Negativity without negation[1]

Lisa HOFMANN — *University of Stuttgart*

Abstract. This paper addresses the anaphoric polarity sensitivity of negativity-tags, challenging the idea that they are only licensed by sentential negation (e.g., Klima 1964; Kramer and Rawlins 2009; Farkas and Bruce 2010; Brasoveanu et al. 2013, 2014; Roelofsen and Farkas 2015), arguing instead that they are sensitive to counterfactual propositional content in discourse. This is supported by data showing that negativity-tags are licensed without overt negation and their acceptability is influenced by contextual factors. The emerging notion of discourse-polarity has theoretical implications: The discourse-effect of negation is tied to its anti-veridical semantics, and characterizing negative antecedents requires a discourse-level representation that integrates information from both semantic representations and pragmatic inferences.

Keywords: negation, anaphora, discourse interpretation.

1. Introduction

Klima (1964) identified a set of negativity-tags: complex anaphoric expressions that require their antecedent to be in some sense negative, like English *neither*-tags, which may have negative antecedents (1a), but not affirmative ones (2a). The same contrast has been observed for (1/2b) agreeing uses of the negative polarity particle (PolP) *no* (Pope 1972), and (c) factive[2] uses of elliptical *'why not'*-questions (Hofmann 2018, 2022; Anand et al. 2021).

(1) **Negative antecendent:**
Sue didn't dance at the party.
a. Neither did Mary.
b. No, she really didn't.
c. but she didn't explain why not.

(2) **Positive antecendent:**
Sue danced at the party.
a. # Neither did Mary.
b. # No, she really did.
c. # but she didn't explain why not.

The expressions in (a–c) are interpreted relative to propositional content in discourse: *neither*-tags (a) have an additive presupposition (Heim 1992), while *no* (b) depends on a previous utterance (treated as propositional anaphora in e.g. Farkas and Bruce 2010; Krifka 2013). *why not* in (c) involves clausal ellipsis (Kramer and Rawlins 2009; Hofmann 2018), often linked to contextual entailment of the elided proposition (e.g. Merchant 2001; Kroll 2019). Since the standard view of propositions as sets of possible worlds doesn't distinguish the polarity of the introducing clause, the question arises of how to explain the sensitivity to polarity.

A related question—what makes a propositional antecedent behave as negative in discourse—arises because the class of licensing contexts is syntactically and semantically diverse. Negative markers (*not*), and expressions of varying negative strength can license negativity-tags, in various syntactic positions (Klima 1964; Brasoveanu et al. 2013, 2014): They include negative

[1]Thanks to Adrian Brasoveanu, Jim McCloskey, Donka Farkas, Amanda Rysling, and Jorge Hankamer, whose teachings inspired and guided the development of the ideas presented here; Jess Law, Hans Kamp, Judith Tonhauser, Tom Roberts, Chris Barker, Morwenna Hoeks, the UCSC S-Circle and the NYU semantics group, as well as audiences and reviewers at SuB and XPrag in September 2022, for helpful discussion and valuable feedback.

[2]Factive uses of *why not* appear in information-seeking questions or veridical contexts. They differ from the free modal reading, illustrated below, found in (zero-answer) rhetorical questions (Hofmann 2018; Anand et al. 2021).
(i) A: *Let's go to the movies!* B: *Why not?* (There is no reason that we wouldn't/shouldn't go to the movies.)

 Negativity without negation. In: Maria Onoeva, Anna Staňková, and Radek Šimík (eds.): Proceedings of Sinn und Bedeutung 27, pp. 309-327 Praha: Charles University.

quantifiers (like *never, no one*), certain downward-entailing quantifiers (*rarely, few people*), and negative proximatives (*hardly, hardly anybody*). In adverbial positions (3a), or argument positions (3b–c), these can license subsequent negativity-tags (4).

(3) a. Pat {never/rarely/hardly} dances.
b. {No one/few people/hardly anybody} dance(s).
c. Pat dislikes {no one/few people/ hardly anybody}.

(4) a. Neither do I.
b. No, they really don't.
c. but I'm not sure why not.

The various quantifiers in (3) have received a unifying analysis as contributing sentential negation and a positive (upward-monotone) quantifier to the representation of their clause. (e.g. Klima 1964; Haegeman and Zanuttini 1991; Penka 2007, though see De Swart 2000). This decompositional analysis is mainly motivated by split-scope readings (e.g. Penka 2007), but also their ability to license negativity-tags (Klima 1964; Brasoveanu et al. 2013, 2014). Accordingly, negativity-tags are often taken to be sensitive to sentential negation. For instance, the feature-based account of PolPs (Farkas and Bruce 2010; Roelofsen and Farkas 2015) assumes that negative antecedents are introduced by clauses involving sentential negation, based on Jackendoff's (1969) semantic characterization. Similarly, ellipsis-based accounts of PolPs (Kramer and Rawlins 2009; Holmberg 2013) rely on syntactic reflexes of negation.

This paper argues that an analysis of discourse-negativity and negativity-tags requires a level of representation that goes beyond explicit clausal content, incorporating information from the literal semantic content and pragmatic inferences. Section 2 presents data showing that negativity-tags can be licensed implicitly, without overt negation. It develops the generalization that negativity-tags are sensitive to negation because they are licensed in a discourse that makes available counterfactual propositional content. Section 3 provides an overview of previous approaches to anaphoric polarity sensitivity, evaluating them in light of the generalization developed in Section 2. While none of these accounts can fully explain it, a synthesis of their key insights leads to the conclusion that the negative antecedent requirement of negativity-tags arises from two factors: (i) an anaphoric expression in an anti-veridical embedding, and (ii) a requirement to interpret the utterance containing the anaphor in conjunction with the utterance containing its antecedent. Section 4 illustrates how these factors manifest in the three negativity tags in (1–2). Section 5 presents an analysis of discourse-negativity—the anaphoric potential of counterfactual content. It is implemented in Intensional CDRT, a dynamic intensional framework with propositional discourse referents (drefs). Section 6 derives the polarity-sensitivity of the anaphoric negativity-tag *why not* as a combination of multiple discourse-links. For reasons of space, I only provide a formal analysis for *why not* in this paper, but I lay the groundwork for extending the analysis to the other negativity-tags discussed. Section 7 concludes the paper.

2. Negativity without negation

This section shows that antecedents under neg-raising or anti-veridical attitudes can license negativity-tags without overt negation. The proposed generalization is that discourse-negative utterances introduce propositional content, which is (possibly implicitly) interpreted as counterfactual (i.e., false according to speaker commitments). We begin by looking to Kroll (2019, 2020), who identifies neg-raising as a context giving rise to polarity-reversing sluices, where the antecedent is an embedded clause without negation, but ellipsis receives a negative paraphrase:

(5) Kroll 2019: (2)
I don't think that California will comply but I don't know why ~~California won't comply~~.

Based on cases like (5), Kroll argues for a pragmatic constraint on ellipsis licensing: If the neg-raising inference is derived pragmatically (e.g. based on Gajewski 2007), the negative interpretation for the ellipsis is based on this pragmatic inference as well. Similarly, embedded clauses under neg-raising (6) also provide antecedents for the negativity-tags in (7).

(6) **Neg-raising antecendent:**
I don't think that Sue danced at the party…

(7) a. …but she didn't say why not
b. No, she really didn't.
c. …and neither did Mary.

The same argument applies here: If neg-raising is derived pragmatically, then so is the licensing of the tags in (7). While there is evidence that neg-raising with *think/believe* could be derived syntactically, (7a–c) may also follow neg-raising in island contexts, like the NP-island in (8), which is uncontroversially derived pragmatically (Collins and Postal 2018).

(8) I don't get the impression that Sue danced at the party…

While *neither*-tags and factive *why not* unambiguously require a negative antecedent, *no* (7b) after (6) or (8) could be characterized as a reversing use (based on Farkas and Bruce 2010). However, the German PolP *doch*, designated for rejecting negative antecedents, is also available in these types of contexts. (9) shows that neg-raising with *glaube* ('believe', A1) and island neg-raising with *Eindruck* ('impression', A2) can be followed by the PolP negativity-tag *doch* (B).

