Neg-raising: Focus and Implicatures[1]

Yimei XIANG — *Harvard University*

Abstract. This paper offers a 'feature-based' account of neg-raising (NR) phenomenon. This account integrates Alternative Semantics for focus into the grammatical view of scalar implicatures (SIs). I argue that the derivation of an NR or non-NR reading is determined by the distributions of two features, namely, the SI feature [+σ] of NR predicates and the [+F] feature of focused items.

Keywords: Neg-raising, Negation, Focus, Implicatures, Alternatives

1. Introduction

'Neg-raising' (henceforth NR) is a phenomenon that the clause-mate negation of a sentence-embedding verb is intuitively interpreted as taking scope in the complement clause. For instance, sentences like (1a) and (2a) are intuitively interpreted as (1b) and (2b), respectively. The narrow scope readings in (1b) and (2b) are called 'NR readings', and the wide scope readings in (1c) and (2c) are called 'non-NR readings'. Predicates that allow NR readings are called 'neg-raising predicates' (henceforth NRPs).

(1) a. John doesn't believe that it is raining.
b. $\rightsquigarrow$ John believes that it isn't raining.
c. $\not\rightsquigarrow$ It isn't the case that John believes that it is raining.

(2) a. John doesn't want to leave here.
b. $\rightsquigarrow$ John wants not to leave here.
c. $\not\rightsquigarrow$ It isn't the case that John wants to leave here.

NR was firstly conceived as a syntactic phenomenon (Fillmore 1963, Lakoff 1969, a.o.). This early syntactic approach postulates that negation in a sentence like (1a) is generated and interpreted in the embedded clause, but it raises into the main clause and is pronounced there. While being popular in its early period, the syntactic view is challenged by various pragmatic and semantic approaches. Representative studies following a semantic/pragmatic vein include: the conventional implicature-based approach (Horn 1978), the presupposition-based approach (Bartsch 1973; Gajewksi 2005, 2007), the scalar implicature (SI)-based approach (Romoli 2012, 2013) and the PPI-based approach (Homer 2012). My proposal is developed from the presuppositional-based account and the SI-based account.

[1]For helpful comments and criticism, I want to thank Gennaro Chierchia, Kai von Fintel, Danny Fox, Michael Franke, Jim Huang, Manfred Krifka, Jacopo Romoli, Uli Sauerland, Benjamin Spector and the audiences at SuB 18.

In the presupposition-based approach, Gajewksi (2005, 2007) adopts Bartsch's (1973) view that an NRP gives rise to an excluded middle homogeneity inference, which says that the subject is opinionated about the truth or falsity of the complement clause. For instance, the excluded middle for '*John believes* ϕ' is '*either John believes* ϕ *or John believes* $\neg\phi$'. To explain the defeasibility of excluded middles, Gajewski (2007) assumes that excluded middles are 'soft presuppositions' in the sense of Abusch (2002, 2010), derived from the lexical entries of NRPs.

Romoli (2012, 2013) adopts the viewpoint that NRPs trigger excluded middles. However, he argues that excluded middles are not presuppositions but instead SIs, more precisely, optional indirect SIs activated by *relevance*. My proposal follows his idea that excluded middles are SIs. However, my approach differs greatly in the way of assuming how excluded middles get activated: instead of attributing the activation to the notion of relevance, I argue that a predicate taking an SI-feature [+σ] activates an excluded middle obligatorily. In addition, I consider non-NR readings as results of intervention effects from focus. To highlight the core difference between this two SI-based accounts, I will call Romoli's (2012, 2013) approach as 'relevance-based' while mine as 'feature-based'.

The primary goal of this paper is to explain the distributions of NR and non-NR readings based on the distributions of features. Gajewski (2005) observes that it is marked to negate an NRP without assuming an excluded middle. To say that the subject is not opinionated at the value of the embedded clause, there must be a stress on the negative auxiliary or on the NRP. For instance, sentences in (3a-b) do not imply the NR inference in (3c) (capitals indicate stress). These facts suggest that the derivation of an NR or non-NR inference is related to F(ocus)-marking.

(3) a. John DOESn't believe that it is raining.
b. John doesn't BELIEVE that it is raining.
c. $\not\leadsto$ John believes it isn't raining.

The rest of the paper is organized as follows. In section 2, I summarize the SI-based account from Romoli (2012, 2013) and the grammatical theory of SIs he adopts. In section 3, I offer a feature-based account to explain the derivations of NR and non-NR interpretations, and in particular, the cancellation effects from focus. This feature-based account can be summarized in two sentences. First, the requirements of feature-checking and avoiding G-triviality provide a group of eligible EXH-structure candidates. Second, a set of OT constraints select out the winning EXH-structures and the preferred interpretations.

2. The relevance-based approach

2.1. Excluded middles

Bartsch (1973) indicates that the NR phenomenon is influenced by a pragmatic presupposition, called 'excluded middle'. This inference says that the subject is opinionated about the truth or falsity of its complement clause ϕ. In a positive case like (4), the excluded middle is asymmetrically entailed by the assertion and hence doesn't affect the overall meaning. In a negative context like (5), the excluded middle projects over negation, and an NR reading is derived as a logical consequence of the negative assertion and the excluded middle.

