ONLY: An NPI-licenser and NPI-unlicenser

Yimei Xiang – Harvard University

Abstract It is commonly argued that weak negative polarity items (NPIs) (e.g., any) can occur in any (Strawson) downward-entailing environment. This generalization, however, is challenged by Wagner’s (2006) observations with the NPI-licenser only: although an only-clause is (Strawson) downward-entailing in its unfocused part, NPIs are not necessarily licensed there. In particular, DP-only does not license an NPI that appears within the unfocused part of its left argument (as seen in *Only a chair of any HUMANITIESF department met with the president), and VP-only does not license an NPI if this NPI and the focused item appear within the same island (as seen in *The president only met with a chair of any HUMANITIESF department). These observations suggest that the licensing status of an NPI in an only-sentence is not just determined by the polarity pattern of the environment where this NPI gets interpreted.

To explain Wagner’s (2006) observations, I argue that only is not just an NPI-licenser but also an “NPI-unlicenser.” Following Chierchia (2006, 2013), I assume that an NPI carries a domain feature [D] which activates domain alternatives, and that an NPI is unlicensed if exhaustifying its domain alternatives yields a contradiction. I further propose that only can check off the [D] feature of an NPI that appears within its syntactic argument. If the argument of only is downward-entailing with respect to an NPI, using only to check off the [D] feature of this NPI would return an inference that contradicts the prejacent presupposition and make this NPI unlicensed. In the case of VP-only association, if an NPI is not focused and doesn’t appear within a focus-contained island, the contradiction can be avoided by F-movement.

1. Introduction

Negative polarity items (henceforth NPIs), such as the emphatic expression any, are known to be licensed in downward-entailing (henceforth DE) environments (Fauconnier 1975, 1979; Ladusaw 1979). Prototypical DE environments are the following:

(1) Under the semantic scope of negation
   a. John didn’t read any papers.
   b. * John read any papers.

(2) Within the scope of negative quantifiers
   a. Few/no/at most 3 students read any papers.
   b. * Many/most students read any papers.

(3) In the left argument of universal quantifiers
   a. Every student who has read any papers passed the exam.
   b. * Every student who has read some papers passed any exams.
c. * Some student who has read any papers passed the exam.

(4) In the antecedent of conditionals
   a. If John knows any big names, he will be invited.
   b. * If John is invited, he will know any big names.

An environment is DE if and only if it supports downward inferences from a superset to a subset. For instance, observe the contrast between (5a) and (5b) that a downward inference holds from a term student to a more specific term smart student in the left argument of every, but not in the left argument of some.

(5) a. Every student passed the exam. → Every smart student passed the exam.
   b. Some students passed the exam. /\ Some smart students passed the exam.

Following von Fintel (1999) and Gajewski (2007), I define DE functions and DE environments as in (6), where the arrow ‘⇒’ stands for generalized entailment.1

(6) a. **Downward entailing functions**
   A function f of type ⟨σ, τ⟩ is DE if and only if
   for all x and y of type σ such that x ⇒ y, f(y) ⇒ f(x).

b. **Downward entailing environments**
   If α is of type δ and A is a constituent that contains α, then:
   A is DE with respect to α if and only if the function λx.[A[α/v]δ]g[υδ→x] is DE.
   [A[α/v] is the result of replacing α with a trace v in A.]

---

1Generalized entailment is cross-categorically defined for items of any entailing type. Entailing types are defined recursively as in (1), adopted from Chierchia (2013: 204). Accordingly, t, ⟨e, t⟩, ⟨e, et⟩, and any type of the form ⟨… t⟩ are entailing types.

(1) **Entailing type**
   a. t is a basic entailing type.
   b. If τ is an entailing type, then for any type σ, ⟨σ, τ⟩ is an entailing type.

Following von Fintel (1999), I define generalized entailment as in (2).

(2) **Generalized entailment ‘⇒’**
   a. If ϕ, ψ are of type t, then: ϕ ⇒ ψ if and only if ϕ is false or ψ is true.
   b. If β, γ are of an entailing type ⟨σ, τ⟩, then:
      β ⇒ γ if and only if for all α such that α is of type σ: β(α) ⇒ γ(α).

The basic case (2a) is defined based on truth values: a truth-value entails another if and only if it is not the case that the first is true and the second is false. In a generalized case, as schematized in (2b), a function entails another if and only if the result of applying the first function to any argument entails the result of applying the second function to the same argument. For example, [[smart student]] and [[student]] are functions of type ⟨e, t⟩. [[smart student]] ⇒ [[student]], because for any x of type e, [[smart student]](x) ⇒ [[student]](x). All the aforementioned cases can also be understood from a set-theoretic perspective: for any two sets A and B, A ⇒ B if and only if A is a subset of B (written as ‘A ⊆ B’).
For example, the function $\lambda x.\{\text{every } A/A' \text{ is } B\}^{g[A'-\tau x]}$ is DE, and hence the sentence “every $A$ is $B$” is DE with respect to $A$. The DE analysis of NPI-licensing can now be summarized as follows:

(7) **The DE analysis of NPI-licensing**

An NPI is grammatical if and only if it appears in a constituent that is DE with respect to this NPI.

Nevertheless, the NPI-licensing effect of the exclusive focus particle *only*, first observed by Klima (1964), casts doubt on the DE analysis of NPI-licensing: while the NPI *any* is licensed in (8a) and (9a), these contexts do not seem to support downward inferences (Atlas 1993, 1996), as shown in (10a) and (10b), respectively.\(^2\) In comparison, the ungrammaticality of (8b) and (9b) shows that the NPI *any* can be licensed by the overt exclusive particle *only*, but not by a covert exhaustifier.\(^3\)

(8) Right argument of DP-*only*
   a. Only $\text{JOHN}_F$ read any papers.
   b. * $\text{JOHN}_F$ read any papers.

(9) Unfocused part under VP-*only*
   a. Mary only gave any books to $\text{JOHN}_F$.
   b. * Mary gave any books to $\text{JOHN}_F$.

(10) a. Only $\text{JOHN}_F$ ate *vegetables* for breakfast. ↠ Only $\text{JOHN}_F$ ate *kale* for breakfast.
    b. Mary only gave *fruit* to $\text{JOHN}_F$. ↠ Mary only gave *apples* to $\text{JOHN}_F$.