(9) Neg-raising antecedents for *doch*: (German)
A1: Ich glaube nicht, dass die Baustelle vor nächstem Jahr fertig wird.
I believe not that the construction site before next year finished gets
'I don't believe that the construction will be finished before next year.'
A2: Ich habe nicht den Eindruck, dass die Baustelle vor nächstem Jahr fertig wird.
I have not the impression that the construction site before next year finished gets
'I don't get the impression that the construction will be finished before next year.'
B: Doch, das wird sie bestimmt.
DOCH that will it certainly
'Yes, it certainly will.'

In (9), the negativity-tag *doch* is interpreted in relation to the embedded clause without negation. Licensing by neg-raising thus present a strong case that the relevant notion of polarity should be characterized on the level of discourse, not clausal representations. The main empirical contribution of this paper is the discourse-based generalization stated in (10).

(10) **Discourse-Negativity:**
An utterance patterns as negative for the purposes of negativity-tags, if it explicitly introduces propositional content into the discourse, which is interpreted as counterfactual (based on literal content or pragmatic inference.)

(10) states that the propositional content anaphorically picked up by negativity-tags needs to be introduced into the discourse explicitly, whereas the information about the speaker's epistemic

stance towards that content, i.e. whether it is interpreted as (counter-/non-)factual, may be implicitly inferred in discourse. The requirement of explicit introduction captures that entailment is insufficient to license negativity tags. This is illustrated for the tags in (12), which are not available following (11), even though the sentence entails the negative paraphrase *Susan didn't pass the exam*, modulo satisfaction of presuppositions (Hofmann 2019b).

(11) Sue failed the exam.
(↔ Susan didn't pass the exam.)

(12) a. # …but I'm not sure why not
b. # No, she really didn't.
c. # …and neither did Mary.

While the propositional content itself has to be introduced explicitly, the generalization in (10) further predicts that overt non-negative clauses embedded in semantically anti-veridical (AV) attitude contexts (13) can introduce negativity, and are acceptable with the tags in (7).

(13) a. You were wrong…
b. It is a lie…
c. It was just a rumor…
d. It is false…
…that Sue danced at the party.

While judgments for (7) following (13) vary, experimental data from a forced-choice continuation task in Hofmann (2022) provides evidence that *why not* can be used following AV-attitudes, as well as neg-raising antecedents with *think/believe* or NP-islands.

Further supporting evidence for the claim that the implicit counterfactual interpretation of propositional content in dicourse licenses negativity-tags comes from the availability of negativity tags following sarcastic utterances (14), or when accommodating a negative answer to a polar question (15), illustrated there for *why not* and *neither*-tags.

(14) A (sarcastic): I will totally be able to afford buying a house by 35 in this economy.
a. B: Yeah, me neither.
b. B: I can see why not.

(15) Do you think this was a good idea?
a. Why or why not?
b. Well, neither do I.

3. Approaches to anaphoric polarity-sensitivity

This section surveys approaches to polarity-sensitive propositional anaphora, evaluates their applicability to implicitly negative antecedents, and highlights the insights that are integrated into the proposed analysis. I mainly focus on the literature on PolPs, as they have been explored in the greatest formal depth, discussing the feature-based analysis (Farkas and Bruce 2010; Roelofsen and Farkas 2015), the saliency account (Krifka 2013; Claus et al. 2017), and briefly touch on ellipsis-based accounts of PolPs (Kramer and Rawlins 2009; Holmberg 2013). Another formally explicit approach to the interaction of anaphora to propositional content with polarity involves local contextual entailment (Kroll 2019, 2020; Hofmann 2022), discussed here in some detail for Kroll's (2019) analysis of licensing clausal ellipsis.

3.1. Feature-based account of PolPs

The feature-based account of PolPs (Farkas and Bruce 2010; Roelofsen and Farkas 2015; Farkas and Roelofsen 2019) builds on Pope's (1972) typological study of PolP answering systems,

which identified two dimensions of information signalled by PolPs across languages: (i) the response being positive or negative, and (ii) the response agreeing or disagreeing with the antecedent. The account explains Pope's generalization by proposing that PolPs across languages may morphologically realize *absolute* or *relative* polarity features, or a combination of both:

(16) Presuppositions of polarity features (simplified, after Roelofsen and Farkas 2015)
   a. Absolute features:
      (i) [+]: The response has positive polarity.
      (ii) [−]: The response has negative polarity.
   b. Relative features:
      (i) [AGREE]: The response and the antecedent have the same polarity and they are semantically equivalent.
      (ii) [REVERSE]: The response and the antecedent have opposite polarity and they are contradictory opposites.

The typology of answering systems is governed by morphological feature realization potentials, which specify which features PolPs may realize. For English *yes* and *no*, these are as in (17).

(17) Feature realization potential of English PolPs (Roelofsen and Farkas 2015)
   a. [AGREE] and [+] can be realized by *yes*
   b. [REVERSE] and [−] can be realized by *no*

The account assumes that PolPs can express one or both of the associated features and can be interpreted in combination with another feature not explicitly realized by the particle based on markedness considerations. Thus, agreeing uses of *no* are predicted to be negative, and matching the polarity of the antecedent, which in turn predicts their negativity-tag behavior.

This analysis offers core insights into the pragmatics and typology of answering systems by parametrizing PolPs based on the polarity of the response and their relationship with the antecedent utterance. However, cases of negativity without explicit negation raise questions about the representational assumptions and proposed semantics of polarity features. Roelofsen and Farkas (2015) associate negative polarity features with propositional discourse referents (drefs) introduced by semantically negative clauses, where contradictory negation is the highest-scoping operator in the semantic representation of the clause (based on Jackendoff 1969). However, this generalization doesn't cover the cases discussed in Section 2, where counterfactual propositions expressed affirmatively act as negative antecedents in discourse. To account for these cases while retaining the insights from the feature-based account, we redefine the notion of (discourse) polarity in terms of counterfactual content rather than sentential negation, while adopting the idea that PolPs and anaphoric negativity-tags impose a dual constraint: They convey information about the (absolute) polarity of the anaphoric utterance and (relative) polarity, i.e., the relationship between the utterances containing the anaphor and its antecedent.

3.2. PolPs as propositional anaphora

The saliency account of PolPs (Krifka 2013; Claus et al. 2017) explains the interpretation of PolPs by analogizing them to pronominal propostional anaphora. It assumes that negative utterances introduce two propositional drefs: one for the matrix clause and another for the negated content (see also Stone 1999; Murray 2014; Snider 2017). This is motivated by the independent

observation that pronominal propositional anaphora can pick up either of these drefs:

(18) [Ede didn't [~~Ede~~ steal the cookie]$^{\phi_2}$]$^{\phi_1}$ … Adapted from Krifka 2013: (24)
a. …and he can actually prove it$_{\phi_1}$. it$_{\phi_1}$: that Ede didn't steal the cookie
b. …even though people believed it$_{\phi_2}$ it$_{\phi_2}$: that Ede stole the cookie

The account posits that *yes* and *no* are propositional anaphora, where *yes* asserts its antecedent, while *no* asserts the negation of its antecedent. The patterns for English PolPs are analyzed by assuming that they may be anaphoric to either the matrix content ϕ_1 or the negated content ϕ_2.

(19) A: [Ede didn't [~~Ede~~ steal the cookie]$^{\phi_2}$]$^{\phi_1}$

a. B: Yes$_{\phi_1}$, he didn't. (agreement)
b. B: Yes$_{\phi_2}$, he did. (disagreement)
c. B: No$_{\phi_2}$, he didn't. (agreement)
d. B: No$_{\phi_1}$, he did. (disagreement)

No negates its antecedent, so when it refers to the matrix content (ϕ_1), it entails disagreement. Therefore, agreeing uses of *no* always point to a proposition in the scope of negation (ϕ_2). This explains why such uses are restricted to negative antecedents. These representational assumptions readily apply to our data, as long as we extend the class of negative antecedents to encompass counterfactual propositions more broadly, not just those in the scope of negation.