(4) a. John believes that it is raining. $\mathbf{bel}\phi$
b. John has a opinion as to whether it is raining. $\mathbf{bel}\phi \vee \mathbf{bel}\neg\phi$

(5) a. John doesn't believe that it's raining. $\neg\mathbf{bel}\phi$
b. John has an opinion as to if it's raining. $\mathbf{bel}\phi \vee \mathbf{bel}\neg\phi$
c. John believes that it isn't raining. $\mathbf{bel}\neg\phi$

Gajewski (2005, 2007) adopts the idea of excluded middles. However, instead of grouping excluded middles as pragmatic presuppositions, he argues that they are soft presuppositions lexically specified for NRPs. The label 'soft' is coined by Abusch (2002, 2010) to describe the group of presuppositions that are defeasible.

In the very recent work by Romoli (2012, 2013), he argues that excluded middles are not presuppositions but instead SIs. This claim is supported by the non-presuppositional projection status of excluded middles: although excluded middles survive under negation, they do not project in other embedding contexts. As illustrated in (7), when the NRP *believe* is embedded in the antecedent of a conditional, under an epistemic modal, or in an interrogative, the excluded middle is not implied. For sake of comparison, as shown in (8), the existential presupposition of an *it*-cleft is implied in all those embeddings.

(6) a. Bill believes that Sue is here.
b. Bill doesn't believe that Sue is here.
c. $\rightsquigarrow$ Bill is opinionated with respect to if Sue is here.

(7) a. If Bill believes that Sue is here, he will come.
b. Perhaps Bill believes that Sue is here.
c. Does Bill believe that Sue is here?
d. $\not\rightsquigarrow$ Bill is opinionated with respect to if Sue is here.

(8) a. It was Mary who killed Bill.

b. It wasn't Mary who killed Bill.
c. If it was Mary who killed Bill, she should confess.
d. Perhaps it was Mary who killed Bill.
e. Was it Mary who killed Bill?
f. $\rightsquigarrow$ Someone killed Bill.

Furthermore, Romoli (2012, 2013) notices that SIs behave in the same way as excluded middles do. The implicature in (9c) survives as an indirect SI in the negative sentence (9b), but it does not project out when the scalar item is embedded inside the antecedent of a conditional like (10a).

(9) a. Every student came.
b. Not every student came.
c. $\rightsquigarrow$ Some student came.

(10) a. If every student came, the party was a success.
b. $\not\rightsquigarrow$ Some student came.

2.2. The grammatical view of scalar implicatures

SIs were primarily considered as a wholly pragmatic phenomenon in the Gricean framework (Grice 1975). However, in the recent works by Chierchia et. al. (2013) a.o., SIs are conceived as grammatical matters. This grammatical view assumes that a scalar item triggers a set of alternatives. The alternative set are computed in the same way as questions (Hamblin 1973) or focus (Rooth 1985, 1992). A schematized recursive definition is given below, adopted from Chierchia (2013).

(11) Basic Clause: For any lexical entry α, $\mathscr{A}lt(\alpha)$ =
a. $\{[\![\alpha]\!]\}$ if α is lexical and does not belong to a scale;
b. $\{[\![\alpha_1]\!], ..., [\![\alpha_n]\!]\}$ if α is lexical and part of a scale $\langle[\![\alpha_1]\!], ..., [\![\alpha_n]\!]\rangle$.
Where $\mathscr{A}lt$ is a function from expressions to a set of interpretations.

(12) Recursive Clause: $\mathscr{A}lt(\beta(\alpha)) = \{b(a) : b \in \mathscr{A}lt(\beta) \text{ and } a \in \mathscr{A}lt(\alpha)\}$

Alternatives keep growing until factored into meaning via a covert exhaustivity operator EXH. This operator, with a meaning akin to *only*, affirms the prejacent and negates an excludable subset of the alternative set (notation: $\mathscr{E}xcl(p)$).[2] The excludable alternatives are all the ones that can be

[2]Here and throughout the paper, the symbols EXH and p are sloppily used for both syntactic phrases and truth conditions. A stricter semantic representation for EXH should be as follows, where S is the c-commanded phrase of EXH.

consistently negated with the assertion on its own.[3]

(13) a. $\text{EXH}(p) = \lambda w.p(w) \wedge \forall q \in \mathscr{E}xcl(p)[\neg q(w)]$
b. $\mathscr{E}xcl(p) = \{q \in \mathscr{A}lt(p) : \lambda w[\neg q(w)] \cap p \neq \emptyset\}$

As a simple illustration, the indirect SI in (14a) is derived as an entailment of global exhaustification, as schematized in (14c). Note that the local exhaustification structure in (14d) is out, because it yields a meaning equivalent to the plain assertion, violating the MaxStrength condition (see section 3.4.1 for extensive discussions).

(14) a. Not every student came. $\rightsquigarrow$ Some student came.
b. $\mathscr{A}lt(\neg\phi_{\text{SOME}}) = \{\neg\phi_{\text{SOME}}, \neg\phi_{\text{EVERY}}\}$
c. $\text{EXH}(\neg\phi_{\text{EVERY}}) = \neg\phi_{\text{EVERY}} \wedge \neg\neg\phi_{\text{SOME}} = \neg\phi_{\text{EVERY}} \wedge \phi_{\text{SOME}}$ (√)
d. $\neg\text{EXH}(\phi_{\text{EVERY}}) = \neg\phi_{\text{EVERY}}$ (×)

Adopting the grammatical theory of SIs, Romoli (2012, 2013) proposes that an NRP is associated with two alternatives, the assertion itself and an excluded middle. Those alternatives must be used up by a c-commanding EXH-operator. In a positive case, the excluded middle is asymmetrically entailed by the assertion and hence the EXH is semantically vacuous, as schematized in (15). In a negative case, exhaustifying over negation denies the negated excluded middle, giving rise to an NR reading, as schematized in (16).