In responding to the NPI-licensing effect of *only*, many different positions have been defended in the literature, each of which weakens the strict DE condition to some extent for weak NPIs. For example, the Strawson-DE analysis grants the presuppositions of the consequence when a downward-inference is assessed (von Fintel 1999, Wagner 2006, Hsieh 2012). The grammatical view of NPI-licensing ignores presuppositions and implicatures when the meaning of a weak NPI is evaluated (Gajewski 2012, Chierchia 2013; see fn. 13). The pseudo-anti-additivity analysis (Atlas 1996) and the non-veridicality analysis (Giannakidou 2006) each use a weaker condition for the licensing of weak NPIs.

It is further observed that the distributional pattern of NPIs in *only*-clauses is restricted by F-association: *only* cannot license an NPI that appears in its F-associate, as exemplified in (11). For simplicity, I call this fact “Licensing Asymmetry.”

\(^2\)Here and henceforth, CAPITAL letters mark stressed items, and the subscript ‘$F$’ marks semantic focus.

\(^3\)The grammatical view of NPI-licensing (see section 3 and the references therein) assumes that focused items are always associated with overt or covert exhaustifiers. Accordingly, for example, the LF of (8b) contains a covert exhaustivity $O$-operator associated with the semantic focus $\text{JOHN}_F$. Under this view, the contrast between (8a) and (8b) thus suggests that the covert $O$-operator cannot license NPIs. This contrast is expected by the grammatical view: the $O$-operator is non-monotonic; unlike *only* which asserts only an exhaustivity inference, the $O$-operator also asserts a prejacent inference. See more details in footnote 13.
(11) a. Only [some/*any students]_F saw John.
   b. Mary only gave [some/*any books]_F to John.

Nevertheless, Drubig (1994) and Wagner (2006) observe that Licensing Asymmetry is not solely determined by F-association. In both examples in (12), only does not license the NPI any, although the F-associate of only is just the NP-complement of any.

   b. John only read [some/*any PAPERS_F], (he didn’t read any books).

Similar facts are observed in (13) and (14). In these examples, only is associated with narrow focus, and the NPI any appears outside the focus: in (13), the NP-complement of any contains narrow focus; in (14), the anyP serves as an of-argument of the focused NP CHAIR; in (15), the anyP serves as the possessor of the focused NP ADVISOR.

(13) a. Only a chair of [a/*any HUMANITIES_F department] met with the president.
   b. The president only met with a chair of [a/*any HUMANITIES_F department].

(14) a. Only [a CHAIR_F of a/*any humanities department] met with the president.
   b. The president only met with [a CHAIR_F of a/*any humanities department].

(15) a. Only [someone’s/*anyone’s ADVISOR_F] met with the president.
   b. The president only met with [someone’s/*anyone’s ADVISOR_F].

Examples (12) to (15) suggest two generalizations. First, DP-only does not license an NPI that appears within its left argument, regardless of whether this NPI is part of its associated focus. Second, VP-only does not license an NPI if this NPI and the focused item appear within the same island. In examples from (12) to (15), I used ‘[●]’ to enclose the minimal island that contains the semantic focus.\(^4\) I call these facts uniformly “Licensed-Nonfocal Mismatch”, meaning that the part of an only-clause where an NPI can appear does not fully match up with the part that is not F-associated with only.\(^5\)

In sum, to capture the NPI-licensing effect of only, we need to answer at least three questions:

\(^4\)Determiner phrases like anyP can be considered islands, to the extent that the NP complement of a determiner cannot be moved out alone. According to Abels (2003), the complement of a phasal head – such as the D head any – cannot be moved by itself and strand its embedding phasal head, but must pied-pipe that phasal head.

\(^5\)There are some seeming counter-examples to this generalization. For example, the NPI any is licensed in (1a), although it appears within the left argument of DP-only. Interestingly, replacing the plural candidates with the singular existential indefinite a candidate makes any not licensed, as seen in (1b). Such a contrast is also observed with VP-only, as shown in (2). I thank Brian Buccola, Gary Thoms, and an anonymous reviewer for the observation.

(1) a. Only [candidates with some/*any relation to NY_F] met with the pope.
   b. Only [a candidate with some/*any relation to NY_F] met with the pope.

(2) a. The pope only met with [candidates with some/*any relation to NY_F].
   b. The pope only met with [a candidate with some/*any relation to NY_F].
(A) *Only* is not a prototypical DE-operator; why does it license NPIs?

(B) Why is it that the NPI-licensing effect of *only* is subject to Licensing Asymmetry?

(C) Why is it that Licensing Asymmetry is subject to Licensed-Nonfocal Mismatch?

The rest of this paper is organized as follows. Section 2 and section 3 will review two representative theories on the NPI-licensing effect of *only*. One is the focus (F)-movement theory by Wagner (2006), and the other is the grammatical (G-)view of NPI-licensing by Chierchia (2006, 2013). I will show that both analyses have clear advantages but neither of them properly predicts the distributional pattern of NPIs under *only*. In section 4, I will propose a hybrid analysis that incorporates features of both theories. The main ingredients of the new analysis are the following:

(A) Following the G-view of NPI-licensing, I assume that an NPI is licensed if and only if assessing its domain feature [D] does not yield a G-triviality.

(B) I argue that *only* is not only an NPI-licenser, but also an “NPI-unlicenser”: if an NPI is interpreted within the syntactic complement of *only* and if the prejacent of *only* is not downward-entailing with respect to this NPI, using *only* to assess the [D] feature of this NPI returns an inference contradicting the prejacent presupposition of *only*.

(C) In the case of VP-*only* association, the requirement of avoiding G-trivialities motivates F-movement. The semantic contradiction in (B) can be avoided if and only if the NPI is interpreted in the remnant VP.

2. The theory of F-movement and its shortcomings

2.1. The Strawson DE-condition

To capture the NPI-licensing effect of *only*, von Fintel (1999) proposes a Strawson Downward Entailing (Strawson DE henceforth) analysis of NPI-licensing, as summarized in (16). Unlike the strict DE condition, the Strawson DE condition grants all presuppositions of the consequent sentence when the validity of a downward inference is assessed.

(16) The Strawson DE analysis of NPI-licensing

I argue that the licensing of *any* in (1a) and (2a) is irrelevant to *only*; instead, it correlates to the fact in (3) that an NPI can be licensed when it appears within a *with*-adjunct or a relative clause of a plural item.

