The Puzzle of *Reflexive Belief Construction* in Spanish[1]

Amir ANVARI — *IJN, IEC, ENS, EHESS, PSL Research University, CNRS*
Mora MALDONADO — *Center for Language Evolution, University of Edinburgh*
Andrés SORIA RUIZ — *IJN, IEC, ENS, EHESS, PSL Research University, CNRS*

Abstract. The Reflexive Belief Construction in Spanish, built by adjoining the reflexive pronoun *se* to the verb *creer* (*believe*), displays a puzzling interpretive 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 (Heim 1991), 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. 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. We leave the resolution of this dilemma to future work.

Keywords: (contra-)factivity, belief reports, neg-raising, presuppositions.

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).[2] 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.[3,4]

[1]We would like to thank Paul Egré, Benjamin Spector, Rachel Dudley, and the audiences at SuB 23, the Workshop on Modal and Evaluative Expressions held at SADAF (Argentina), and the 19th Szklarska Poreba Workshop.

[2]The 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 (Di Tullio 2018), but their relation with other uses of the reflexive remains to be explicated.

[3]The 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.

[4]Several recent studies focus on attitudinal predicates that show a similar "negative bias" as *creerse* (Kierstead 2014 for *akala* in Tagalog, Hsiao 2017 for *liah-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 *liah-tsun* share substantial similarities, although there are at least two crucial differences (see footnotes 10 and 17), and data for a complete comparison are sorely lacking. Clearly much more detailed semantic fieldwork is required in this area.

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

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.*
⇝ it is not raining
b. Juan believes that it is raining.
⇝ it is not raining

The standard account of the inferences in (3) involves the principle *Maximize Presupposition!* (hf. MP, Heim 1991). 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 2.1, and in subsections 2.2 to 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 3 and 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 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 3.1, allows us to explain the differences between RBCs and regular belief reports discussed in section 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*.[5]

(4) *Juan no **se** cree que está lloviendo.*
'Juan doesn't REFL believe that it's raining.'
⇝ it is raining

[5]There 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 licenses 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 (Egre 2009).

As we will argue in subsection 3.2, the contrafactivity analysis can only account for this "polarity reversal" effect under negation if it is coupled with the *syntactic* account of neg-raising (Collins and Postal 2014).[6]

In section 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 4.2. 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 5, we take a step back and provide a discussion of some further empirical issues. First, in subsection 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. In subsection 5.2, we point out some differences between RBCs and their dative alternatives. Finally, in subsection 5.3, we report some preliminary evidence suggesting that the predicate *s'imaginer* in French has certain properties in common with RBCs. Section 6 concludes.

Before we move on, a cautionary note 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.

2. Why not *Maximize Presupposition!*

2.1. Background on *Maximize Presupposition!*

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

(5) *Maximize Presupposition!* (hf. MP): Use the alternative that has the strongest presup-

[6] It 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.

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

position, 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. $[\![\,\text{believe}\,]\!]^w = \lambda P_{st}\lambda x_e.\,\text{BEL}_x^w(P)$
b. $[\![\,\text{know}\,]\!]^w = \lambda P_{st}\lambda x_e : P(w) = 1.\,\text{BEL}_x^w(P)$

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 (2008) calls the *epistemic step*. It essentially consists of the assumption that the speaker is opinionated 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

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

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 2.2); that MP-type inferences are epistemically weak while RBC-inferences are not (subsection 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 2.4).

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 can*not* 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 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!'

[9] As 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 that (=*yǐwéi*) she's a billionaire . . . and she actually is.'

[10] As 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 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'

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 (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* believes that *P* is true (in particular, if *x* believes that *P* is false) then *x* does *not* believe that *P* is common ground. This

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

[12] In 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 5.1). We have to leave this issue to future research.

[13] As 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.'

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.

2.4. Projection

MP-inferences are known to project existentially from the scope of universal quantifiers (Sauerland 2008, a.o.). More specifically, and putting aside epistemic issues for simplicity, a plain belief report embedded under a universal quantifier such as *every 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.
'Every student believes that she passed'.

The third difference between RBCs and regular belief reports is that the former show a different behavior in the same environment, as witnessed by the infelicity of (17a). 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.*
'Every 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.*
'Every 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.*
'Every swimmer believes that she has won the race'.

[14] Again, 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.'

b. # *Cada una de las nadadoras* ***se*** *cree que ha ganado la carrera.*
'Every 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)).

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.[15]

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 3 we explore the hypothesis that RBCs are "contrafactive", in that they presuppose the falsity of the clausal complement. In section 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.