(15) a. $\mathscr{A}lt\ (\mathbf{bel}\phi) = \{\mathbf{bel}\phi, \mathbf{bel}\phi \vee \mathbf{bel}\neg\phi\}$
b. $\text{EXH}\ (\mathbf{bel}\phi) = \mathbf{bel}\phi$

(16) a. $\mathscr{A}lt\ (\neg\mathbf{bel}\phi) = \{\neg\mathbf{bel}\phi, \neg[\mathbf{bel}\phi \vee \mathbf{bel}\neg\phi]\}$
b. $\text{EXH}\ (\neg\mathbf{bel}\phi) = \neg\mathbf{bel}\phi \wedge \neg\neg[\mathbf{bel}\phi \vee \mathbf{bel}\neg\phi] = \mathbf{bel}\neg\phi$

2.3. Activation of alternatives and non-NR readings

Romoli (2012, 2013) subdivides alternatives into two classes, obligatory ones and optional ones. The activation of the former class is determined by the value of its corresponding grammatical features, while that of the latter is subject to relevance. In Chierchia's (2006) convention, for instance, scalar alternatives are activated when the SI-feature $[\sigma]$ takes the '+' value, and are

(1) $[\![{}_{\text{EXH}\ \ S}]\!]^{w,g} = [\![S]\!](w) \wedge \forall S' \in \mathscr{E}xcl(S)[\neg[\![S']\!](w)]$

[3]Fox (2007) gives a different lexical entry for the EXH-operator, and argues that EXH only negates the *Innocently Excludable* (IE) alternatives. See Chierchia (2013) and section 3.3 for details.

inactive when $[\sigma]$ gets the '-' value. An alternative is said to be obligatory iff the corresponding feature always takes the '+' value.

Romoli claims that factive alternatives are obligatorily activated by an agreement feature $[\pi]$. This feature always takes the '+' value and always activates a factive alternative. In a negative sentence like (17b), the negated factive alternative is stronger than the negative assertion and is excludable. Hence, as schematized in (19), exhaustification proceeding via a global EXH negates the negated factive alternative, giving rise to a meaning that entails the factive inference.

(17) a. John knows that it is raining.
b. John doesn't know that it is raining.
c. $\rightsquigarrow$ It is raining.

(18) a. $\mathscr{A}lt$ $(\mathbf{know}_{[+\pi]}\phi) = \{\mathbf{know}\phi, \phi\}$
b. EXH $(\mathbf{know}_{[+\pi]}\phi) = \mathbf{know}\phi$

(19) a. $\mathscr{A}lt$ $(\neg\mathbf{know}_{[+\pi]}\phi) = \{\neg\mathbf{know}\phi, \neg\phi\}$
b. EXH $(\neg\mathbf{know}_{[+\pi]}\phi) = \neg\mathbf{know}\phi \wedge \neg\neg\phi = \neg\mathbf{know}\phi \wedge \phi$

As for optional alternatives, Romoli assumes that an alternative falling in this class is inactive unless it is relevant to the current question. According to the standard assumption that a question is associated with a partition of the common ground, 'relevance' can be defined as in (20), where Q is the partition set associated with the question. This definition says an assertion is relevant iff it does not discriminate between cell-mates. Namely, for each partition associated with the question, a relevant assertion must eliminate either all the worlds in that cell or none of them.

(20) **Relevance** : A proposition p is relevant to a question Q iff p is (contextually equivalent to) the union of some subset of Q. Heim (2011)

Romoli further classifies excluded middles as optional alternatives. He claims that the assertions in (21a) and (22a) can be thought of as answers to the questions in (21b) and (22b), respectively. The associated partition sets of these two questions are given in (21c) and (22c), of which only the former contains an alternative derived from excluded middles (viz. $\neg[\mathbf{Bel}\phi \vee \mathbf{Bel}\neg\phi]$). On a relevance-based account, we can say that the negated excluded middle indiscriminates in Q_1 and discriminates in Q_2. Based on this idea, Romoli concludes that the reason for the non-NR interpretation in (22a) is that here the (negated) excluded middle alternative is inactive.

(21) a. Bill doesn't believe that it is raining.
b. What does Bill believe about whether it is raining?
c. $Q_1 = \{\mathbf{Bel}\phi, \mathbf{Bel}\neg\phi, \neg[\mathbf{Bel}\phi \vee \mathbf{Bel}\neg\phi]\}$

(22) a. John DOESn't believe that it is raining.
b. Does John believe that it is raining?
c. $Q_2 = \{\mathbf{Bel}\phi, \neg\mathbf{Bel}\phi\}$

As for the non-NR case in (23) which has narrow focus on the NRP, the (negated) excluded middle is relevant and activated. Romoli (2012, 2013) claims that here the EXH-operator has to be applied locally, so as to be consistent with the denial of excluded middle in the continuation.

(23) a. John doesn't BELIEVE that it is raining. He isn't sure.
b. What does John do with respect to whether it is raining?
c. $Q_3 = \{\mathbf{Bel}\phi, \mathbf{Bel}\neg\phi, \neg\mathbf{Bel}\phi \wedge \neg\, \mathbf{Bel}\neg\phi, \mathbf{hope}\phi, ...\}$

2.4. Problems with the relevance-based account

To sum things up, Romoli (2012, 2013) manipulates the NR reading in (24a) as a logical consequence of global exhaustification. In response to the non-NR readings in (24b), he postulates that the excluded middle is inactive in (24a) because of its irrelevance to the current question. As for the non-NR reading in (24c), he assumes that here the EXH-operator has stay locally, so as to stay consistent with the negated excluded middle in the continuation. I agree with his analysis regarding to NR readings but disagree with those to non-NR readings.