(3) a. (The) candidates with some/any relation to NY\(_F\) met with the pope.

b. A candidate with some/*/any relation to NY\(_F\) met with the pope.

c. (The) candidates who had some/any relation to NY\(_F\) met with the pope.

d. A candidate who had some/*/any relation to NY\(_F\) met with the pope.
a. A function $f$ of type $\langle \sigma, \tau \rangle$ is Strawson DE if and only if for all $x$ and $y$ of type $\sigma$ such that $x \Rightarrow y$ and $f(x)$ is defined: $f(y) \Rightarrow f(x)$.

b. An NPI is only grammatical if it is in the scope of a function $f$ such that $f$ is Strawson DE.

Further, von Fintel (1999) argues that “only+NP” is a Strawson DE function: only presupposes the truth of its propositional prejacent (Horn 1969); the scope of “only+NP” is DE when the prejacent presupposition of only is satisfied, as shown in (17).

(17) Kale is a vegetable. 
John ate kale for breakfast. 
Only JOHN$_F$ ate vegetables for breakfast. 
$\therefore$ Only JOHN$_F$ ate kale for breakfast

2.2. The theory of F-movement

The Strawson DE condition, however, still cannot explain why VP-only association is subject to Licensing Asymmetry and why Licensing Asymmetry in general exhibits Licensed-Nonfocal Mismatch. In responding to these questions, Wagner (2006) proposes a theory of F-movement. This theory has components in both syntax and semantics, as outlined in the following.

2.2.1. The LF syntax of only

First, both DP-only and VP-only have two syntactic arguments, a restrictor and a scope. In the case of DP-only, the restrictor and the scope correspond to its left argument and right argument, respectively. In the case of VP-only, F-association always invokes covert phrasal movement of the focused expression to the syntactic restrictor of only; thus the restrictor and the scope correspond to the F-moved phrase and the remnant VP, respectively.

(18) a. DP-only
Only JOHN$_F$ came.

b. VP-only
Mary only invited JOHN$_F$.

Second, F-movement is island-sensitive (compare Anderson 1972, Jackendoff 1972, Rooth 1985); therefore, when only is associated into an island, the F-moved phrase ought to be the minimal island that contains the focus-associate of only (see also Drubig 1994). For instance, the F-moved item in (19a) ought to be the entire complex DP, and the one in (19b) ought to be the entire when-clause.
(19)  a. Dr. Smith only rejected [the proposal that \(JOHN_F\) submitted].  
     b. Dr. Smith only complains [when \(BILL_F\) leaves the lights on].

2.2.2. The semantics of only

Wagner (2006) defines the lexical entry of *only* as in (20). *Only* is a two-place predicate; it asserts an exhaustivity inference and presupposes an existential premise. In (20), the arguments \(\alpha\) and \(P\) correspond to the syntactic restrictor and the scope of *only*, respectively. The variable \(C\) stands for the exhaustification domain; its value is determined by both linguistic and contextual factors.

(20)  a. \([\text{only}](\alpha)(P) = \forall a \in C [P(a) \to \neg\text{\(\alpha\)} \subseteq \neg P(a)]\)  
     b. Presupposition: \(\exists x. P(x)\)

Note that the presupposition (20b) is unconventional. It clearly differs from the standard prejacent presupposition (i.e., \(P(\alpha)\)) assumed by Horn (1969). It is also weaker than the existential presupposition assumed by Horn (1996), which is generated by abstracting over the semantic focus, not the entire syntactic restrictor. See (21) for a simple comparison of these three presuppositions. Underlining marks the syntactic restrictor of *only*.

(21) Only a female inhabitant of [Twin Earth]\(_F\) met Particle Man.  
     a. A female inhabitant of Twin Earth met Particle Man. Prejacent (Horn 1969)  
     b. A female inhabitant from somewhere met Particle Man. Existential (Horn 1996)  
     c. Someone met Particle Man. Existential (Wagner 2006)

The main purpose for Wagner to weaken the presupposition of *only* is to capture the phenomenon of Licensed-Nonfocal Mismatch. If we assume the prejacent presupposition (21a), we would predict an unfocused part of the syntactic restrictor of *only* to be Strawson DE. Compare the following two lines of reasoning for illustration. In both lines of reasoning, the premise grants the prejacent inference of the consequent *only*-sentence. In (22), the downward inference from the focused NP *student\(_F\)* to its subset *junior student\(_F\)* does not hold, while in (23), the downward inference from the unfocused item *inhabitant* to its subset *female inhabitant* does hold.

(22) A junior student is a student. \(x \Rightarrow y\)  
     A junior student met Particle Man. \(f(x)\) is defined  
     Only a *student\(_F\)* met Particle Man. \(f(y)\)  
     \(\not\Rightarrow\) Only a *junior student\(_F\)* met Particle Man. \(\not\Rightarrow f(x)\)

(23) A female inhabitant is an inhabitant. \(x \Rightarrow y\)  
     A female inhabitant of Twin Earth met Particle Man. \(f(x)\) is defined  
     Only an *inhabitant* of [Twin Earth]\(_F\) met Particle Man. \(f(y)\)  
     \(\therefore\) Only a *female inhabitant* of [Twin Earth]\(_F\) met Particle Man \(\therefore f(x)\)
The same problem arises if we adopt Horn’s (1996) existential presupposition (21b). For instance, in (24), if the premise grants the truth of this presupposition, a downward inference holds from the unfocused item inhabitant to its subset female inhabitant.

(24) A female inhabitant is an inhabitant.  \[ x \Rightarrow y \]
A female inhabitant from somewhere met Particle Man.  \[ f(x) \text{ is defined} \]
Only an inhabitant of [Twin Earth]\_F met Particle Man.  \[ f(y) \]
\[ \therefore \text{Only a female inhabitant of [Twin Earth]\_F met Particle Man} \]
\[ \therefore f(x) \]

Hence, if assuming that only triggers Horn’s (1969) prejacent presupposition or Horn’s (1996) existential presupposition, the Strawson DE analysis of NPI-licensing would predict that the unfocused part of the restrictor can license NPIs, contra fact in (25).

(25) * Only any inhabitant of [Twin Earth]\_F met Particle Man.

To solve this problem, Wagner assumes that only presuppose just a weak existential presupposition, which abstracts over the entire syntactic restrictor, and argues that the syntactic restrictor of only is not Strawson DE. For instance, in (26), the downward inference from the unfocused item inhabitant to female inhabitant seems to be invalid.

