura Económica.
- Chemla, Emmanuel. 2008a. An epistemic step for anti-presuppositions. *Journal of Semantics* 25(2). 141–173.
- Chemla, Emmanuel. 2008b. An epistemic step for anti-presuppositions. *Journal of Semantics* 25(2). 141–173. doi:10.1093/jos/ffm017.
- Chemla, Emmanuel, Paul Égré & Benjamin Spector. 2017. Characterizing logical consequence in many-valued logic. *Journal of Logic and Computation* 27(7). 2193–2226.
- Chierchia, Gennaro, Danny Fox & Benjamin Spector. 2009. Hurford’s constraint and the theory of scalar implicatures. In *Presuppositions and implicatures*, 47–62.
- Chomsky, Noam. 1993a. *Lectures on government and binding: The pisa lectures*. Berlin: Mouton de Gruyter.
- Chomsky, Noam. 1993b. A minimalist program for linguistic theory. In Kenneth Hale & Samuel Jay Keyser (eds.), *The view from building 20: Essays on linguistics in honor of Sylvain Bromberger*, 1–52. Cambridge, MA: MIT Press.
- Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA, USA: MIT Press.
- Collins, Chris & Paul M Postal. 2014. *Classical neg raising: An essay on the syntax of negation*. Cambridge, MA: MIT Press.
- Collins, Chris & Paul Martin Postal. 2012. *Imposters: A study of pronominal agreement*. Cambridge, MA, USA: MIT Press.
- Collins, James N. 2019. The polarity of clauses embedded under neg-raising predicates. In *Proceedings of sinn und bedeutung 23*, Universitat Autònoma de Barcelona.
- Deal, Amy Rose. 2017. Shifty asymmetries: universals and variation in shifty indexicality. Berkeley: University of California (manuscript). <http://linguistics>.

- berkeley.edu/~ardeal/papers/Deal-indexicals-2017.pdf.
- Dieuleveut, Anouk, Emmanuel Chemla & Benjamin Spector. 2016. Weak and strong quantity implicatures—an experimental investigation. Ms. LSCP & IJN.
- DiTullio, A. 2018. Verbos estativos y su combinación con el se no argumental. In *Encuentro de gramática generativa*, Buenos Aires.
- Dubinsky, Stanley & Robert Hamilton. 1998. Epithets as Antilogophoric Pronouns. *Linguistic Inquiry* 29(4). 685–693. doi:10.1162/002438998553923.
- Ebert, Cornelia & Christian Ebert. 2014. Gestures, demonstratives, and the attributive / referential distinction. Handout of a talk given at Semantics and Philosophy in Europe (SPE 7), Berlin.
- Egré, Paul. 2008. Question-embedding and factivity. *Grazer Philosophische Studien* 77(1). 85–135.
- von Fintel, Kai. 1999. Npi licensing, strawson entailment, and context dependency. *Journal of Semantics* 16(2). 97–148.
- von Fintel, Kai. 2008. What is Presupposition Accommodation, Again? *Philosophical Perspectives* 22(1). 137–170. doi:10.1111/j.1520-8583.2008.00144.x.
- Fox, Danny. 1999. Reconstruction, Binding Theory, and the Interpretation of Chains. *Linguistic Inquiry* 30(2). 157–196. doi:10.1162/002438999554020.
- Fox, Danny. 2002. Antecedent-Contained Deletion and the Copy Theory of Movement. *Linguistic Inquiry* 33(1). 63–96. doi:10.1162/002438902317382189.
- 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 anality 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.
- Gajewski, Jon Robert. 2007. Neg-raising and polarity. *Linguistics and Philosophy* 30(3). 289–328. doi:10.1007/s10988-007-9020-z.
- Glass, Lelia. 2019. The curious case of the negatively biased Mandarin belief verb “yiwei”. Ms. Available at <https://ling.auf.net/lingbuzz/002600>. <https://ling.auf.net/lingbuzz/002600>.
- Haida, Andreas. 2007. *The Indefiniteness and Focusing of Wh-Words*. Berlin: Humboldt-Universität dissertation.
- Harley, Heidei & Elizabeth Ritter. 2002. Person and Number in Pronouns: A Feature-Geometric Analysis. *Language* 78(3). 482–526. doi:10.1353/lan.2002.0158.
- Heim, Irene. 1983a. File change semantics and the familiarity theory of definiteness. *Semantics: Critical Concepts in Linguistics* 108–135.
- Heim, Irene. 1983b. On the projection problem for presuppositions. In *Wccfl 2*, 114–125.