(24) a. John doesn't believe that it is raining.
b. John DOESn't believe that it is raining.
c. John doesn't BELIEVE that it is raining.

First of all, it is unclear how 'relevance' determines whether an alternative is activated, since empirically there is no discernible difference between obligatory alternatives and optional alternatives in terms of the way of being activated. Compare (24a-b) with the examples in (25). The factive inference, which Romoli claims to be obligatorily activated, is cancelled in exactly the same way as the excluded middle is: the factive inference is denied-able when the negative auxiliary is F-marked, and is not denied-able in the absence of this F-marking.

(25) a. John doesn't know it is raining. # It is not raining.
b. John DOESn't know it is raining, since it is not raining.

Second, to get the non-NR reading in (24c), the excluded middle doesn't have to be negated in the continuation, as illustrated in (26). What's more, the continuation in (26) sounds even more natural than the negated exclude middle in (23a). Therefore, being consistent with a continuation shouldn't be the cause for an EXH to be applied locally.

(26) John doesn't BELIEVE that it is raining, he KNOWS that it is raining.

Third, it is problematic to say that the scope of an EXH can be decided by the continuation, which incorrectly predicts that excluded middles can be suspended even in the basic negation case in (27).

(27) John doesn't believe that it is raining, # he is not sure.

3. My proposal: a feature-based account

I offer a feature-based account to explain the distributions of NR and non-NR readings. I follow Romoli's (2012, 2013) view that excluded middles are SIs, but argue that their activations are not bound to the confines of relevance. More importantly, the account that I assume highlights the role of focus on the cancellation of excluded middles, and provides principles to restrict the landing position of EXHs.

3.1. An overview

I propose that the status of an NR inference is determined by the distributions of two features, a lexically endowed SI feature [σ] and a contextually dependent focus feature [F]. The [σ] feature is adopted from Chierchia's (2006) analysis on scalar items. I assume that when a predicate (**P**) has an [+σ] feature in its lexicon, it activates an excluded middle alternative, as schematized in (28). In this sense, NRPs are predicates containing an [+σ] feature.[4]

[4]A similar assumption has been drawn in Romoli (2012, 2013). However, he assumes that the value of the [σ] feature in NRPs is restricted by relevance. In my view, although the [σ] feature of NRPs can take the '-' value, it isn't decided by relevance. As far as I can see, the value of [σ] varies in the following ways. First, it varies cross-linguistically. For instance, English *hope* takes [+σ], while German *hoffen* takes [-σ] in the third person use. Second, the value of [σ] of a predicate can be affected by the meaning of the complement. For instance, *want* is uncontroversially NR, however, it can also be paraphrased as the non-NR predicate *desire*, especially when its complement refers to some high expectation. Consider the examples quoted from Homer (2012), (1a) doesn't imply (1b). The variation with *want* can be reduced to Marantz's (1984) observation that the interpretation of a predicate is sensitive to its object. Third, even in a basic negation case like (2), some native speakers suggest that it is marginally acceptable to interpret a canonical NRP as a non-NR one.

(1) a. My great-grandparents didn't want to spend all their spare time on the internet.
b. My great-grandparents wanted not to spend all their spare time on internet.

(2) John doesn't believe it is raining, ?he isn't opinionated.

(28) a. $\mathscr{A}lt\ (\mathbf{P}_{[-\sigma]}) = \{\lambda x \lambda \phi . \mathbf{P}(\phi)(x)\}$ Non-NRPs

b. $\mathscr{A}lt\ (\mathbf{P}_{[+\sigma]}) = \{\lambda x \lambda \phi . \mathbf{P}(\phi)(x), \lambda x \lambda \phi . [\mathbf{P}(\phi)(x) \vee \mathbf{P}(\neg\phi)(x)]\}$ NRPs

The idea of focus feature [+F] is inspired by Rooth's (1985, 1992, 1996) Alternative Semantics for focus. I assume that an F-marked item is assigned with an [+F] feature. This feature activates a set of focus-related alternatives $\mathscr{A}lt_F(p)$, namely, a subset of $[\![p]\!]^f$ (the focus value of p) containing the prejacent p and particular contextually selected elements. [5]

All of the activated alternatives need to be used up by an EXH-operator, which affirms the prejacent p and negate all the alternatives that are not entailed by the prejacent (Magri 2010, Chierchia et al. 2013). In a basic negation case like (29a), I adopt Romoli's proposal that here the EXH-operator takes scope over negation, yielding an NR reading. However, as for the non-NR readings in (30a) and (31a), different from Romoli's view, I argue that the LF structures of these two sentences must take double exhaustification and local exhaustification, respectively.

(29) a. John doesn't believe that it is raining.

b. EXH $\neg$ [John believes$_{[+\sigma]}$ it's raining]

(30) a. John DOESn't believe that it is raining, he isn't sure.

b. EXH $\neg_{[+F]}$ EXH [John believes$_{[+\sigma]}$ it's raining], he isn't sure.

(31) a. John doesn't BELIEVE that it is raining, he knows it.

b. $\neg$ EXH [John believes$_{[+\sigma,+F]}$ it's raining], he knows it.

The heart of my proposal is that, where to insert an EXH is determined by two unviolatable conditions and two OT constraints, listed out in (32) and (33), respectively. In the next section, I will show how to select out the winning LF structures based on those conditions and constraints.