(26) A female inhabitant is an inhabitant.  \[ x \Rightarrow y \]
Someone met Particle Man.  \[ f(x) \text{ is defined} \]
Only an inhabitant of [Twin Earth]\_F met Particle Man.  \[ f(y) \]
\[ \not\therefore \text{Only a female inhabitant of [Twin Earth]\_F met Particle Man} \]
\[ \not\therefore f(x) \]

2.2.3. Motivating F-movement

Wagner (2006) assumes that F-movement is triggered to strengthen the existential presupposition of only, in the spirit of the Maximise Presupposition (MP) Principle:

(27) **Maximise Presupposition Principle** (Heim 1991)

Out of two sentences which are presuppositional alternatives and which are contextually equivalent, the one with the stronger presuppositions must be used if its presuppositions are met in the context.

This idea brings up Wagner’s assumptions on syntax and semantics: “F-movement minimizes the size of the syntactic restrictor, which may have an effect on the strength of the statement that is grammatically encoded by the sentence.” Consider the basic only-sentence in (28) for illustration. F-movement changes the size of the syntactic complement of only, as marked by underlining: the syntactic complement of only encloses the entire VP “played BASKETBALL\_F” in (28a) but just the semantics focus “BASKETBALL\_F” in (28b). Next, the existential presupposition triggered by only is obtained by abstracting over the entire VP in (28a), while over the focused noun phrase in (28b). Hence, the only-sentence obtains a stronger existential presupposition based on the F-moved form (28b). Finally, in the spirit of the MP Principle, the F-moved form is more preferable than the non-F-moved form.
(28) John only played BASKETBALL$_F$.
   a. John only played BASKETBALL$_F$. [Without F-movement]
      Presupposition: $\exists x.\text{John }x$-ed.
   b. John only played BASKETBALL$_F$. [With F-movement]
      Presupposition: $\exists x.\text{John played }x$.

2.2.4. Consequences

Wagner (2006) claims that the theory of F-movement is followed by two general restrictions on the NPI-licensing effect of VP-$only$, namely, the Island Restriction and the Head Restriction, as described in the following:

(29) **Head Restriction**
   If $only$ is associated with the head of a constituent, it does not license an NPI in the complement of the head.

(30) **Island Restriction**
   Association with a constituent within an island cannot license an NPI that appears in the same island.

The Head Restriction comes from the standard view that F-movement is phrasal movement: a head cannot take phrasal movement; therefore, when $only$ is associated with a head, the F-moved item has to be the entire projection of this head. For instance, in (31), the F-moved element has to be the entire VP, including the anyP; therefore, the NPI any is interpreted within the syntactic restrictor of $only$ and is not licensed.

(31) * John only [CUT$_F$ any vegetables].

Wagner (2006: 310)

The Island Restriction follows Wagner’s assumption that F-movement is sensitive to islands. For instance, in (32a), $only$ is associated into the because-clause, which is an island; therefore, the F-moved item must be the entire because-clause, which contains an NPI any. In (32b), in contrast, the NPI anyone is outside the because-clause and therefore is not involved in F-movement.\(^6\)

(32) a. * Mary only gave a book to John [because BILL$_F$ gave any book to him].
   b. She only gave anything to anyone [because YOU$_F$ did].

(Wagner 2006: 313)

\(^6\)The island effect of Licensing Asymmetry under VP-$only$ might not be the real reason (or at least not the only reason) why the NPI any is not licensed in (32a). Chierchia (2013) indicates that because has an intervening effect on the licensing of NPIs. In (1b), even clause-mate negation cannot license the NPI anybody across because.

(1) a. Mary doesn’t believe that John criticized anybody.
   b. * Mary doesn’t believe that John was arrested because he criticized anybody.
2.3. Problems with Wagner (2006)

Wagner (2006) has convincingly argued that F-movement is needed for predicting the NPI-licensing effect of *only*, especially for unifying the Licensed-Nonfocal Mismatch behaviors of DP-*only* and VP-*only*. DP-*only* does not license an NPI that appears in its left argument, even if it is not associated with this NPI. Analogously, VP-*only* does not license an NPI if this NPI and the focused item appear within the same island, regardless of whether this NPI itself is focused. By virtue of F-movement, whatever explanation that works for DP-*only* will also work for VP-*only*.

Nevertheless, Wagner’s explanation of Licensed-Nonfocal Mismatch is infeasible: contrary to what he claims, with the weak existential presupposition (20b), an *only*-clause actually is Strawson DE with respect to any unfocused position (section 2.3.1). Moreover, in Wagner’s account, the F-movement operation is not well-motivated or properly controlled (section 2.3.3 and 2.3.3).

2.3.1. Problems with Licensed-Nonfocal Mismatch

Recall that Wagner weakens the presupposition of *only*, so as to avoid overly predicting an unfocused position in the syntactic restrictor of *only* to be Strawson DE. This attempt of eliminating Strawson DE-ness, however, is infeasible. The seeming failure of licensing a Strawson-downward inference in (26), repeated below, involves confusions between natural language semantics and meta-language interpretations.

\[
\begin{align*}
\text{A female inhabitant is an inhabitant.} & \quad x \Rightarrow y \\
\text{Someone met Particle Man.} & \quad f(x) \text{ is defined} \\
\text{Only an inhabitant of } \text{Twin Earth} \text{ met Particle Man.} & \quad f(y) \\
\text{Only a female inhabitant of } \text{Twin Earth} \text{ met Particle Man} & \quad ? \not\Rightarrow f(x)
\end{align*}
\]

To be more concrete, I suggest that the reason why the above reasoning is invalid is that the actual presupposition of the consequent *only*-sentence is stronger than what has been provided in the premise. The actual presupposition might be Horn’s (1969) prejacent presupposition, Horn’s (1996) existential presupposition, or Ippolito’s (2008) conditional presupposition (see (42)), as described in the following:

\[
\begin{align*}
\text{“Only a female inhabitant of } \text{Twin Earth} \text{ met Particle Man” presupposes at least one of} \\
\text{the following:} & \\
a. \text{A female inhabitant of Twin Earth met Particle Man. } & \text{[Prejacent]} \\
b. \text{A female inhabitant of some planet met Particle Man. } & \text{[Regular existential]} \\
c. \text{If a female inhabitant of some planet met Particle Man, then a female inhabitant of Twin Earth met Particle Man. } & \text{[Conditional]}
\end{align*}
\]