The saliency account further suggests that ambiguity resolution with bare PolPs, and production choices when two distinct PolPs can express the same answering relation are modulated by salience, assuming that matrix and embedded propositions have different levels of salience. However, the account lacks an explicit operationalization of propositional salience, and experiments in Claus et al. (2017) found no effect of contextually modulated salience on the acceptability of PolPs. Furthermore, Farkas and Roelofsen (2019) argue that the salience-based constraints cannot explain the typology of answering systems, particularly the behavior of negative PolPs in languages where they cannot affirm negative utterances. Additional assumptions would be required to exclude the possibility of a counterfactual dref as a potential antecedent, specifically in these languages and to the exclusion of other propositional anaphora.

I adopt the representational assumptions of this account: negative antecedent utterances are ones introducing a certain kind of embedded propositional content into the discourse. Based on the generalization from Section 2, this extends to counterfactual content beyond just content introduced under negation. On the anaphoric side, the saliency account posits that negative PolPs involve anaphoric expressions under negation. We adopt this assumption for negativity-tags more broadly—they involve anaphora in anti-veridical contexts. However, given the typological limitations of the saliency account, I contend that the constraints on PolPs can be better understood by parametrization based on absolute/relative polarity.

3.3. Ellipsis-based accounts of PolPs

According to ellipsis-based accounts of PolPs (e.g. Kramer and Rawlins 2009; Holmberg 2013), PolPs connect to their antecedent by appering in utterances with clausal ellipsis, and their sensitivity to antecedent polarity arises from a negative syntactic dependency (e.g., based on Zeijlstra 2004). This dependency links polarity features associated with the PolP to syntactic expressions of polarity in the ellipsis site. Since this type of account relies on a morphosyntactic realization of negation, it cannot readily be extended to cases of negativity without explicit negation.

3.4. Local contextual entailment

The analysis of polarity-reversing sluices in Kroll (2019, 2020) explains the interaction between clausal ellipsis and negation at the discourse level, through a pragmatic licensing condition of local contextual entailment (20).

(20) **Local givenness**: (simplified version, adapted from Kroll 2020: 62)
A clause α can be deleted only if α expresses a proposition p, such that p is contextually entailed in the local context of α.

Kroll uses (20) to explain the interpretation of polarity-reversing sluices with neg-raising antecedents in (5), repeated here.

(5) Kroll 2019: (2)
I don't think that California will comply but I don't know why ~~California won't comply.~~

The polarity-reversing interpretation is derived in a dynamic semantic framework, by interpreting the ellipsis in a context that results from an update with the first conjunct, which contextually entails that *California won't comply*. This inference is derived by two assumptions: (i) the neg-raising inference arises from an excluded-middle presupposition associated with neg-raising predicates (Gajewski 2007), and (ii) a speaker's self-ascription of belief leads to discourse commitment. The analysis is fomulated in a version of update semantics (Heim 1983), with relevant definitions given in (21). The notation below differs slightly from Kroll (2019, 2020).

(21) a. Discourse update (assertion): $c + p$
  (i) If c entails the presuppositions of p (i.e. $c \subseteq ps(p)$), then $c + p = c \cap p$.
  (ii) If c does not entail the presuppositions of p, then $c + p$ is undefined.
  (iii) If $c + p$ is undefined and c does not contradict the presuppositions of p, presuppositions may be accommodated:
  $c + p = (c + ps(p)) + p$
 b. Negation: $c + \neg p = c \setminus (c + p)$
 c. Dynamic conjunction: $c + p + q = (c + p) + q$

The interpretation of (5) is based on a standard treatment of *think* as a doxastic attitude in a Hintikkan semantics for propositional attitudes (based on Heim 1992).

(22) *think* as a doxastic attitude:
 a. $[\![think(x_e)(p_{wt})(w_w)]\!]^{M,g} = [\![\textsc{Dox}_x(w) \subseteq p]\!]^{M,g}$, where:
 b. $[\![\textsc{Dox}_x(w)]\!]^{M,g} = \{w' \in D_w : w'$ conforms to what $[\![x]\!]^{M,g}$ believes in $[\![w]\!]^{M,g}\}$

Besides its assertive content, the neg-raising verb *think* contributes an excluded-middle presupposition (Gajewski 2007), which we assume is accommodated (though Kroll does not explicitly make this assumption). Further, speaker self-ascription of belief leading to pragmatic assertion is captured by adding the embedded content to the superordinate context. Here, we state this discourse-effect as the presupposition in (23b-ii). Based on these assumptions, the content in the scope of negation in the first conjunct in (5) is stated as (23).

(23) a. Asserted content: $\lambda w'.think_{w'}(\mathbf{sp})(\lambda w.comply_w(cal))$
 b. Presuppositions:
  (i) Excluded middle:
  $\lambda w'.[think_{w'}(\mathbf{sp})(\lambda w.comply_w(cal) \vee think_{w'}(\mathbf{sp})(\lambda w.\neg comply_w(cal))]$

(ii) Self-ascription of belief: $\lambda w.\forall p[think_w(\mathbf{sp})(p) \rightarrow p(w)]$

The update of the first conjunct in (5) is derived as follows:

(24)
a. Starting context (by assumption): $c = W$
b. Negation: $c + (5) = c \setminus c'$, where $c' = (c + (23))$
c. Presupposition failure: $c' = (c + (23))$ is undefined, because $c \not\subseteq ps(23)$.
d. Accommodation: For $c \not\subseteq \neg ps(23)$, we get (global) accommodation:[3]
$c + (5) = c'' + \neg$(23a), where:
$c'' = (c + ps(23)) = (c + (23\text{bi}) + (23\text{bii}))$
e. Negation: $c'' + \neg(23\text{a}) = c'' \setminus c'''$, where $c''' = (c'' + (23\text{a}))$

Due to global accommodation of the presuppositions and negation of the assertive content in (23), the resulting context entails $\lambda w.\neg comply_w(cal)$. Kroll suggests that this functions as the local context for the ellipsis in the second conjunct. Thus, the clause may be elided when interpreted as expressing this proposition, provided that syntactic isomorphism is also satisfied.

The analysis uses local contextual entailment at a discourse-level representation, which can be enriched by pragmatic inferences like neg-raising. This suggests that clausal ellipsis behaves similarly to other types of anaphora, as local contextual entailment within dynamic semantics has also been proposed as a condition on interpreting of pronominal anaphora, both for individuals (Stone 1999; Brasoveanu 2006; Hofmann 2019a, 2022), and propositions (Hofmann 2022). Extending a discourse-based approach with local contextual entailment to negativity-tags will allow us to include implicitly negative antecedents in the class of discourse-negative utterances.

Further, when treating negativity-tags as involving anaphora in non-veridical contexts, a condition of local contextual entailment can help us understand why they need counterfactual propositional antecedents. However, for ellipsis in local contexts under negation, like in factive *why-not*-interrogatives, Kroll's condition is slightly too strong. The example of clausal ellipsis under *not* in (25b) suggests that local contextual entailment should be considered a condition on ellipsis interpretation rather than licensing.

(25)
a. Mary didn't dance,
b. but she didn't say why not ~~[Mary danced]~~
c. not ~~[Mary danced]~~

The *why*-interrogative in (25b) is factive (Fitzpatrick 2005), and presupposes the truth of its prejacent (i.e. its propositional argument), given in (25c). We can therefore assume that the local context for its prejacent is the global context after updating with (25a), and illustrate the problem for an update with (25a) + (25c).