(32) a. Avoid unchecked features and syntactically vacuous EXH-operators.

b. Avoid G-triviality.

(33) a. **ExclF**: there must be some excludable F-alternative.

b. **MaxStrength**: do not exhaustify S in [$_{S'}$ ... S ...] if it leads to a reading that is weaker than or equivalent to S'.

[5] The [+F] feature and the $\mathscr{A}lt_F(p)$ set resemble the focus interpretation operator '$\sim$' and the C variable in Rooth (1996: 279): "ϕ is a syntactic phrase and C is a syntactically covert semantic variable, $\phi \sim C$ introduces the presupposition that C is a subset of $[\![\phi]\!]^f$ containing $[\![\phi]\!]^0$ and at least one other element."

3.2. Feature-checking

An eligible EXH-structure should be first syntactically well-formed, especially needs to satisfy two requirements on feature-checking. First, a well-formed structure shouldn't contain any unchecked feature. According to Chierchia (2006, 2013), every feature that takes the '+' value is forced to enter into an agreement relation with a c-commanding EXH; in absence of such an EXH, a feature can only take the '-' value. This requirement rules out the EXH-structures in (34), each of which contains an unchecked feature [+σ]. Second, a well-formed structure shouldn't contain any syntactically vacuous EXH-operator. Namely, every occurrence of EXH should probe for some feature that takes the '+' value.[6] This requirement rules out the EXH-structures in (35).

(34) a. *[... some$_{[+\sigma]}$]
b. *[... some$_{[+\sigma]}$ ... [EXH [... some$_{[+\sigma]}$ ...]]]

(35) a. *EXH [... some$_{[-\sigma]}$ ...]
b. *EXH [EXH [... some$_{[+\sigma]}$...]]

Syntactically well-formed EXH-structures for sentences containing negation and NRPs are list below. In section 3.3, I will show that the sentence in (37a) must choose the double exhaustification structure in (37c), since its competitive, the global exhaustification one, yields a G-trivial reading. In section 3.4, I will show that global exhaustification is optimal for (36a), and that local exhaustification is strongly preferred for (38a).

(36) a. John doesn't believe that it is raining.
b. EXH $\neg$ [John believes$_{[+\sigma]}$ it's raining]
c. $\neg$ EXH [John believes$_{[+\sigma]}$ it's raining]

(37) a. John DOESn't believe that it is raining
b. EXH $\neg_{[+F]}$ [John believes$_{[+\sigma]}$ it's raining]
c. EXH $\neg_{[+F]}$ EXH [John believes$_{[+\sigma]}$ it's raining]

(38) a. John doesn't BELIEVE that it is raining
b. EXH $\neg$ [John believes$_{[+\sigma,+F]}$ it's raining]
c. $\neg$ EXH [John believes$_{[+\sigma,+F]}$ it's raining]

[6] This description is still sloppy. A stricter way to state this requirement is as follows: insert an occurrence of EXH to an assertion α iff α contains an expression β such that β takes an alternative-sensitive feature and that β doesn't fall within the scope of an occurrence of EXH appearing in α.

3.3. Avoid G-triviality

After feature-checking, syntactically well-formed structures will then be sent to the process of semantic compositions. The most important filter in this process is avoiding G-triviality. Chierchia (2013) defines G-triviality as a special case of L-triviality: L-trivial sentences refer to those that are tautologous or contradictory in the traditional sense, while G-triviality says that a sentence receives the same value (1 or 0) regardless how the lexical terminals are replaced in the structure. As a simple illustration, the sentence in (40a) is both G-trivial and L-trivial, while the one in (40b) is merely L-trivial. Unlike the infelicitous sentence in (40b) which can be utterable especially under certain embeddings, the sentence in (40a) is always perceived as "ungrammatical".

(39) **G-triviality**: A sentence ϕ is G-trivial iff for any situation s and model M, $[\![\phi']\!]^{M,s}$ = same (where same is either 1 or 0) and ϕ' is obtained from ϕ by an arbitrary substitution of its lexical terminal nodes. Chierchia (2013)

(40) a. * some student but John smokes.
b. # John smokes and doesn't smokes.

In the grammatical view of SIs, a sentence is considered as G-trivial once applying exhaustification yields a semantic contradiction (or a tautology). Based on this idea, Chierchia (2006, 2013) claims that the reason why NPI *any* must stay in DE context is that, exhaustifying the D-alternatives of *any* in a non-DE context yields a contradiction. I won't get into more details about this idea, but the crucial point is that the meaning of an EXH-structure cannot be G-trivial.

There are different approaches to avoid G-triviality. Besides the way of changing the monotonicity pattern of the embedding context, we can also manipulate the scope or the quantity of EXH-operators. In the case of focused negation, we have seen two LFs that satisfy the feature-checking requirements: one is to insert an occurrence of EXH above negation, as in (41b), and the other is to insert an EXH both above and below negation, as in (41c). I call the former structure "global exhaustification", and the latter "double exhaustification".

(41) a. John DOESn't believe that it is raining
b. EXH $\neg_{[+F]}$ [John believes$_{[+\sigma]}$ it's raining] (×)
c. EXH $\neg_{[+F]}$ EXH [John believes$_{[+\sigma]}$ it's raining] (√)

Global exhaustification yields G-triviality, as shown by the schematized derivation in (42). In this EXH-structure, the alternative set used by the global EXH includes both the affirmed and the negated excluded middles (as underlined), both of which are excludable on their own. However, negating these two alternatives gives rise to a semantic contradiction.