To see whether Wagner’s definition of *only* in (20) predicts an unfocused position of the restrictor to be Strawson DE, we should compare the predicted interpretations of the two *only*-sentences in (33), as shown in the following:

\[
\begin{align*}
\text{a. Only an inhabitant of } \text{Twin Earth} \text{ met Particle Man.} & \\
\text{b. Only a female inhabitant of } \text{Twin Earth} \text{ met Particle Man.} & \text{[Condition]}
\end{align*}
\]
From the interpretations above, we can see that Wagner’s account actually predicts the sentence “only an α of [Twin Earth]F met Particle Man” to be Strawson DE as well as strictly DE with respect to α. It is Strawson DE because the downward inference holds in the asserted component (namely, (35a-i) entails (35b-i)). It is also strictly DE because (35a) and (35b) have exactly the same presupposition. Hence, Wagner cannot capture the phenomenon of Licensed-Nonfocal Mismatch in (25), no matter whether he follows the Strawson DE analysis or the DE analysis.

This problem also extends to the case of VP-only association. As the interpretations in (36) show, using the definition of only in (20), Wagner would predict that the sentence “John only CUTF x” is Strawson DE as well as DE with respect to the object position x. It is Strawson DE because a downward inference holds in the asserted component (i.e., (36a-i) entails (36b-i)). It is also strictly DE because the two only-sentences, (36a) and (36b), have the same presupposition. Hence, Wagner cannot predict the Head Restriction in (31).

(36)  a. John only CUTF vegetables.
   i. Assertion: If John did any action to vegetables, that action is no more than cutting.
      ii. Presupposition: John did something.
   b. John only CUTF kale.
      i. Assertion: If John did any action to kale, that action is no more than cutting.
      ii. Presupposition: John did something.

In sum, the weak existential presupposition of only cannot help with explaining the phenomenon of Licensed-Nonfocal Mismatch. Wagner’s account predicts that an only-clause is Strawson DE as well as DE with respect to any unfocused position.

2.3.2. Problems with the presupposition of only and “why move?”

Recall that Wagner uses the MP Principle to motivate F-movement: the presupposition of only is obtained by abstracting the syntactic restrictor; F-movement reduces the size of the syntactic restrictor and hence strengthens the existential presupposition. In this section, I will show that

\footnote{Wagner (2006: 310) himself writes the presupposition of (36b) as “John did something with kale”, which however conflicts with his main assumption that the existential import abstracts over the entire complement of only.}
only triggers either a prejacent presupposition or a conditional presupposition, the involvement of which makes the MP-based analysis of motivating F-movement untenable.

Wagner has noticed that the definition of only in (20) is too weak to account for sentences like (37), where only is associated with a conjunction. Intuitively, the sentence in (37) cannot be uttered if Sue invited only John, although the asserted exhaustivity inference in (37a) and the assumed existential presupposition in (37b) are both satisfied in this context. To predict this infelicity, Wagner adopts McCawley’s (1993: 311) idea and proposes that the sentence (37) conversationally implies the prejacent inference in (37c). This implicature is derived roughly as follows. If the speaker believes that Sue didn’t invite John or believes that Sue didn’t invite Mary, she could have been more informative by leaving the person out of the list. Next, assuming the speaker is being cooperative in the Gricean sense and is well-informed, we conclude that the speaker believes that Sue invited both John and Mary.

(37) Sue only invited [John and Mary]_{F}.
   a. Assertion: Sue didn’t invite anyone who is neither John nor Mary.
   b. Presupposition: Sue invited someone.
   c. Conversational implicature: Sue invited both John and Mary.

As argued by Ippolito (2008: 59), however, the implicature analysis of the prejacent inference has at least two problems. First, it cannot explain why a negative only-sentence also introduces a prejacent inference, as in (38). If the speaker believes that Sue invited Mary but not John, saying “Sue invited someone who is not John” is less informative than the assumed exhaustive assertion “Sue invited someone who is neither John nor Mary.” Hence, we are not able to derive the expected prejacent inference of (38) based on the conversational maxims.

(38) Sue not only invited [John and Mary]_{F}

Second, the implicature analysis cannot explain why the prejacent inference cannot be cancelled without an epistemic operator, as shown in (39).

(39) a. # Only Mary can speak French – in fact, not even she can.
   b. Only Mary can speak French, and maybe not even she can.

Related to the second problem pointed out by Ippolito, the contrast between the following two conversations also suggests that the prejacent inference from only is more robust than an implicature from the weak scalar item some: unlike the scalar implicature from some, the prejacent inference affects the truth conditions of the only-clause and is much more difficult to cancel.

(40) Did John invite some of the speakers to the dinner?
   a. Yes. Actually he invited all of them.
   b. # No. He invited all of them.

(41) Did John only invite Mary?
a. # Yes. Actually he didn’t invite anybody.
b. No. He didn’t invite anybody.

Hence, the prejacent inference should not be treated as an independent implicature. Instead, we can analyze it as a presupposition (Horn 1969; Rooth 1985, 1992) or adopt Ippolito’s (2008) analysis that the prejacent inference is a logical consequence of the scalar implicature (42a) together with a conditional presupposition (42b).

(42) Only $A$ is $B$.
   a. Implicature: Someone is $B$.
      (Derived by negating the stronger alternative No one is $B$.)
   b. Presupposition: If someone is $B$, $A$ is $B$.
   c. $\Rightarrow A$ is $B$.

In this section, I don’t take a position as to which analysis correctly characterizes the presupposition(s) of only, but just show that the MP-based analysis of motivating F-movement is infeasible whichever analysis of presuppositions we follow. Under the presupposition analysis, the existential presupposition – regardless of its strength – is always entailed by and hence collapses under the prejacent presupposition, and thus the MP Principle cannot motivate F-movement. Alternatively, with Ippolito’s (2008) conditional presupposition, the MP Principle would predict that the un-F-moved form is more preferable over the F-moved form: the existential inference is the antecedent of the conditional presupposition; therefore, strengthening the existential inference would weaken the presupposition. To be more concrete, consider the example (28) again, repeated below. If following Ippolito’s (2008) analysis, we would have conditional presuppositions as follows. Observe that the presupposition in (43b) from the F-moved form is semantically weaker than the one in (43a) triggered in absence of F-movement.