- Heim, Irene. 1991a. Artikel und Definitheit. In Arnim von Stechow & D. Wunderlich (eds.), *Handbuch der Semantik*, Berlin: de Gruyter.
- Heim, Irene. 1991b. Artikel und Definitheit. In Arnim Von Stechow (ed.), *Semantics: An international handbook of contemporary research.*, 487–535. Berlin: de Gruyter.
- Heim, Irene. 1991c. Class handouts on ‘Control’ (Seminar taught by Heim and Higginbotham, Spring 1991): ‘Interpretation of PRO’ (Feb. 22), ‘The first person’ (March 8), and a handout on ‘connectedness’ (April 12).
- Heim, Irene. 2008. Features on bound pronouns. In Daniel Harbour, David Adger & Susana Béjar (eds.), *Phi theory: Phi-features across modules and interfaces*, 35–56. Oxford University Press.
- Holton, Richard. 2017. I—Facts, Factives, and Contrafactuals. *Aristotelian Society*

- Supplementary Volume* 91(1). 245–266. doi:[10.1093/arisup/akx003](https://doi.org/10.1093/arisup/akx003).
- Horn, Larry. 1978. Remarks on neg-raising. In P. Cole (ed.), *Syntax and semantics 9: Pragmatics*, New York: Academic Press.
- Hsiao, Pei-Yi Katherine. 2017. On counterfactual attitudes: a case study of Taiwanese Southern Min. *Lingua Sinica* doi:[10.1186/s40655-016-0019-7](https://doi.org/10.1186/s40655-016-0019-7).
- Kahnemuyipour, Arsalan. 2001. On *Wh*-questions in Persian. *Canadian Journal of Linguistics / Revue Canadienne De Linguistique* 46(1–2). 41–62. doi:[10.1017/S000841310001793X](https://doi.org/10.1017/S000841310001793X).
- Kaplan, David. 1977. Demonstratives. In Almog, Perry & Wettstein (eds.), *Themes from Kaplan (1989)*, Oxford University Press.
- 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*, .
- Kendon, Adam & Laura Versante. 2003. Pointing by Hand in “Neapolitan”. In Sotaro Kita (ed.), *Pointing: Where Language, Culture, and Cognition Meet*, chap. 6. New York, USA: Psychology Press. doi:[10.4324/9781410607744](https://doi.org/10.4324/9781410607744).
- Kierstead, G. 2014. The interaction between propositional attitude verbs and reportative evidentiality in tagalog. In *Proceedings of chicago linguistic society 48*, .
- Kratzer, Angelika. 2009. Making a Pronoun: Fake Indexicals as Windows into the Properties of Pronouns. *Linguistic Inquiry* 40(2). 187–237. doi:[10.1162/ling.2009.40.2.187](https://doi.org/10.1162/ling.2009.40.2.187).
- Križ, Manuel. 2015. *Aspects of homogeneity in the semantics of natural language*: University of Vienna dissertation.
- 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.
- Lazard, Gilbert. 1992. *A grammar of contemporary persian*. Mazda Publishers.
- Lewis, David. 1979a. Attitudes *De Dicto* and *De Se*. *The Philosophical Review* 88(4). 513–543. doi:[10.2307/2184843](https://doi.org/10.2307/2184843).
- Lewis, David. 1979b. Scorekeeping in a language game. *Semantics from Different Points of View* 172–187. doi:[10.1007/978-3-642-67458-7_12](https://doi.org/10.1007/978-3-642-67458-7_12).
- Lewis, Davis. 1980. Index, context, and content. *Philosophy and Grammar* 79–100. doi:[10.1007/978-94-009-9012-8_6](https://doi.org/10.1007/978-94-009-9012-8_6).
- Magri, Giorgio. 2009a. A theory of individual-level predicates based on blind mandatory scalar implicatures. *Natural Language Semantics* 2009(7). 245–297.
- Magri, Giorgio. 2009b. A theory of individual-level predicates based on blind mandatory scalar implicatures. *Natural Language Semantics* 17(3). 245–297. doi:[10.1007/s11050-009-9042-x](https://doi.org/10.1007/s11050-009-9042-x).
- Magri, Giorgio. 2009c. *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. 2017a. *Implicatures in the DP domain*: MIT dissertation.
- Marty, Paul P. 2017b. *Implicatures in the DP domain*. Cambridge, MA, USA: Massachusetts Institute of Technology dissertation. <http://hdl.handle.net/1721.1/1113778>.

- Mayr, Clemens. 2017. Predicting polar question embedding. In Robert Truswell, Chris Cummins, Caroline Heycock, Brian Rabern & Hannah Rohde (eds.), *Proceedings of sinn und bedeutung 21*, 863–880. University of Edinburgh.