(26)
a. $c + (25\text{a}) = c \setminus (c + (\lambda w.dance_w(mary))) = c'$
b. $c' + (25\text{c}) = c' \setminus (c' + $ ~~$[Mary\ danced]$~~$)$

The context c', created by updating with the first conjunct (26a), contains only worlds where *Mary didn't dance*. Since the semantics of negation in (24e) assumes that the superordinate context is passed down as the local context under negation, the ellipsis is interpreted in c'. Because c' does not entail the elided proposition *Mary danced* (but its opposite), ellipsis is not predicted to be possible with the attested interpretation.

[3] Heim (1983) suggests that global accommodation, if possible, is strongly preferred, and this assumption will be sufficient for our current purposes.

To salvage this, we could assume that the local context under negation doesn't have to the superordinate context. Instead, we could consider it to be the universe containing all possible worlds. This way, the content of the elided presupposition could be consistently locally accommodated, but it would still not be previously entailed. Another option would be adopting a Lewisian counterfactual semantics for counterfactual local contexts, i.e. picking out a set of worlds that is like the global context, except that the locally expressed proposition is true (e.g. Heim 1992; Stone 1999). But using this as a way of recovering interpretations for ellipsis creates circularity in the explanation. We would have to claim that the proposition is entailed in the local context because it is interpreted in the context, and the ellipsis is interpreted as expressing the proposition because it is entailed in the context. In either case, the previous semantic and pragmatic context, without the elided content, does not already entail the elided proposition.

3.5. Interim conclusion

A discourse-negative utterance overtly expresses a proposition, which is (possibly implicitly) interpreted as counterfactual. This characterization builds on Krifka's observation that negated content is available for anaphora, the generalization in Section 2, and Kroll's dynamic account of how propositional content in discourse can be pragmatically construed as counterfactual.

Negativity-tags can be seen as having a negative absolute polarity. When defining discourse-polarity in terms of counterfactual content, this means they involve anaphoric elements in a anti-veridical context. Interpreting the anaphor locally in such a context is one aspect of explaining their negative antecedent requirement. The second aspect of the explanation can be seen as a reflection of relative polarity, requiring that the utterances containing the anaphor and its antecedent convey non-contradictory matrix propositions. The following section will elaborate on this aspect, suggesting that negativity-tags are complex anaphoric expressions involving the the interplay of multiple discourse-dependencies and negation.

4. A recipe for negativity-tags: anti-veridical anaphora and non-contradiction

Combining insights from the generalization in Section 2 and previous approaches to polarity-sensitive propositional anaphora (Section 3), we understand negativity-tags as complex anaphoric expressions. They require a negative antecedent due to the interplay of an anaphoric element in an anti-veridical context and a second discourse-requirement that the utterances containing the negativity-tag and its antecedent are non-contradictory. This section explores the implications of this characterization for *why not*, *neither*-tags, and agreeing uses of negative PolPs.

The requirement for the antecedent of *'why not ~~p~~'* to be negative arises from a combination of: (i) clausal ellipsis under negation, requiring an antecedent for the ellipsis of the clause expressing the proposition *p*, and (ii) the factivity associated with information-seeking uses of *why*, presupposing $\neg p$ (Fitzpatrick 2005). Together, these two components require that *p* has been introduced counterfactually in the discourse, whereas individually, they do not impose the same requirement. For example, a factive use of a negative *why*-question without ellipsis (27) does not require a negative antecedent.

(27) A: Mary sat in her chair all night at the party.
B: Why didn't Mary dance?

(27) shows that the factive presupposition of negative *why*-questions can be met by an affirma-

tive utterance that entails the prejacent of *why*. Likewise, negative polarity ellipsis (i.e., clausal ellipsis of the complement of a high polarity head, e.g., Kramer and Rawlins 2009; Gribanova 2017; McCloskey 2017) in non-factive contexts does not require a negative antecedent (27).

(28) Maybe Mary danced at the party
  a. …but maybe/probably not. (maybe/probably she didn't dance)
  b. If not, it must have been a boring party. (if Mary didn't dance)

Although we focus here on analyzing *why not*, we can test the predictions for other negativity-tags by identifying similar conspiring discourse-requirements. In the case of negative additive tags, there is verb-phrase ellipsis (VPE) under negation, which is licensed by an overt verb-phrase in the discourse, while the negative additive presupposes that the proposition expressed by that VP is false (i.e. interpreted counterfactually). In combination, we get a negative-antecedent requirement, not only for *neither*-tags, as Klima (1964) observed, but also for *not either* + VPE, i.e. both utterances in (30) are acceptable after (29a), but not (b).

(29) a. Mary didn't pass the test.
  b. Mary failed the test.

(30) a. Neither did Sue ~~pass the test~~.
  b. Sue didn't either ~~pass the test~~.

Either one of these discourse-links on its own does not lead to a negative antecedent requiement, as illustrated for VPE under negation in (31), and negative additives in (32).

(31) Mary danced,
but Agatha didn't.

(32) Mary aced syntax.
  a. She didn't fail any of her other classes either.
  b. Neither did she fail any of her other classes.

Agreeing uses of *no* can be viewed similarly: Following Krifka's (2013) assumption that negative PolPs negate the antecedent proposition, they can be seen as propositional anaphors in anti-veridical contexts. The second discourse-link comes from the agreeing interpretation: If the PolP were to negate a (veridical) matrix proposition, the response couldn't be interpreted as affirming the previous utterance. In case of designated particles for rejecting negative utterances (like German *doch*), we can follow Krifka's assumption that they are positive PolPs, asserting their antecedent, and carry an additional presupposition forcing their antecedent to be a counterfactual proposition (p 15). While Krifka simply states that *doch* should pick up a negated proposition, we might propose that the presupposition of *doch* demands that the response is interpreted as disagreeing with the previous utterance. This aligns with our understanding that polarity-sensitivity arises as a reflection of specifying the relationship with the prior utterance.

Finally, if the proposed analysis of negativity-tags is right, we expect pronominal anti-veridical anaphora in a non-contradictory discourse relations to exhibit the a negative antecedent requirement as well. This is borne out, illustrated in (33): In an anti-veridical context, the pronoun *that* may have a counterfactual antecedent (33a), but not a veridical one (33b). That applies, when they are subsequent assertions by the same speakers, which requires the two utterances to be non-contradictory. In case of utterances by different speakers (33c), an anti-veridical anaphor can pick up the matrix assertion, as this discourse allows for an interpretation where the two utterances contradict each other.

(33) a. Mary didn't dance at the party. That's a lie. (✓*that:* that Mary danced)
  b. #Mary danced at the party. That's a lie. (✗ *that:* that Mary danced)
  c. A: Mary danced at the party. B: That's a lie. (✓*that:* that Mary danced)

The anaphoric polarity-sensitvity of negativity-tags presents itself as a complex phenomenon, which has many researchers led to assume anaphoric mechanisms that make reference to clausal negation. However, the argument presented here suggests that it comes about as a combination of regular kinds of anaphoric dependencies that are otherwise well-behaved, like propositional anaphora and ellipsis, and their interaction with their semantic and pragmatic context. The central contrasts can be explained by appealing to a generalization that anaphoric expressions are interpreted in their local context, and the basic principle of discourse consistency.

5. Analysis of discourse-negative utterances

While negativity-tags generally involve anaphoric elements in anti-veridical contexts, Section 4 showed that the nature of these elements varies between tags. For instance, ellipsis in *why not* interacts with the discourse in a way that can be modeled using Heimian local contexts, while PolPs and pronominal anaphora can be seen as picking up propositional drefs. To address this diversity, Section 5.1 sets up the analysis of discourse-negativity in a dynamic system where propositional operators systematically provide Heimian local contexts for interpreting their prejacent and also introduce propostional drefs (based on Stone 1999; Brasoveanu 2010; Snider 2017; Hofmann 2019a, 2022). The analysis of semantically counterfactual content is then extended to implicitly negative propositions under neg-raising in Section 5.2. Although Section 6 focuses on deriving the anaphoric requirements of *why not*, this framework lays the foundation for extending it to other negativity-tags in future research.