(43) John only played $\underline{BASKETBALL}_F$.
   a. John only played $\underline{BASKETBALL}_F$. [Without F-movement]
      Presupposition: If $\exists x$ [John $x$-ed], then John played basketball.
   b. John only played $\underline{BASKETBALL}_F$. [With F-movement]
      Presupposition: If $\exists x$ [John played $x$], then John played basketball.

2.3.3. Problems with “when move?”

Wagner (2006) claims that F-movement is mandatory for VP-only association. I argue that this strong claim leads to two undesired consequences.

First, it incorrectly predicts that an NPI associated with only is always unlicensed, even if this NPI appears under some other licenser. Recall that Wagner’s assumptions yield the prediction that only does not license an NPI appearing inside the F-moved constituent. This prediction, together with the claim that F-movement is mandatory for VP-only association, implies a stricter constraint as follows: “VP-only cannot be associated with any NPIs or with/into any NPI-contained islands
within which the NPIs are not licensed.” This constraint, however, is too strong for cases like (44), where only is associated with an anyP across another NPI-licenser (i.e., clause-mate negation).

(44) Mary only didn’t give [any books]$_F$ to John. (She did her best to help him.)

The stricter constraint predicts that the sentence (44) has to take the LF in (45). In this LF, however, the NPI any cannot be licensed: the anyP, as the minimal F-contained island, is moved to the syntactic restrictor of only, a context that is non-Strawson DE and cannot license NPIs. \(^{89}\)

(45)

\[\begin{array}{c}
\text{only} \\
\text{DP}_i \\
\text{any books} \\
\text{Mary didn’t gave } t_i \text{ to John}
\end{array}\]

Second, as pointed out by Büring and Sharvit and reported by Wagner (2006: fn. 20), assuming F-movement to be mandatory overly rules out possible interpretations for sentences like (46), where only is associated with a scopal element at most 3 students across another scopal element want. If F-movement were mandatory, the focused item would have to take only a wide scope reading.

\[^{8}\]Note that the NPI any is not licensed if only merely associates with any, as exemplified below. This is so because (1) doesn’t have any excludable alternatives, failing to satisfy the additive presupposition of only.

(1) How much books is such that Mary didn’t give to John?

*Mary only didn’t give ANY$_F$ books to John.

Consider the dialogue in (2) for illustration of the additive presupposition. The which-question with a restricted domain ensures that the exhaustification domain of only contains exactly three members, namely, I will invite John, I will invite Mary, and I will invite both John and Mary. The answer (2b) is infelicitous because the prejacent is the strongest one among the alternatives, which therefore makes the additive presupposition of only unsatisfied. In contrast, the answer (2c) is fully acceptable, because a covert exhaustification does not have an additive presupposition.

(2) Which of John and Mary will you invite?

a. Only JOHN$_F$, (not Mary/ not both).

b. * Only BOTH$_F$.

c. BOTH$_F$.

Likewise in (1), the exhaustification domain contains only propositions of the following two forms: (i) it is not the case that John read some paper in $D’$ ($D’ \subseteq D$); (ii) it is not the case that John read most/all/... the papers in $D$. All of these alternatives are entailed by the prejacent and are not excludable. For more discussions on the additive presupposition, see Klinedinst (2005), Beaver and Clark (2009), among others.

\[^{9}\]An anonymous reviewer of Journal of Semantics points out the possibility of moving the entire negative VP complement to the restrictor of only. This movement is in general possible, but is disallowed under Wagner’s framework: to enable this movement, we need to stipulate that the licensing of an NPI occurs prior to the satisfaction of the MP Principle; but Wagner assumes that the failure of licensing an NPI is due to F-movement, an operation motivated by the MP Principle.
(46) She only wanted to kiss [at most 3 students]F.

To account for the correct scope readings, Wagner (2006: fn. 20) proposes that the F-moved item undergoes obligatory semantic or syntactic reconstruction. Nevertheless, if reconstruction were always feasible, then the focused anyP in (47) should also be reconstructed and licensed within the scope of VP-only, contra fact.

(47) * Mary only wanted to read [any books]F.

A defender of the reconstruction analysis might argue that only must be reconstructed in company with the F-moved phrase. Under this assumption, (47) would be interpreted as ‘*Mary wanted to read only [any books]F.’ Nevertheless, this assumption is incorrect, because VP-only takes rigid scope reading (Taglicht & Randolph 1984, Rooth 1985, Bayer 1996), as exemplified in (48).

(48) John is only required to meet MARYF. (OK only > required; #required > only)

3. The G-view of exhaustifications

3.1. The G-view of scalar implicatures

The grammatical (G-)view (Fox 2007, Chierchia et al. 2012, among others) was first introduced to analyze scalar implicatures. This view argues that the generation of scalar implicatures is not purely due to pragmatics (compare Grice 1975), given the fact that scalar implicatures can be generated in embedding contexts.

The main ingredients of the G-view are summarized as follows. First, propositions with scalar items are associated with sets of alternatives, which are computed in the same way as answer sets of questions (Hamblin 1973) and alternative sets of focus (Rooth 1985, 1992, 1996). A recursive definition of alternative sets is schematized as follows:

(49) Alternative sets

a. For any lexical entry α, ALT(α) =
   i. \{[α]\} if α is lexical and does not belong to a scale;
   ii. \{[α₁], ..., [αₙ]\} if α is lexical and part of a scale \{[α₁], ..., [αₙ]\};

   Where ALT is a function from expressions to a set of interpretations.

b. ALT(β(α)) = \{b(a) : b ∈ ALT(β), a ∈ ALT(α)\}

Next, alternatives keep growing until factored into meaning via a covert exhaustivity operator O (or written as “EXH”). This O-operator affirms the prejacent and negates all the alternatives that are not entailed by the prejacent, as schematized in (50). The non-entailed alternatives are also called *excludable* alternatives.

---

10In this formula, the propositional letter p is sloppily used as a syntactic expression in ALT(p). A stricter definition of the O-operator is as follows, where S is the sentential complement of O.

---
\[
O(p) = \lambda w[p(w) = 1 \land \forall q \in ALT(p)[p \not\subseteq q \rightarrow q(w) = 0]]
\]

Accordingly, a scalar implicature is derived as a logical consequence of applying an \(O\)-operator over a sentence containing a scalar item. For instance, in (51), applying an \(O\)-operator over the bare \emph{some}-sentence \(\phi_{\text{some}}\) affirms the prejacent \(\phi_{\text{some}}\) and negates the stronger scalar alternative \(\phi_{\text{all}}\), yielding the scalar implicature \(\neg\phi_{\text{all}}\).