3. The Contrafactivity Hypothesis

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) $[\![\text{creer}]\!]^w = \lambda P \lambda x.\, \text{BEL}_x^w(P)$

(20) $[\![\text{saber}]\!]^w = \lambda P \lambda x : P(w) = 1.\, \text{BEL}_x^w(P)$

(21) $[\![\text{creerse}]\!]^w = \lambda P \lambda x : P(w) = 0.\, \text{BEL}_x^w(P)$

[15] There are implementations of *Maximize Presupposition!* that come closer to capturing RBC-inferences. Magri (2009) and Marty (2017) 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 paper are concerned, although the issue needs to be considered more carefully.

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.[16] 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 demosntrate that this prediction is borne out.

In subsection 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 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.

(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 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) *Every student knows that he passed the exam.*
$\rightsquigarrow$ every student passed the exam

(26) *Cada estudiante se cree que ha aprobado.*
'Every student REFL believes that she passed.'
$\rightsquigarrow$ 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

[16] Holton 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.

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.

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 1, the polarity of RBC inferences reverses under negation.[17]

(29) a. *Juan se cree que está lloviendo.*
'Juan REFL believes that it's raining.'
⇝ it is not raining
b. *Juan **no** se cree que está lloviendo.*
'Juan doesn't REFL believe that it's raining.'
⇝ 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.

[17] As 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*.

(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.'

(31) *Ningún estudiante se cree que ha suspendido.*
No student REFL believes that she has failed.
⇝ 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 and 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]

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) $[\![\text{ creer/believe }]\!]^w = \lambda P \lambda x : \textsc{Bel}_x^w(P) \vee \textsc{Bel}_x^w(\neg P).\, \textsc{Bel}_x^w(P)$

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: $\textsc{Bel}_j^w(r) \vee \textsc{Bel}_j^w(\neg r)$
ASSERTION: $\neg \textsc{Bel}_j^w(r)$
$\Rightarrow$ true iff $(\textsc{Bel}_j^w(r) \vee \textsc{Bel}_j^w(\neg r)) \wedge \neg \textsc{Bel}_j^w(r)$ iff $\textsc{Bel}_j^w(\neg r)$

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

(38) $[\![\text{ creerse }]\!]^w = \lambda P \lambda x : P(w) = 0 \wedge (\textsc{Bel}_x^w(P) \vee \textsc{Bel}_x^w(\neg P)).\, \textsc{Bel}_x^w(P)$

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. $\textsc{Bel}_j^w(r) \vee \textsc{Bel}_j^w(\neg r)$
2. $r(w) = 0$
ASSERTION: $\neg \textsc{Bel}_j^w(r)$
$\Rightarrow$ true iff $r(w) = 0 \wedge (\textsc{Bel}_j^w(r) \vee \textsc{Bel}_j^w(\neg r)) \wedge \neg \textsc{Bel}_j^w(r)$ iff $r(w) = 0 \wedge \textsc{Bel}_j^w(\neg r)$

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. The Enriched Excluded Middle Hypothesis

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. $[\![\text{ creer/believe }]\!]^w = \lambda P \lambda x : \textsc{Bel}_x^w(P) \vee \textsc{Bel}_x^w(\neg P).\, \textsc{Bel}_x^w(P)$
b. $[\![\text{ creerse }]\!]^w = \lambda P \lambda x : (P(w) = 0 \wedge \textsc{Bel}_x^w(P)) \vee (P(w) = 1 \wedge \textsc{Bel}_x^w(\neg P)).\textsc{Bel}_x^w(P)$

(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 2.4, and the fact that RBC inferences are epistemically strong, as discussed in 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 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.

4.2. The problem with Polar Questions

The major difficulty pertains to projection. As pointed out in at the end of section 3.1, RBC inferences project from polar questions.

(41) *¿Se cree Juan que está lloviendo?*
'Does Juan REFL believe that it is raining?'
$\rightsquigarrow$ 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.

4.3. 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.

5. Some Further Empirical Observations

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.[18]

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.[19] 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.'
$\rightsquigarrow$ María falsely believes that she is very bad.

[18]The 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. 2018, a.o.). For space limitations we have to leave this to future work.

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

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 1993), we consider these constructions plain belief reports, and the corresponding inference is just an MP-type inference.

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?[20]

(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 2008). The RBC in Spanish could then be analyzed by saying that the "inducer" 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*.

(1) *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:

(47) ?? *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.

(48) 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 (48a) trigger the inference that it is, in fact, raining. Crucially, this inference does not arise from the polar question in (48b), indicating that it is not as strong as the RBC-inference.