5.1. Counterfactual content in a propositional logic of change

5.1.1. Updates

I implement the analysis in a version of intensional CDRT (based on Muskens 1996; Brasoveanu 2006; Hofmann 2022), where propositional operators introduce drefs for the set of worlds in which their prejacent is true, and their truth-functional meaning is expressed in terms of relations over these drefs (Stone 1999; Brasoveanu 2010; Krifka 2013; Snider 2017). Discourse states store the information about propositional drefs and relations between them, allowing us to keep track of speaker commitments about drefs, and capture the generalization about anaphora and propositional antecedents in anti-veridical contexts.

The basic types are: t (truth-values), e (entities), w (worlds), s (variable assignments). A propositional dref ϕ is a function from assignments i_s to propositions p_{wt}. Utterances are interpreted as discourse updates, which are interpreted within an underlyingly static logic as relations of type $s(st)$ between input and output states i_s, j_s, where discourse states are variable assignments. Updates are represented as in (34c). They contain a list of new drefs, introduced by variable update (34a), and properties of the output state, which are imposed as output conditions (34b).

(34) a. Variable update:
$i[\phi]j$ is true iff i and j differ at most wrt the values assigned to the variable ϕ

b. Output conditions: $Dance_{\phi}\{Mary\} := \lambda i_s.\forall w \in \phi(i).dance_w(mary)$

c. Updates: $[\phi_1,\ldots,\phi_n \mid C_1,\ldots,C_n] := \lambda i_s.\lambda j_s.i[\phi_1,\ldots,\phi_n]j \wedge C_1(j) \wedge \ldots \wedge C_n(j)$

Variable update (34a) is defined as random assignment of values to a variable, following Groenendijk and Stokhof (1991). In (34b), the dynamic predicate *Dance* takes two arguments: an individual discourse constant (*Mary*) and a propositional dref (ϕ), indicated as a subscript on the

predicate. In this propositional logic of change, we simplify by using only discourse constants (type se) associated with proper names and no individual variables.[4] Individual constants point to the same entity across assignments (e.g. $Mary_{se} := \lambda i.mary_e$). (34b) holds of some i_s, iff the corresponding static predicate (in lowercase) holds of $mary_e$ in each world in $\phi(i)$. Stone (1999) introduced this mechanism of point-wise checking across all worlds of evaluation, proposing that lexical meanings encapsulate universal quantification over possible worlds in their local context. Dynamic concjunction is defined in the usual way, as relation composition:

(35) Dynamic conjunction: $D_{s(st)};D'_{s(st)} := \lambda i_s.\lambda j_s.\exists h_s.[D(i)(h)](D'(h)(j))$

5.1.2. Assertion

To embed a version of Heimian (propositional) update in a system with propositional drefs, we need propositional drefs representing our global and local contexts. For the global context, we assume a indexical propositional dref ϕ_{DC_S} pointing to the discourse-commitment set of the speaker S (Gunlogson 2004).

Assertion is modeled as intersective update of the global context (Stalnaker 1978; Heim 1983) by imposing an output condition that ϕ_{DC_S} entails the asserted proposition at the output. Here, we attribute this function to a declarative sentential mood operator that combines with a propositional prejacent, following Bittner (2009); Murray (2014). The prejacent of a propositional operator is represented as a 'dynamic proposition' (type $s(wt),s(st)$), which takes a propositional variable as an argument to return an update (e.g. $\lambda\phi.[Dance_\phi\{Mary\}]$). The declarative operator provides this argument as a propositional dref, and imposes the condition that the proposition is entailed by speaker commitments. (36) illustrates the assertion of *Mary danced.*

(36) a. S: $\text{DEC}_S(Mary\ danced) \rightsquigarrow$ $[\phi \mid \phi_{DC_S} \Subset \phi];[Dance_\phi\{Mary\}]$
b. Dynamic inclusion: $\phi_1 \Subset \phi_2 := \lambda i_s.\phi_1(i) \subseteq \phi_2(i)$
c. Indexical dref for speaker commitments: $\phi_{DC_S} := \lambda i.DC_S$

In (36a), the condition $Dance_\phi\{Mary\}$ ensures that the output state j can only be one where the values for $\phi(j)$ are sets of worlds such that *Mary danced* in each world. The condition $\phi_{DC_S} \Subset \phi$ indexically invokes the commitment set of the speaker S, and states that $\phi(j)$ is entailed by S's discourse commitments at j. This is defined in terms of dynamic set inclusion (36b). As a result, all ϕ_{DC_S}-worlds are ϕ-worlds, and S is committed that *Mary danced*. This means that ϕ is a *veridical propositional dref* relative to S.

5.1.3. Negation

For negation, the assumed propositional relation is one of non-overlap (contrary negation). The discourse-effect of a negative utterance is captured in (37).

(37) Mary didn't dance $\rightsquigarrow$ $[\phi_1 \mid \phi_{DC_S} \Subset \phi_1];[\phi_2 \mid \phi_1 \Cap \phi_2 = \varnothing];[Dance_{\phi_2}\{Mary\}]$

An update with (37) introduces two propositional drefs ϕ_1 and ϕ_2. The condition $Dance_{\phi_2}\{Mary\}$ restricts ϕ_2 to worlds where *Mary danced*. The condition $\phi_1 \Cap \phi_2 = \phi_\varnothing$ states that ϕ_1 and ϕ_2 have no overlap, and is defined in terms of dynamic intersection over drefs, while invoking a discourse constant for the empty set, as defined in (38).

[4] See Hofmann (2019b, 2022) for details on how the presented system can incorporate individual quantification.

(38) a. Dynamic intersection: $\phi_1 \Cap \phi_2 := \lambda i_s.\phi_1(i) \cap \phi_2(i)$
b. Discourse constant for empty set: $\phi_\varnothing := \lambda i.\varnothing$

We implicitly assume maximization over propositional drefs throughout (discussed in detail in Hofmann 2022, pp. 155–160), based on Brasoveanu (2006). Maximization takes scope over the sentence radical of a clause, as illustrated in (39), so ϕ_2 is the maximal proposition where *Mary danced*, and ϕ_1 is the maximal proposition where she did not.

(39) Mary didn't dance $\rightsquigarrow$
$[\phi_1 \mid \phi_{DC_S} \Subset \phi_1]; \mathbf{max}_{\phi_1}([\phi_2 \mid \phi_1 \Cap \phi_2 = \varnothing]; \mathbf{max}_{\phi_2}([Dance_{\phi_2}\{Mary\}]))$

This results in a complementation relation, implementing the standard boolean truth-functional meaning of negation (contradictory negation). Accordingly, ϕ_1 points to the set of worlds in which *Mary didn't dance*. By assertion, all worlds in ϕ_{DC_S} are ϕ_1-worlds. A crucial consequence of this it that no world in ϕ_{DC_S} will be a ϕ_2-world, rendering ϕ_2 a *counterfactual propositional dref* for S, i.e. one referring to a proposition that S is committed to being false. For the rest of this paper, I will leave the maximization assumed for propositional drefs implicit. Therefore, we can understand (37) as an abbreviation of (39).

5.1.4. AV-attitudes

Like negation, anti-veridical attitude introduces a counterfactual dref for the embedded content, here exemplified for the AV-predicate *be wrong* (40).

(40) Sue was wrong that Mary danced $\rightsquigarrow$
$[\phi_1 \mid \phi_{DC_S} \Subset \phi_1]; [\phi_2 \mid Wrong_{\phi_1}\{Sue, \phi_2\}]; [Dance_{\phi_2}\{Mary\}]$

Following Anand and Hacquard (2014), I treat *be wrong* as an anti-veridical assertive attitude:[5]

(41) a. $[\![wrong_w(x_e)(p_{wt})]\!]^{M,g} = [\![\textsc{Assertive}_x \subseteq p \wedge \neg p(w)]\!]^{M,g}$

The condition $Wrong_{\phi_1}\{Sue, \phi_2\}$ holds of an output state j iff for each world w in $\phi_1(i)$, Sue asserted $\phi_2(i)$, and $\phi_2(i)(w) = 0$. Accordingly $\phi_2(i)$ is false in each ϕ_1-world, which commits the speaker that ϕ_2 is false. Here, the propositional dref ϕ_2 for the embedded content (*Mary danced*) is counterfactual due to the relation introduced by *be wrong*.