(51)  
a. Some of the students came. \(\rightarrow\) Not all of the students came.
\[\text{b. ALT}(\phi_{\text{some}}) = \{\phi_{\text{some}}, \phi_{\text{all}}\}\]
\[\text{c. } O(\phi_{\text{some}}) = \phi_{\text{some}} \land \neg\phi_{\text{all}}\]

3.2. The G-view of NPI-licensing

Chierchia (2006, 2013) extends the G-view to NPI-licensing with assumptions compatible with the strict DE condition. He proposes that the NPI \emph{any} is an indefinite existential item like \emph{some} but is lexically encoded with a grammatical feature \([D]\). This feature obligatorily activates a set of domain (D)-alternatives and must be checked off by a c-commanding \(O_D\)-operator.

Exercising an \(O_D\)-operator over a sentence containing an occurrence of \emph{any} has consequences in both syntax and semantics. In syntax, it checks off the \([D]\) feature in the lexicon of \emph{any}, just like a regular feature-checking operation. In semantics, it affirms the assertion of the prejacent and negates all the excludable D-alternatives, namely, the D-alternatives that are not entailed by the assertion of the prejacent.

A schematic example for the total domain \(D\) and its corresponding D-alternative sets is given in (52). The D-alternative set includes the prejacent, while the proper D-alternative set does not.

(52)  
a. Total-D: \{a, b\} \quad \text{Assertion} = \exists x \in \{a, b\} f(x)
\[\text{b. Sub-D: } \{a, b\}, \{a\}, \{b\} \quad \text{D-ALT} = \{\exists x \in \{a, b\} f(x), \exists x \in \{a\} f(x), \exists x \in \{b\} f(x)\}\]
\[\text{c. Proper sub-D: } \{a\}, \{b\} \quad \text{Proper D-ALT} = \{\exists x \in \{a\} f(x), \exists x \in \{b\} f(x)\}\]

Consider the basic positive sentence (53) to see how the G-view captures the DE condition of NPI-licensing. In LF (53a), an \(O_D\)-operator is mandatorily present so as to check off the \([D]\) feature of \(\text{any}_D\). The meaning of this LF proceeds as follows. First, the prejacent clause of \(O_D\) (labeled ‘S’) asserts an existential inference (53b) and is associated with a set of D-alternatives (53c). These D-alternatives are activated by the \([D]\) feature of \(\text{any}_D\) and are generated by substituting the total domain variable \(D\) with a subdomain \(D’\). Next, applying \(O_D\) affirms the prejacent \([S]\) and negates the excludable D-alternatives. Here, since the prejacent \(S\) is upward-entailing (henceforth UE) with respect to the domain variable \(D\), the proper D-alternatives are all stronger than \([S]\) and are all excludable. Hence, applying \(O_D\) yields the exhaustivity inference in (53d). Crucially, the

(1)  
\[\text{[O(S)]} = \lambda w[[S](w) = 1 \land \forall \phi \in ALT(S)[[S] \not\subseteq \phi \rightarrow \phi(w) = 0]]\]

Note that this \(O\)-operator is distinct from the one used by Fox (2007), which negates only alternatives that can be negated consistently (viz., the so-called “innocently excludable alternatives”).
exhaustivity inference in (53d) contradicts the prejacent inference \[ S \]. Such a contradiction makes the sentence (53) ungrammatical and the NPI *any* unlicensed.

(53) * John read any papers.
   a. \( OD [S \text{ John read any}_D \text{ papers}] \)
   b. \([S] = \exists x \in D[P(x) \land R(j,x)]\) (John read some papers in the total domain \( D \).)
   c. \( \text{D-ALT}(S) = \{ \exists x \in D'[P(x) \land R(j,x)] \mid D' \subseteq D \} \)
   d. \( \forall D'[D' \subseteq D \rightarrow \neg \exists x \in D'[P(x) \land R(j,x)]\) (for any proper subdomain \( D' \), John read no paper in \( D' \).)
   e. \([OD(S)] = [S] \land (53d) = \bot\) (# John read some papers in \( D \), but for any proper subdomain \( D' \), he read no paper in \( D' \).)

The contradiction in (53e) is in essence different from the one in (54). The former makes an utterance ungrammatical, while the latter makes an utterance infelicitous but not ungrammatical.

(54) # It is raining and it isn’t raining.

To distinguish these two types of contradictions, Chierchia (2013: 49) adopts notions from Gajewski (2002) and describes the contradiction in (53e) as a “Grammatical (G-)triviality”. G-triviality is a special form of Logical (L-)triviality: L-trivialities are tautologies or contradictions in the traditional sense; G-triviality refers to a characteristic that a sentence receives the same truth value (true or false) regardless of how the lexical terminals in the structure are replaced. Compare the sentences in (55) for illustration. Expressions like *John*, *smokes*, and *student* are lexical terminals, and the rest are functional terminals. The contradiction in (55a) can be avoided by substituting the two occurrences of *smoke* with distinct lexical items (as in *John smokes and doesn’t dance*.). In contrast, the meaning of (55b) is always contradictory no matter which lexical items are used. Therefore, we identify (55a) as L-trivial, while (55b) as both L-trivial and G-trivial.