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

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*):

(49) *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 2). 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*):[21]

(50) *Juan se piensa que está lloviendo.*
'Juan REFL thinks that it is raining.'
⇝ it's not, in fact, raining

Crucially, (50) 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 (Di Tullio 2018). The construction with *pensar* was left outside the scope of this paper because it presents some minor differences with RBC, specifically under 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.

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 Egré). As illustrated in (51), French *s'imaginer* gives rise to similar inferences to Spanish *creerse*.

(51) a. *Jean s'imagine qu'il est en train de pleuvoir.*
'Jean REFL thinks it's raining.'
⇝ it is not raining.

b. *Jean ne s'imagine pas qu'il est en train de pleuvoir.*
'Jean doesn't REFL think it's raining. '
⇝ 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

[21] 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*.

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.

(52) a. *Me imagino que está lloviendo.*
'I REFL think it's raining.'
b. *Juan se imagina que tengo una hermana, y tiene razón!*
'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.

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 5, a more sophisticated, compositional account of RBCs needs to be made to account for both cross-linguistic and within-language data.

References

Anvari, A. (2018). Logical integrity: from *Maximize Presupposition!* to *Mismtaching Implicatures*.

Bartsch, R. (1973). Negative transportation" gibt es nicht. *Linguistische Berichte 27*(7), 1–7.

Bogard, S. (2006). El clítico se. Valores y evolución. In *Sintaxis histórica de la lengua española. Primera parte: La frase verbal*, pp. 755–870.

Chemla, E. (2008). An epistemic step for anti-presuppositions. *Journal of Semantics 25*(2), 141–173.

Chomsky, N. (1993). *Lectures on government and binding: The Pisa lectures*. Number 9. Walter de Gruyter.

Collins, C. and P. M. Postal (2014). *Classical NEG raising: An essay on the syntax of negation*, Volume 67. MIT Press.

Collins, J. N. (2019). The polarity of clauses embedded under neg-raising predicates. In *Proceedings of Sinn und Bedeutung 23*, Universitat Autónoma de Barcelona.

Di Tullio, (2018). Verbos estativos y su combinación con el se no argumental. In *Encuentro de Gramática Generativa*, Buenos Aires.

Egre, P. (2009). Question-embedding and factivity. In *MIT Working papers in linguistics*, pp. 79–99.

Gajewski, J. R. (2007). Neg-raising and polarity. *Linguistics and Philosophy 30*(3), 289–328.

Glass, L. (2019). The curious case of the negatively biased Mandarin belief verb "yiwei".

Heim, I. (1991). Artikel und Definitheit. In *Semantics: An International Handbook of Contemporary Research.*, pp. 487–535:. Berlin: de Gruyter.

Holton, R. (2017). I—Facts, Factives, and Contrafactives. *Aristotelian Society Supplementary Volume 91*(1), 245–266.

Horn, L. (1978). Remarks on neg-raising. In P. Cole (Ed.), *Syntax and Semantics 9: Pragmatics*. Academic Press.

Hsiao, P.-Y. K. (2017). On counterfactual attitudes: a case study of Taiwanese Southern Min. *Lingua Sinica*.

Kierstead, G. (2014). The interaction between propositional attitude verbs and reportative evidentiality in tagalog. In *Proceedings of Chicago Linguistic Society 48*.

Križ, M. (2015). *Aspects of homogeneity in the semantics of natural language*. Ph. D. thesis, University of Vienna.

Magri, G. (2009). A theory of individual-level predicates based on blind mandatory scalar implicatures. *Natural Language Semantics*.

Marty, P. (2017). *Implicatures in the DP domain*. Ph. D. thesis, MIT.

Mayr, C. (2017). Predicting polar question embedding. In *Proceedings of Sinn und Bedeutung 21*.

Percus, O. (2006). Antipresuppositions. *Theoretical and empirical studies of reference and anaphora: Toward the establishment of generative grammar as an empirical science 52*.

Quer, J. (2009). Twists of mood: The distribution and interpretation of indicative and subjunctive. *Lingua 119*(12), 1779–1787.

Romoli, J. (2013). A scalar implicature-based approach to neg-raising. *Linguistics and Philosophy 36*(4), 291–353.

Sauerland, U. (2008). Implicated Presuppositions. *Language, Context and Recognition: The Discourse Potential of Underspecified Structures* (June), 581–600.

Schlenker, P. (2012). Maximize Presupposition and Gricean reasoning. *Natural Language Semantics*.

Spector, B. and Y. Sudo (2017). Presupposed ignorance and exhaustification: how scalar implicatures and presuppositions interact. *Linguistics and Philosophy*.

Theiler, N., F. Roelofsen, and M. Aloni (2018). Picky predicates: Why *believe* doesn't like interrogative complements and other puzzles.