5.2. Neg-raising and accommodation

The proposed system embeds Heimian-style intensional contexts in a referential context by using propositional drefs. This allows us to analyze neg-raising antecedents, by adapting Kroll's (2019; 2020) approach (Section 3.3) to this system. In the following, I show how a counterfactual dref can be introduced under neg-raising, and illustrate how discourse inferences can enrich our discourse-semantic representations. Importantly, propositional drefs for embedded content are introduced explicitly (c.f. Heim's, 1982 formal link condition), while information about previously introduced drefs can be conveyed implicitly. Neg-raising utterances provide a dref

[5] We may assume the following definition of an assertive information-state:
(i) $[\![\textsc{Assertive}_x(w)]\!]^{M,g} = \{w' \in D_w : w'$ conforms to the (communicative) commitments of some assertion made by $[\![x]\!]^{M,g}$ in $[\![w]\!]^{M,g}\}$
This is a simplification. E.g., Anand and Hacquard (2014) suggest the relevant information state is a projected common ground (in the sense of Farkas and Bruce 2010) associated with a reported communicative event.

for the embedded clause, which would be non-veridical (but not counterfactual), if interpreted based on the semantic at-issue content alone, as illustrated in (42).

(42) I_S don't think that Mary danced $\rightsquigarrow$
$[\phi_1 \mid \phi_{DC_S} \Subset \phi_1]; [\phi_2 \mid \phi_1 \Cap \phi_2 = \phi_\varnothing]; [Think_{\phi_2}\{S, \phi_3\}]; [Dance_{\phi_3}\{Mary\}]$

(42) introduces a dref ϕ_3 for a set of worlds where *Mary danced*, and a dref ϕ_2 for a set of worlds where *S thinks* ϕ_3. Based on (42), ϕ_2 can contain worlds where ϕ_3 is true and worlds where it is false. As a result, the same applies to its negation, ϕ_1, and the assertion does not commit the speaker to the truth or falsity of ϕ_3. This makes ϕ_3 a *hypothetical dref* (but not a counterfactual one), based on the semantic content alone. Speaker commitment to ϕ_3 being false is derived pragmatically, by accommodating an excluded middle presupposition associated with uses of *think*, and self-ascription of belief leading to discourse commitment. This paper does not offer an account of projection or the contextual requirements of presuppositions. It merely illustrates how the inferences can be stated and accommodated as presuppositions within the proposed system: The expressions in (42) are interpreted as conditions on a discourse state i, in relation a propositional variable ϕ provided in discourse.

(43) Presuppositions
a. Excluded middle: $\langle \phi(i) \subseteq \lambda w.[think_w(S(i))(\phi_3(i)) \vee \neg think_w(S(i))(\phi_3(i))] \rangle_{i,\phi}$
b. Self-ascription of belief: $\langle \phi(i) \subseteq \lambda w.[think_w(S(i))(\phi_3(i)) \rightarrow \phi_3(i)(w)] \rangle_{i,\phi}$

While the contents in (43) are not dynamically active like explicit content, and cannot introduce new drefs, they can provide information about existing ones. Thus, the content α of $\langle \alpha \rangle_{i,\phi}$ is stated as a condition in the underlying static logic. Recall that we assume accommodation in the global context, so the propositional variable ϕ is interpreted as pointing to ϕ_{DC_S}:

(44) I_S don't think that Mary danced $\rightsquigarrow$
$[\phi_1 \mid \phi_{DC_S} \Subset \phi_1]; [\phi_2 \mid \phi_1 \Cap \phi_2 = \phi_\varnothing];$
$\langle \phi_{DC_S}(i) \subseteq \lambda w.[think_w(S(i))(\phi_3(i)) \vee \neg think_w(S(i))(\phi_3(i))] \rangle_i;$
$\langle \phi_{DC_S}(i) \subseteq \lambda w.[think_w(S(i))(\phi_3(i)) \rightarrow \phi_3(i)(w)] \rangle_i; [Think_{\phi_2}\{S, \phi_3\}]; [Dance_{\phi_3}\{Mary\}]$

Setting aside questions of presupposition satisfaction or failure here, we state an accomodation update which states a discourse-condition on the input assignment:

(45) $\langle \alpha \rangle_i := \lambda i.\lambda j.[i = j \wedge \alpha]$

Assuming that the contents in (43) are accommodated globally, and combining this with the assertion of (42), the output j will be s.t. $\phi_{DC_S}(j) \subseteq \lambda w.[\neg \phi_3(i)(w)]$. This means that all worlds $w \in \phi_{DC_S}(j)$ are non-ϕ_3-worlds, and ϕ_3 is a counterfactual dref for S. While this does not provide an account of presupposition projection or interpretation, it serves as a demonstration of how presupposition accommodation can lead to pragmatic enrichment in the proposed system.

5.3. Interim conclusion

We have characterized discourse-negative utterances as ones that introduce a counterfactual propositional dref, and presented a formal analysis within a version of intensional CDRT. This dynamic system provides a level of representation that is based on semantic content, but amenable to pragmatic enrichment due to discourse inferences. It can address utterances with sentential negation, but also other kinds of counterfactual content under AV-attitudes and neg-raising.

6. *why not* — licensing clausal ellipsis under negation

To derive that the negativity-tag *why not* is sensitive to the presence of a counterfactual proposition in the discourse, I give an analysis of the contrast in (46).

(46) a. Mary didn't dance. Sue explained why not.
b. Mary danced. # Sue explained why not.

6.1. *why*-questions

We assume that the unavailability of (46b) results from a combination of clausal ellipsis under negation and the factivity associated with *why*-questions (as discussed in Section 4). To develop the formal analysis, let us first make explicit our assumptions about a non-elliptical case in (47).

(47) a. S: Mary didn't dance. $\rightsquigarrow$ $[\phi_1 \mid \phi_{DC_S} \Subset \phi_1]; [\phi_2 \mid \phi_1 \Cap \phi_2 = \phi_\varnothing]; [Dance_{\phi_2}\{Mary\}]$
b. S: Sue explained why Mary didn't dance $\rightsquigarrow$
$[\phi_3 \mid \phi_{DC_S} \Subset \phi_3]; [\phi_4 \mid Explain_{\phi_3}\{Sue, \phi_4\}]; [\phi_5, \phi_6 \mid Because_{\phi_4}\{\phi_5, \phi_6\}];$
$\langle \phi_{DC_S}(i) \subseteq \phi_5(i) \rangle_i; [\phi_7 \mid \phi_5 \Cap \phi_7 = \phi_\varnothing]; [Dance_{\phi_7}\{Mary\}]$

For simplicity, we treat *explain* as a basic assertive attitude, glossing over some aspects of the semantics of the predicate that are not directly relevant here. Accordingly, the condition $Explain_{\phi_3}\{Sue, \phi_4\}$ states that the ϕ_4-worlds are the ones consistent with what Mary explained in ϕ_3. The interpretation of *why*-interrogatives invokes the condition $Because_{\phi_4}\{\phi_5, \phi_6\}$ states that ϕ_6 is a reason for ϕ_5 in ϕ_4. Here, ϕ_5 is the proposition corresponding to the prejacent of *why*. We make the simplifying assumption that ϕ_6, the proposition serving as a reason for ϕ_5, is introduced as a dref, sidestepping other aspects of the interrogative semantics here. We capture the factivity of *why*-questions analogously to the presuppositional contents in Section 5.2. We assume that they carry a presupposition $\langle \phi(i) \subseteq \phi_5(i) \rangle_{i,\phi}$, which is interpreted in ϕ_{DC_S}. This allows us to state that the content is interpreted globally.