- Merchant, Jason. 2000. Economy, The Copy Theory, and Antecedent-Contained Deletion. *Linguistic Inquiry* 31(3). 566–575. doi:10.1162/002438900554460.
- Meyer, Marie-Christine. 2013. *Ignorance and grammar*. Cambridge, MA, USA: Massachusetts Institute of Technology dissertation. <http://hdl.handle.net/1721.1/84420>.
- Milsark, Gary L. 1974. *Existential sentences in English*: Massachusetts Institute of Technology dissertation. <http://hdl.handle.net/1721.1/13021>.
- Nouwen, Rick. 2003. *Plural Pronominal Anaphora in Context: Dynamic Aspects of Quantification*. Utrecht, Netherlands: Utrecht Institute of Linguistics dissertation. https://www.lotpublications.nl/Documents/84_fulltext.pdf.
- Patel-Grosz, Pritty. 2014. Epithets as *De Re* Pronouns. In Christopher Piñón (ed.), *Empirical issues in syntax and semantics*, vol. 10, http://www.cssp.cnrs.fr/eiss10/index_en.html.
- Percus, Orin. 2000. Constraints on Some Other Variables in Syntax. *Natural Language Semantics* 8(3). 173–229. doi:10.1023/a:1011298526791.
- Percus, Orin. 2006a. 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. 2006b. Antipresuppositions. In A. Ueyama (ed.), *Theoretical and empirical studies of reference and anaphora: Toward the establishment of generative grammar as empirical science*, Tokyo: Japan Society for the Promotion of Science.
- Percus, Orin. 2010. Antipresuppositions revisited. Handout of a talk given at CRISCO, Université Caen Normandie. https://www.academia.edu/1468709/Antipresuppositions_Revisited_Handout_2010_Caen_talk_.
- Percus, Orin & Uli Sauerland. 2003a. On the LFs of Attitude Reports. In Matthias Weisgerber (ed.), *Proceedings of the Conference “SuB7–Sinn und Bedeutung”*, vol. 114 Arbeitspapiere des Fachbereichs Sprachwissenschaft, 228–242. Universität Konstanz, Germany. <http://ling.uni-konstanz.de/pages/conferences/sub7/>.
- Percus, Orin & Uli Sauerland. 2003b. Pronoun movement in dream reports. In Makoto Kadowaki & Shigeto Kawahara (eds.), *Proceedings of the 33th North East Linguistic Society (NELS) Conference*, Massachusetts Institute of Technology.
- Perry, John. 1979. The Problem of the Essential Indexical. *Noûs* 13(1). 3–21. doi:10.2307/2214792.
- Podobryaev, Alexander. 2014. *Persons, Imposters, and Monsters*. Cambridge, MA, USA: Massachusetts Institute of Technology dissertation. <http://hdl.handle.net/1721.1/87498>.
- Quer, Josep. 2009. Twists of mood: The distribution and interpretation of indicative and subjunctive. *Lingua* 119(12). 1779–1787. doi:10.1016/j.lingua.2008.12.003.
- Romoli, Jacopo. 2013. A scalar implicature-based approach to neg-raising. *Linguistics and Philosophy* 36(4). 291–353. doi:10.1007/s10988-013-9136-2.
- van Rooij, Robert & Katrin Schulz. 2004. Exhaustive interpretation of complex sentences. *Journal of Logic, Language and Information* 13(4). 491–519. doi:10.1007/s10849-004-2118-6.
- Sauerland, U. 2008a. Implicated Presuppositions. In Anita Steube (ed.), *The discourse potential of underspecified structures*, 581–600. Berlin: Mouton de Gruyter.

- Sauerland, Uli. 2004. The interpretation of traces. *Natural Language Semantics* 12(1). 63–127. doi:10.1023/b:nals.0000011201.91994.4f.
- Sauerland, Uli. 2008b. Implicated presuppositions. In A. Steube (ed.), *Sentence and context: Language, context, and cognition*, .
- Sauerland, Uli. 2013. Presuppositions and the Alternative Tier. In Todd Snider (ed.), *Proceedings of the 23rd Semantics and Linguistic Theory (SALT) Conference*, vol. 23, 156–173. doi:10.3765/salt.v23io.2673.
- Schlenker, Philippe. 1999. *Propositional attitudes and indexicality (a crosscategorical approach)*. Cambridge, MA, USA: Massachusetts Institute of Technology dissertation. <http://hdl.handle.net/1721.1/9353>.
- Schlenker, Philippe. 2003. A Plea for Monsters. *Linguistics and Philosophy* 26(1). 29–120. doi:10.1023/a:1022225203544.