(42) a. $\mathscr{A}lt\ (\neg_{[+F]}\ \mathbf{bel}_{[+\sigma]}\phi) = \{\mathbf{bel}\phi,\ \neg\mathbf{bel}\phi,\ \underline{\mathbf{bel}\phi \vee \mathbf{bel}\neg\phi},\ \underline{\neg[\mathbf{bel}\phi \vee \mathbf{bel}\neg\phi]}\}$

b. $\text{EXH}\ [\neg_{[+F]}\ \mathbf{bel}_{[+\sigma]}\phi] = \neg\mathbf{bel}\phi \wedge \neg\mathbf{bel}\phi \wedge \neg[\mathbf{bel}\phi \vee \mathbf{bel}\neg\phi] \wedge \neg\neg[\mathbf{bel}\phi \vee \mathbf{bel}\neg\phi] = \bot$

The contradiction can be avoided by applying double exhaustification. According to the principle in (43) from Chierchia (2013), the excluded middle has been (vacuously) used up by the local EXH, and hence is not available for the global EXH. This structure yields a non-NR reading, as expected.

(43) $\mathscr{A}lt\ (\text{EXH}(\phi)) = \{\text{EXH}(\phi)\}$
(once an EXH applies to some expression ϕ, ϕ's alternatives are no longer available)

(44) a. $\mathscr{A}lt(\neg_{[+F]}\ \text{EXH}\ \mathbf{bel}_{[+\sigma]}\phi) = \{\neg\ \text{EXH}\ \mathbf{bel}\phi,\ \text{EXH}\ \mathbf{bel}\phi\}$

b. $\text{EXH}\ \neg_{[+F]}\ \text{EXH}\ \mathbf{bel}_{[+\sigma]}\phi = \neg\ \text{EXH}\ \mathbf{bel}\phi \wedge \neg\ \text{EXH}\ \mathbf{bel}\phi = \neg\ \text{EXH}\ \mathbf{bel}\phi = \neg\mathbf{bel}\phi$

To sum up, as for the case with focused negation, using one single global EXH-operator to check off the [+F] feature on negation and the [+σ] feature on NRP yields a semantic contradiction. Hence, the LF has to take double exhaustification, giving rise to a non-NR reading.

3.4. An optimality theory for NR

3.4.1. The conditions

In this part, I offer an OT system to explain why an unmarked negative sentence prefers an NR reading, and why a sentence with an F-marked NRP has to be interpreted as non-NR. Two constraints are considered in this system, ranked in order of priority.

(45) a. **ExclF**: there must be some excludable F-alternative.

b. **MaxStrength**: do not exhaustify S in $[_{S'}$ ... S ...$]$ if it leads to a reading that is weaker than or equivalent to S'.

The ExclF condition is motivated by excludability inference of overt *only*, which requires the focused constituent to trigger at least one excludable alternative. As a simple illustration, the answer in (46c) is infelicitous because all the F-alternatives are entailed by the prejacent. As for the case of the covert EXH-operator, such a requirement is less rigid, as one can stress the strongest alternative as in (46d) without yielding any infelicity. [7] However, we can still observe the effect

[7]Note that the ExclF condition is different from the AvoidF principle in Schwarzschild's (1999): "F-mark as little as possible (without violating GIVENness)."The former concerns the most as to whether an F-marked item contrasts to any alternative, while the latter cares the most as to whether a non-given expression should be F-marked. In other words, the former is about the semantic consequence of F-marking, while the latter determines whether to and where to assign an F-mark.

of ExclF in negative sentences. For an utterance like (47a), the global EXH structure in (47b) is bad because under such a structure the focused scalar item doesn't trigger any excludable F-alternatives. Thus the EXH-operator has to be applied locally, giving rise to a cancellation effect on the SI $\neg\phi_{\text{ALL}}$.

(46) a. A: Which of John and Mary are you going to invite?
b. B: Only JOHN, (not Mary/both).
c. B: # Only BOTH.
d. B: BOTH.

(47) a. I didn't see SOME of the students. (I saw ALL of the students.)
b. *EXH $[\neg\phi_{\text{SOME}_{[+\sigma,+F]}}]$
c. $\neg$EXH $[\phi_{\text{SOME}_{[+\sigma,+F]}}] = \neg\ [\phi_{\text{SOME}} \wedge \neg\phi_{\text{ALL}}] = \neg\phi_{\text{SOME}} \vee \phi_{\text{ALL}}$
(Either I didn't see some of the students, or I saw all of the students.)

The MaxStrength condition (also named as "Strongest Meaning Hypothesis" in the literature) has been extensively discussed in a bunch of works, including Chierchia, Fox and Spector (2013), Fox and Spector (2009), Magri (2011), Romoli (2012), among the others. However, each implementation has its own nuances and makes different predictions. For instance, Chierchia, Fox and Spector (2013) provides two candidates and leaves their final decision open. One candidate is to say that the preferred reading is always the strongest possible one (if there is one) among all the possible readings, and the other candidate is to say that the meaning of $[_{S'}\text{EXH}[S]]$ shouldn't be weaker than S'. Most of the works mentioned above are inclined to the second one, so is MaxStrength. The definition of MaxStrength used in this paper follows Fox and Spector (2009), according to which MaxStrength disallows not only weakening exhaustifications but also semantically vacuous exhaustifications. Without this move, the obligatoriness of NR readings in unmarked sentences cannot be captured.

When the scalar item *some* in (47a) is not F-marked, the ExclF condition becomes irrelevant. Instead, the MaxStrength condition plays a role and requires to exercise global exhaustification. The meaning of a global exhaustification structure is schematized in (49b). It successfully predicts that the SI $\neg\phi_{\text{ALL}}$ cannot be canceled in an unmarked sentence like (48).