6.2. Clausal ellipsis

I adopt a dual approach to clausal ellipsis licensing (Chung 2013; Kroll and Rudin 2017; Rudin 2019; Anand et al. 2023), which involves (limited) syntactic isomorphism over the *v*P-domain, while the syntactic form of the rest of the elided clause is recovered based on semantic and pragmatic information in the discourse. This approach is motivated by observations that syntactic isomorphism is required for the argument structure of the elided clause, whereas mismatches are possible in other parts of the clause. Specifically, I adopt the syntactic condition in (48).

(48) Syntactic isomorphism condition: (adapted from Anand et al. 2023)
The argument-domain in the elided clause (i.e. the smallest phrase denoting a property of eventualities) is syntactically isomorphic to a phrase in the discourse context.

The condition (48) requires the syntactic isomorphism indicated in (49).

(49) Mary didn't [~~Mary~~ dance]$_{vP}$. Sue explained why not ~~[Mary danced]$_{vP}$~~.

Further, the recovery of the parts of the elided clause that lie outside of the *v*P-domain is constrained by local contextual entailment as well as independent pragmatic constraints on discourse interpretation. Clausal ellipsis generally allows for mismatches in modality or polarity (Rudin 2019), and when a modal is introduced in the ellipsis site, it often has vague or underde-

termined force and flavor (Anand et al. 2021). Therefore, we need to rule out the interpretations in (50), where the recovered ellipsis site includes an additional instance of negation, or some modal (indicated here as MODAL).

(50) Unavailable interpretations
  a. Mary didn't [~~Mary~~ dance]$_{vP}$. She explained why not ~~NEG [Mary danced]$_{vP}$~~.
  b. Mary didn't [~~Mary~~ dance]$_{vP}$. She explained why not ~~MODAL [Mary danced]$_{vP}$~~.

Based on the discussion in Section 3.3, we weaken Kroll's local contextual entailment to be understood as a constraint on interpretation, rather than licensing. In the intensional dynamic system presented here, this follows automatically from interpreting the elided clause within its local context, and we do not need to explicitly state local contextual entailment as a condition on ellipsis. An elided clause is interpreted as expressing a proposition which, when interpreted in its local context, does not create an inconsistent discourse (or violate other independent pragmatic constraints). This allows us to rule out the unavailable interpretations: (50a) would create an inconsistent discourse: The first sentence states that *Mary didn't dance*, and the factive use of *why* presupposes its prejacent (here: *Mary danced*). (50b) can be ruled out as being uninformative, given that the context already entails that *Mary didn't dance*.

6.3. Negative antecedent requirement

The unavailability of *why not* with positive antecedents comes about as a combination of ellipsis in the scope of negation and the factivity of *why*. We now have an explanation of why none of the interpretations in (51) are available.

(51)
  a. [Mary danced]$_{vP}$. She explained why not ~~[Mary dance]$_{vP}$~~
  b. [Mary danced]$_{vP}$. She explained why not ~~MODAL [Mary dance]$_{vP}$~~
  c. [Mary danced]$_{vP}$. She explained why not ~~NEG [Mary dance]$_{vP}$~~

The interpretations in (51a+b) would create an inconstent discourse: Since the first conjunct asserts that *Mary danced*, presupposing that *Mary did/would not dance* is inconsistent. In contrast, (51c) would yield a non-contradictory discourse. Here, we will have to assume that this interpretation is ruled out on the basis of syntax, ruling out the sequence of two instances of negation, or Gricean manner, which might favor an affirmative utterance over such a sequence.

6.4. Extensions: Pronominal anaphora and other negativity-tags

The analysis straightforwardly applies to anti-veridical pronominal anaphora, and explains why they cannot combine with veridical antecedents, when both the anaphor and antecedent are provided by assertions from the same speaker. The unacceptability of (52) arises as a contradiction:

(52)
  a. S: *Mary danced* $\rightsquigarrow$ $[\phi_1 \mid \phi_{DC_S} \Subset \phi_1]; [Danced_{\phi_1}\{Mary\}]$
  b. S: *That is a lie.* $\rightsquigarrow$ $[\phi_2 \mid \phi_{DC_S} \Subset \phi_2]; [Lie_{\phi_2}\{\phi_1\}]$

We treat *lie* as an anti-veridical assertive attitude like *be wrong*. This glosses over some details, but the key aspect here is its anti-veridical nature. The condition $Lie_{\phi_2}\{\phi_1\}$ refers to ϕ_1 in an anti-veridical context, stating that in an output state j, $\phi_1(j)$ is false in all ϕ_2-worlds. Because ϕ_1 and ϕ_2 are assertions by the same speaker S, this leads to inconsistency, and thus inacceptabiliy. In contrast, in (53), the first utterance makes available a counterfactual propostional dref:

(53) a. S: *Mary didn't dance.* $\rightsquigarrow [\phi_1 \mid \phi_{DC_S} \Subset \phi_1]; [\phi_2 \mid \phi_1 \Cap \phi_2 = \phi_\varnothing]; [Danced_{\phi_2}\{Mary\}]$
b. S: *That is a lie.* $\rightsquigarrow [\phi_3 \mid \phi_{DC_S} \Subset \phi_3]; [Lie_{\phi_3}\{\phi_2\}]$

(53b) refers to ϕ_2 in an anti-veridical context: it is false in all ϕ_3-worlds. That is consistent with (53a), which asserts that ϕ_2 is false in all ϕ_1-worlds. *S*'s subsequent assertions of ϕ_1 and ϕ_3 remain non-contradictory. The contrast between (53) and (52) arises from the requirement of discourse consistency, which prevents assertions by the same speaker from contradicting each other. However, this does not apply with anti-veridical discourse relations (in the sense of Asher and Lascarides 2003), such as the disagreement between two speakers *A* and *B* in (54).

(54) a. A: *Mary danced* $\rightsquigarrow [\phi_1 \mid \phi_{DC_A} \Subset \phi_1]; [Danced_{\phi_1}\{Mary\}]$
b. B: *That is a lie.* $\rightsquigarrow [\phi_2 \mid \phi_{DC_B} \Subset \phi_2]; [Lie_{\phi_2}\{\phi_1\}]$

Being asserted by two speakers disageeing with each other, ϕ_1 and ϕ_2 can be construed as contradictory. Here, *A* and *B* may need to resolve their disagreement to continue their conversation, but (54b) is an acceptable utterance. The interplay between propositional anaphora with antecedent polarity is understood by considering the veridicality of the local context of the anaphor, and the discourse relation between the utterances containing anaphor and antecedent—just like we established for negativity-tags. This provides further support for the claim that negativity-tags do not necessitate a special mechanism making reference to antecedent polartiy; rather, they are complex anaphoric expressions subject to regular constraints on discourse interpretation.

To extend the analysis to PolPs, we would assume that differences in absolute polarity stem from a propositional anaphor in a veridical vs. an anti-veridical context, drawing on Krifka's representational assumptions. Future research might explore a semantics of relative polarity features that encodes constraints on veridical (conjunctive) and non-veridical (disjunctive) discourse relations (e.g., in an SDRT framework, see Asher and Lascarides 2003). For *neither*-tags, the analysis might resemble that of *why not*: VPE imposes a condition of isomorphism over the elided VP, while the additive presupposition contributes parallelism with the previous utterance.