- Schlenker, Philippe. 2004. Person and binding (a partial survey). *Rivista di Linguistica* 16(1). 155–218.
- Schlenker, Philippe. 2009a. Local contexts. *Semantics and Pragmatics* .
- Schlenker, Philippe. 2009b. Local Contexts. *Semantics and Pragmatics* 2. 1–78. doi:10.3765/sp.2.3.
- Schlenker, Philippe. 2011. Donkey anaphora: the view from sign language (ASL and LSF). *Linguistics and Philosophy* 34(4). 341–395. doi:10.1007/s10988-011-9098-1.
- Schlenker, Philippe. 2012a. Maximize Presupposition and Gricean reasoning. *Natural Language Semantics* 4(2). 391–429. doi:10.1007/s11050-012-9085-2.
- Schlenker, Philippe. 2012b. Maximize presupposition and gricean reasoning. *Natural Language Semantics* 20(4). 391–429.
- Schlenker, Philippe. 2012c. Maximize Presupposition and Gricean reasoning. *Natural Language Semantics* 20(4). 391–429. doi:10.1007/s11050-012-9085-2.
- Schlenker, Philippe. 2018. Gesture projection and cosuppositions. *Linguistics and Philosophy* 41(3). 295–365. doi:10.1007/s10988-017-9225-8.
- Schlenker, Philippe. to appear. Gesture projection and cosuppositions. *Linguistics & Philosophy* .
- Sharvit, Yael. 2010. Covaluation and unexpected BT effects. *Journal of Semantics* 28(1). 55–106. doi:10.1093/jos/ffq012.
- 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. 2003. Scalar implicatures: Exhaustivity and Gricean reasoning. *European Summer School in Logic, Language and Information (student session)* 8. 277–288.
- 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. 2017a. Presupposed ignorance and exhaustivity: how scalar implicatures and presuppositions interact. *Linguistics and Philosophy* .
- Spector, Benjamin & Yasutada Sudo. 2017b. Presupposed Ignorance and exhaustification: how scalar implicatures and presuppositions interact. *Linguistics and Philosophy* 40(5). 473–517. doi:10.1007/s10988-017-9208-9.

- Spector, Benjamin & Yasutada Sudo. 2017c. Presupposed ignorance and exhaustification: how scalar implicatures and presuppositions interact. *Linguistics and Philosophy* 40(5). 473–517. doi:[10.1007/s10988-017-9208-9](https://doi.org/10.1007/s10988-017-9208-9).
- Stalnaker, Robert. 1978. Assertion. *Syntax and Semantics* 9. 315–332.
- Stalnaker, Robert. 2002. Common ground. *Linguistics and Philosophy* 25(5/6). 701–721. doi:[10.1023/a:1020867916902](https://doi.org/10.1023/a:1020867916902).
- von Stechow, Arnim. 2004. Binding by Verbs: Tense, Person and Mood under Attitudes. In Horst Lohnstein & Susanne Trissler (eds.), *The syntax and semantics of the left periphery*, 431–488. Berlin–New York: Mouton de Gruyter.
- Takahashi, Shoichi. 2010a. Traces or Copies, or Both – Part I: Characterizing Movement Properties. *Language and Linguistics Compass* 4(11). 1091–1103. doi:[10.1111/j.1749-818X.2010.00255.x](https://doi.org/10.1111/j.1749-818X.2010.00255.x).
- Takahashi, Shoichi. 2010b. Traces or Copies, or Both – Part II: Introducing Further Considerations. *Language and Linguistics Compass* 4(11). 1104–1115. doi:[10.1111/j.1749-818X.2010.00256.x](https://doi.org/10.1111/j.1749-818X.2010.00256.x).
- Theiler, Nadine, Floris Roelofsen & Maria Aloni. 2019. Picky predicates: Why *believe* doesn't like interrogative complements and other puzzles. *Natural Language Semantics* 1–40. doi:<https://doi.org/10.1007/s11050-019-09152-9>.
- Tieu, Lyn, Robert Pasternak, Philippe Schlenker & Emmanuel Chemla. 2017. Co-speech gesture projection: Evidence from truth-value judgment and picture selection tasks. *Glossa: A Journal of General Linguistics* 2(1). 102. doi:[10.5334/gjgl.334](https://doi.org/10.5334/gjgl.334).
- Tieu, Lyn, Robert Pasternak, Philippe Schlenker & Emmanuel Chemla. 2018. Co-speech gesture projection: Evidence from inferential judgments. *Glossa: A Journal of General Linguistics* 3(1). 109. doi:[10.5334/gjgl.580](https://doi.org/10.5334/gjgl.580).