(48) I didn't see some of the students. (# I saw all of the students.)
$\rightarrow$ I saw some of the students. (viz. It isn't the case that I didn't see any student.')

(49) a. $^*\neg\text{EXH}\phi_{\text{SOME}_{[+\sigma]}} = \neg\phi_{\text{SOME}} \vee \phi_{\text{ALL}}$
(Either 'I didn't see some of the students', or 'I saw all the students.')
b. EXH $[\text{SOME}_{i,[+\sigma]}\ [\neg\phi_{t_i}]] = [\text{SOME}_i\ [\neg\phi_{t_i}]\ ] \wedge \neg\ [\text{ALL}_i\ [\neg\phi_{t_i}]]$
('I didn't see some of the students, but not that I didn't see any student.')

3.4.2. Distributing NR and non-NR readings

In section 3.3, I have shown that the requirement of avoiding G-triviality requires the case of focused negation to take double exhaustification. In this part, I will show how the OT constraints select out the winning structures for the other two cases.

The basic negation case In absence of F-marked items, the LF structure contains one and only one occurrence of EXH-operator that checks off the [+σ] feature on *believe*. Both the global EXH-structure and the local one are syntactically well-defined, and are both independent from the ExclF constraint. However, the local one leads to a reading that is equivalent to the plain assertion, as schematized in (51), and hence violates the MaxStrength constraint. In contrast, as schematized in (52), global exhaustification gives rise to a stronger reading, namely, the NR reading.

(50) John doesn't believe that it is raining.

	Input: $\mathbf{bel}_{[+\sigma]}\phi$	ExclF	MaxStrength
☞	EXH[$\neg\mathbf{bel}_{[+\sigma]}\phi$]		
	$\neg$[EXH [$\mathbf{bel}_{[+\sigma]}\phi$]]		*!

(51) a. $\mathscr{A}lt\ (\mathbf{bel}_{[+\sigma]}\phi) = \{\mathbf{bel}\phi, \mathbf{bel}\phi \vee \mathbf{bel}\neg\phi\}$
b. *$\neg$EXH [$\mathbf{bel}_{[+\sigma]}\phi$] = $\neg\mathbf{bel}\phi$

(52) a. $\mathscr{A}lt\ (\neg\mathbf{bel}_{[+\sigma]}\phi) = \{\neg\mathbf{bel}\phi, \neg[\mathbf{bel}\phi \vee \mathbf{bel}\neg\phi]\}$
b. EXH [$\neg\mathbf{bel}_{[+\sigma]}\phi$] = $\neg\mathbf{bel}\phi \wedge \neg\neg[\mathbf{bel}\phi \vee \mathbf{bel}\neg\phi] = \mathbf{bel}\neg\phi$

The case of stressed NRP A stressed NRP carries two features, the SI feature [+σ] that activates an excluded middle, and the [+F] feature that triggers a set of focus-alternatives. For a sentence like (53), the focus-alternative $\mathbf{know}\phi$ asymmetrically entails the asserted component $\mathbf{bel}\phi$ in the terminal level, while its negation $\neg\mathbf{know}\phi$ is asymmetrically entailed by $\neg\mathbf{bel}\phi$ in the global level. The landing position of EXH determines at which level the focus-alternative is factored into meaning. The predominant constraint ExclF requires the focused item to activate an excludable alternative. Hence, according to the results drawn in (54d) and (55d), we conjecture that here the EXH-operator has to be applied under negation, giving rise to a non-NR reading that is weaker than the plain assertion.

(53) Bill doesn't BELIEVE it is raining, he KNOWS it is raining.

Input: $\textbf{bel}_{[+\sigma,+F]}\phi$	ExclF	MaxStrength
EXH$[\neg\textbf{bel}_{[+\sigma,+F]}\phi]$		*!
☞ $\neg$[EXH $[\textbf{bel}_{[+\sigma,+F]}\phi]]$		

(54) a. *EXH $(\neg\textbf{bel}_{[+\sigma,+F]}\phi)$

b. $\mathscr{A}lt\ (\neg\textbf{bel}_{[+\sigma,+F]}\phi) = \{\neg[\textbf{bel}\phi \vee \textbf{bel}\neg\phi], \neg\textbf{bel}\phi, \neg\textbf{know}\phi\}$

c. $\mathscr{E}xcl_\sigma\ (\neg\textbf{bel}_{[+\sigma,+F]}\phi) = \{\neg[\textbf{bel}\phi \vee \textbf{bel}\neg\phi]\}$

d. $\mathscr{E}xcl_F\ (\neg\textbf{bel}_{[+\sigma,+F]}\phi) = \emptyset$

(55) a. $\neg$EXH $(\textbf{bel}_{[+\sigma,+F]}\phi)$

b. $\mathscr{A}lt\ (\textbf{bel}_{[+\sigma,+F]}\phi) = \{\textbf{bel}\phi \vee \textbf{bel}\neg\phi, \textbf{bel}\phi, \textbf{know}\phi\}$

c. $\mathscr{E}xcl_\sigma\ (\textbf{bel}_{[+\sigma,+F]}\phi) = \emptyset$

d. $\mathscr{E}xcl_F\ (\textbf{bel}_{[+\sigma,+F]}\phi) = \{\textbf{know}\phi\}$

e. $\neg$EXH $(\textbf{bel}_{[+\sigma,+F]}\phi) = \neg[\textbf{bel}\phi \wedge \neg\textbf{know}\phi] = \neg\textbf{bel}\phi \vee \textbf{know}\phi$

4. Conclusions

In this paper, I proposed a feature-based account to explain the distributions of NR and non-NR readings. This account inherits the merits of the SI-based account from Romoli (2012, 2013), and improves this SI-based view by highlighting the role of focus in cancellations and providing restrictions on the selection of EXH-structures. I argued that NR readings come from global exhaustification, and that non-NR readings result from either double exhaustification and local exhaustification, depending on the location of focus.