7. Conclusion

Anaphoric polarity-sensitivity is sensitivity to speaker commitments about the truth or falsity of contents in discourse. Therefore, the discourse-effect of negation cannot be explained at the clausal level, but is tied to its anti-veridical semantics. Further, discourse-negativity should be analyzed at a discourse-level which allows for pragmatic enrichment. This paper has presented a discourse-level analysis of discourse-negativity and the anaphoric polarity sensitivity of *why not* (Sections 5+6). It has also provided evidence for this view, showing that negativity-tags are licensed by counterfactual propositional content, even in cases where the counterfactual interpretation is supported by pragmatic reasoning (Section 2). Drawing from previous accounts of anaphoric polarity sensitivity (Section 3), I have shown how negativity-tags can be understood as complex anaphoric expressions. They involving an anaphoric element in an anti-veridical context and a second discourse requirement that enforces a non-contradictory interpretation of the utterances containing the anaphor and its antecedent (Section 4).

References

Anand, P. and V. Hacquard (2014). Factivity, belief and discourse. In L. Crnič and U. Sauerland (Eds.), *The Art and Craft of Semantics: A Festschrift for Irene Heim*, pp. 69–90. MIT Working

Papers in Linguistics.
Anand, P., D. Hardt, and J. McCloskey (2021). The Santa Cruz sluicing data set. *Language 97*(1), e68–e88.
Anand, P., D. Hardt, and J. McCloskey (2023). The domain of formal matching in sluicing. *Linguistic Inquiry*, 1–21.
Asher, N. and A. Lascarides (2003). *Logics of Conversation*. Cambridge: Cambridge University Press.
Bittner, M. (2009). Tense, mood, and centering. Manuscript, Rutgers University.
Brasoveanu, A. (2006). *Structured Nominal and Modal Reference*. Doctoral dissertation, Rutgers University.
Brasoveanu, A. (2010). Decomposing Modal Quantification. *Journal of Semantics 27*(4), 437–527.
Brasoveanu, A., K. De Clercq, D. Farkas, and F. Roelofsen (2014). Question Tags and Sentential Negativity. *Lingua 145*, 173–193.
Brasoveanu, A., D. Farkas, and F. Roelofsen (2013). N-words and sentential negation: Evidence from polarity particles and VP ellipsis. *Semantics and Pragmatics 6*(7), 1–33.
Chung, S. (2013). Syntactic identity in sluicing: How much and why. *Linguistic Inquiry 44*(1), 1–44.
Claus, B., A. M. Meijer, S. Repp, and M. Krifka (2017). Puzzling response particles: An experimental study on the German answering system. *Semantics & Pragmatics 10*(19), EA.
Collins, C. and P. M. Postal (2018). Disentangling two distinct notions of NEG raising. *Semantics and Pragmatics 11*(5), EA.
De Swart, H. (2000). Scope Ambiguities with Negative Quantifiers. In K. von Heusinger and U. Egli (Eds.), *Reference and Anaphoric Relations*, Studies in Linguistics and Philosophy, pp. 109–132. Dordrecht: Springer Netherlands.
Farkas, D. and K. Bruce (2010). On reacting to assertions and polar questions. *Journal of Semantics 27*(1), 81–118.
Farkas, D. F. and F. Roelofsen (2019). Polarity particles revisited. *Semantics and Pragmatics 12*(15), 1–25.
Fitzpatrick, J. (2005). The whys and how comes of presupposition and NPI licensing in questions. In *Proceedings of the 24th West Coast Conference on Formal Linguistics*, pp. 138–145.
Gajewski, J. (2007). Neg-Raising and Polarity. *Linguistics and Philosophy 30*(3), 289–328.
Gribanova, V. (2017). Head movement and ellipsis in the expression of Russian polarity focus. *Natural Language and Linguistic Theory 35*(4), 1079–1121.
Groenendijk, J. and M. Stokhof (1991). Dynamic predicate logic. *Linguistics and philosophy*, 39–100.
Gunlogson, C. (2004). *True to Form: Rising and Falling Declaratives as Questions in English*. Routledge.
Haegeman, L. and R. Zanuttini (1991). Negative heads and the Neg Criterion. *The Linguistic Review 8*(2–4), 233–251.
Heim, I. (1982). *The Semantics of Definite and Indefinite Noun Phrases*. Doctoral dissertation, University of Massachusetts Amherst.
Heim, I. (1983). On the projection problem for presuppositions. In P. H. Portner and B. H. Partee (Eds.), *Formal Semantics: The Essential Readings*, pp. 249–260. John Wiley & Sons.
Heim, I. (1992). Presupposition Projection and the Semantics of Attitude Verbs. *Journal of*

Semantics 9(3), 183–221.
Hofmann, L. (2018). Why not? – Polarity ellipsis in *why*-questions. UCSC Manuscript.
Hofmann, L. (2019a). The Anaphoric Potential of Indefinites under Negation and Disjunction. In *Proceedings of the 22nd Amsterdam Colloquium*, pp. 181–190.
Hofmann, L. (2019b). Sentential Negativity and Polarity-Sensitive Anaphora. In *ESSLLI 2019 Student Session*.
Hofmann, L. (2022). *Anaphora and Negation*. Doctoral dissertation, UC Santa Cruz.
Holmberg, A. (2013). The syntax of answers to polar questions in English and Swedish. *Lingua 128*, 31–50.
Jackendoff, R. (1969). An Interpretive Theory of Negation. *Foundations of Language 5*(2), 218–241.
Klima, E. (1964). Negation in English. In J. Fodor and J. Katz (Eds.), *The structure of language*, pp. 246–32. Prentice Hall, New Jersey.
Kramer, R. and K. Rawlins (2009). Polarity particles: An Ellipsis Account. In A. Schardl, M. Walkow, and M. Abdurrahman (Eds.), *NELS 38: Proceedings of the Thirty-Eighth Annual Meeting of the North East Lingusitic Society*, Volume 1, pp. 479–492.
Krifka, M. (2013). Response Particles as Propositional Anaphors. *Semantics and Linguistic Theory (SALT) 23*, 1–18.
Kroll, M. (2019). Polarity Reversals under Sluicing. *Semantics and Pragmatics 12*(18), EA.
Kroll, M. and D. Rudin (2017). Identity and interpretation: Syntactic and pragmatic constraints on the acceptability of sluicing. In A. Lamont and K. Tetzloff (Eds.), *NELS 47: Proceedings of the Forty-Seventh Annual Meeting of the North East Lingusitic Society*, pp. 177–190.
Kroll, M. I. (2020). *Comprehending Ellipsis*. Doctoral dissertation, UC Santa Cruz.
McCloskey, J. (2017). Ellipsis, Polarity, and the Cartography of Verb-Initial Orders in Irish. In E. Aboh, E. Haeberli, G. Puskás, and M. Schönenberger (Eds.), *Elements of Comparative Syntax: Theory and Description*, pp. 99–151. De Gruyter Mouton.
Merchant, J. (2001). *The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis*. Oxford Studies in Theoretical Linguistics. Oxford: Oxford University Press.
Murray, S. E. (2014). Varieties of update. *Semantics and Pragmatics 7*(2), 1–53.
Muskens, R. (1996). Combining Montague Semantics and Discourse Representation. *Linguistics and Philosophy 19*(2), 143–186.
Penka, D. (2007). Uninterpretable negative features on negative indefinites. In *Proceedings of the 16th Amsterdam Colloquium*, pp. 19–22. Citeseer.
Pope, E. (1972). *Questions and answers in English*. PhD Thesis, Diss. Massachusetts Institute of Technology.
Roelofsen, F. and D. Farkas (2015). Polarity particle responses as a window onto the interpretation of questions and assertions. *Language 91*(2), 359–414.
Rudin, D. (2019). Head-based syntactic identity in sluicing. *Linguistic Inquiry 50*(2), 253–283.
Snider, T. N. (2017). *Anaphoric Reference to Propositions*. Doctoral dissertation, Cornell University.
Stalnaker, R. (1978). Assertion. In P. Cole (Ed.), *Pragmatics*, Number 9 in Syntax and Semantics, pp. 315–332. London: Academic Press.
Stone, M. (1999). Reference to possible worlds. *RuCCS Report 46*.
Zeijlstra, H. (2004). *Sentential Negation and Negative Concord*. Doctoral dissertation, University of Amsterdam, Utrecht: LOT Publications.