(56) a. John doesn't believe that it is raining. NR

b. EXH $\neg$ [John believes$_{[+\sigma]}$ it's raining] Global EXH

(57) a. John DOESn't believe that it is raining., he isn't sure. Non-NR

b. EXH $\neg_{[+F]}$ EXH [John believes$_{[+\sigma]}$ it's raining] Double EXH

(58) a. John doesn't BELIEVE that it is raining, he knows it. Non-NR

b. $\neg$ EXH [John believes$_{[+\sigma,+F]}$ it's raining] Local EXH

I proposed that where to insert EXH is restricted by two inviolable principles (feature checking and avoiding G-triviality) and two OT constraints (ExclF and MaxStrength). Alternative structures are untenable for the following reasons. First, as for the basic negation sentence in (56), local exhaustification results in a reading that is equivalent to the assertion, violating the MaxStrength constraint. Second, as for the sentence in (57) which has narrow focus on negation, an LF with a single local exhaustification has an unchecked feature [+F], failing to satisfy the feature-checking requirement, and an LF with a single global exhaustification has to negate both positive and negative excluded middles, yielding G-triviality. Third, as for the sentence in (58) where the NRP is stressed, the focus-alternative set has no excludable member under global exhaustification.

References

Abusch, D. (2002). Lexical alternatives as a source of pragmatic presuppositions. In *Proceedings of SALT*, Volume 12, pp. 1–19.

Abusch, D. (2005). Triggering from alternative sets and projection of pragmatic presuppositions.

Abusch, D. (2010). Presupposition triggering from alternatives. *Journal of Semantics 27*(1), 37–80.

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

Chierchia, G. (2004). Scalar implicatures, polarity phenomena, and the syntax/pragmatics interface. In A. Belletti (Ed.), *Structures and beyond*, Volume 3, pp. 39–103. Oxford: Oxford University Press.

Chierchia, G. (2006). Broaden your views: Implicatures of domain widening and the logicality of language. *Linguistic inquiry 37*(4), 535–590.

Chierchia, G. (2013). *Logic in grammar: Polarity, free choice, and intervention*. Oxford University Press.

Chierchia, G., D. Fox, and B. Spector (2013). The grammatical view of scalar implicatures and the relationship between semantics and pragmatics. In C. Maienborn, K. Von Heusinger, and P. Portner (Eds.), *Semantics: An International Handbook of Natural Language Meaning*, Volume 3.

Fillmore, C. J. (1963). The position of embedding transformations in a grammar. *Word 19*, 208–231.

Fox, D. (2007). Free choice and the theory of scalar implicatures. In U. Sauerland and P. Stateva (Eds.), *Presupposition and Implicature in Compositional Semantics*, pp. 71–120. Basingstoke: Palgrave.

Fox, D. and B. Spector (2009). Economy and embedded exhaustification. *Handout from a talk at Cornell University (http://web. mit. edu/linguistics/people/faculty/fox/Fox_ Spector_Cornell. pdf)*.

Gajewski, J. R. (2005). *Neg-raising: Polarity and presupposition*. Ph. D. thesis, Massachusetts Institute of Technology.

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

Grice, P. (1975). Logic and conversation. In J. M. Peter Cole (Ed.), *Syntax and Semantics III: Speech Acts*. New York: Academic Press.

Hamblin, C. L. (1973). Questions in montague english. *Foundations of language 10*(1), 41–53.

Heim, I. (2011). Lecture notes. *Pragmatics class*.

Homer, V. (2013). Neg-raising and positive polarity: The view from modals. *Semantics and Pragmatics*.

Horn, L. (1978). Remarks on neg-raising. *Syntax and semantics 9*, 129–220.

Horn, L. R. (1989). *A natural history of negation*. University of Chicago Press.

Lakoff, R. A syntactic argument for negative transportation. In *Proceedings of CLS*, Volume 5.

Magri, G. (2009). A theory of individual-level predicates based on blind mandatory scalar implicatures. *Natural Language Semantics 17*(3), 245–297.

Marantz, A. (1984). *On the nature of grammatical relations*. Cambridge: MIT Press.

Romoli, J. (2011). The presuppositions of soft triggers aren't presuppositions. In *Proceedings of SALT*, Volume 21, pp. 236–256.

Romoli, J. (2012). *Soft but Strong: Neg-raising, Soft Triggers, and Exhaustification*. Ph. D. thesis, Harvard University.

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

Rooth, M. (1985). *Association with focus*. Ph. D. thesis, University of Massachusetts, Amherst.

Rooth, M. (1992). A theory of focus interpretation. *Natural language semantics 1*(1), 75–116.

Rooth, M. (1996). Focus. In S. Lappin (Ed.), *The Handbook of Contemporary Semantic Theory*. Blackwell Publishers.

Sauerland, U. (2004). Scalar implicatures in complex sentences. *Linguistics and philosophy 27*(3).

Schwarzschild, R. (1999). Givenness, avoidf and other constraints on the placement of accent*. *Natural language semantics 7*(2), 141–177.