---

11Consider the mini model in (1) for a simpler illustration of this idea. Assume that the total domain \( D \) contains exactly two items, paper \( p_1 \) and paper \( p_2 \). The D-alternative set is thus schematized as in (1b), consisting of three elements: the asserted proposition *John read a paper in \{p_1, p_2\}* and two proper D-alternatives including *John read a paper in \{p_1\}* and *John read a paper in \{p_2\}* . The proper D-alternatives are not entailed by the assertion. Therefore, applying an \( OD \)-operator affirms the assertion and negates both proper D-alternatives, as schematized in (1c), yielding a contradictory inference that *John read \( p_1 \) or \( p_2 \), and he did not read \( p_1 \), and he did not read \( p_2 \).*

(1) a. \( D = \{p_1, p_2\} \)
   b. \( \text{D-ALT} = \{R(j,p_1) \lor R(j,p_2), R(j,p_1), R(j,p_2)\} \)
   c. \( R(j,p_1) \lor R(j,p_2) \land \neg R(j,p_1) \land \neg R(j,p_2) = \bot\)

12Following von Fintel (1993), Gajewski (2002) assumes the truth conditions in (1) for the logical skeleton of (55b), where the backslash ‘\’ stands for subtraction. These truth conditions say that \( \{x\} \) is the unique minimal set whose subtraction from the restrictor of the quantifier (i.e., \( P \)) makes the quantification true.
(55)  a. # John smokes and doesn’t smoke. [ x P and not P ]
   b. * Some student but John smokes. [ some P but x Q ]

Next, Chierchia argues that the type of contradiction in (53e) can be avoided if the constituent that the $O_D$-operator attaches to is DE with respect to the $D$ variable of any $D$. Consider the basic negative sentence in (56) for instance. By virtue of negation, all the D-alternatives are entailed by the assertion and therefore not excludable. The $O_D$-operator, although mandatorily present for the sake of feature checking, has no effect on semantics.

(56) John didn’t read any papers.
   a. $O_D [S$ not [John read any$_D$ papers]]
   b. $[S] = \neg \exists x \in D[P(x) \land R(j, x)]$
      (John read no paper in the total domain $D$.)
   c. $D$-ALT$(S) = \{- \exists x \in D'[P(x) \land R(j, x)] \mid D' \subseteq D\}$
   d. $[O_D(S)] = [S] = \neg \exists x \in D[P(x) \land R(j, x)]$
      (John read no paper in the total domain $D$.)

3.3. Extending the G-view of NPI-licensing to only

Inspired by Krifka (1995) and Lahiri (1998), Chierchia (2006, 2013) extends the G-view of NPI-licensing to the licenser only. These approaches adopt the lexical entry of only from Horn (1969), namely, that only asserts an exhaustivity inference and presupposes the truth of its prejacent. The heart of this view is the following: the unfocused part of the asserted exhaustivity inference is DE and hence forms an NPI-licensing environment.

Using Chierchia’s (2013) schematic notations, we assume the LF in (57a). This LF contains two exhaustification operators, $O_D$ and only. In particular, $O_D$ checks off the [D] feature of the NPI any$_D$, while only checks off the [F] feature of the focused item JOHN$_F$. The prejacent presupposition and the asserted exhaustivity inference of the only-clause are schematized as in (57b) and (57c), respectively. The D-alternatives are generated from the assertion by replacing the total domain $D$ with a subdomain $D'$, as schematized in (57d).

(57) Only JOHN$_F$ read any papers.

\[ \begin{align*}
(1) \quad & \left[ \text{some } P_{(x,t)} \text{ but } x, \neg Q_{(x,t)} \right] = 1 \text{ if and only if } \\
& \text{a. some}(P \setminus \{x\})(Q) = 1; \\
& \text{b. } \forall S_{(x,t)} \left[ \text{some}(P, S)(Q) = 1 \rightarrow x \in S \right] \\
\end{align*} \]

Because “some $P$ $Q$” is UE with respect to the restrictor $P$, condition (1a) entails condition (2). Nevertheless, because $x \notin \emptyset$, (2) contradicts condition (1b). Hence, any sentence of the form “some $P$ but $x$ $Q$” is a contradiction.

\[ \begin{align*}
(2) \quad & \text{some}(P \setminus \emptyset)(Q) = 1, \text{ or equivalently, some}(P)(Q) = 1 \\
\end{align*} \]

For the purposes of this paper, it is enough to know that G-triviality is a type of L-triviality that affects grammaticality.
a. $\mathcal{O}_D [S \text{ only } [\text{JOHN}_F \text{ read any}_D \text{ papers }]]$

b. Presupposition of $S$: $\exists x \in D [P(x) \land R(j, x)]$

   (John read some papers in the total domain $D$.)

c. Assertion of $S$: $\forall y \in D_e \ [\exists x \in D [P_w(x) \land R_w(y, x)] \rightarrow y \leq j]$

   (For any individual $y$, if $y$ read some papers in the total domain $D$, then $y$ is John.)

d. $D$-\text{ALT}(S)$ = $\{[[\text{only } [\text{JOHN}_F \text{ read any}_D' \text{ papers }]] : D' \subseteq D \}

   = \{\lambda w. \forall y \in D_e [\exists x \in D' [P_w(x) \land R_w(y, x)] \rightarrow y \leq j] | D' \subseteq D\}$

The prejacent presupposition (57b), as argued by Gajewski (2011) and extended by Chierchia (2013), is irrelevant for assessing the [D] feature of the weak NPI any. Moreover, the asserted

\[\text{[(1)]\hfill *Only JOHN came in years.}
\]

a. John came in years. $\land \forall x [x \text{ came in years } \rightarrow x \leq j]$

b. John came in $D$. $\land \forall x [x \text{ came in } D \rightarrow x \leq j]$ where $D \subseteq \text{YEARS}$

This idea also explains the lack of NPI-licensing effect of the covert $O$-operator, as seen in (8b) and (9b). In the following, I show that both syntactically well-formed LFs of (2) yield a G-triviality. Under the LF in (2a), according to the definition of $O$ in (50), the prejacent inference $[[S]]$ is asserted by the covert $O_F$-operator, as schematized in (2a-ii). Thus, when the $O_D$-operator is applied, the prejacent inference will also enter into the assessment of the [D] feature, contrary to the case of the only-sentence (57). Crucially, the prejacent inference (the underlined part) and the exhaustivity inference (the rest) are UE and DE with respect to the NPI any, respectively; thus overall the complement of $O_D$ is non-monotonic with respect to the NPI any. Alternatively, under the LF in (2b), the $[F]$ and $[D]$ features will be checked off simultaneously, generating a set of alternatives via point-wise functional application (Hamblin 1973), as schematized in (2bii). Applying $O$ affirms the prejacent and negates all the alternatives except the prejacent itself, yielding a contradiction.

\[\text{[(2)]\hfill *JOHN}_F \text{ read any papers.}
\]

a. $O_D [O_F [S \text{ JOHN}_F, \text{ read any}_D \text{ papers }]]$

   i. $[[S]] = \lambda w. \exists y \in D [P_w(y) \land R_w(j, y)]$

   (John read a paper in the total domain $D$.)

   ii. $[O_F(S)] = \lambda w. \exists y \in D [P_w(y) \land R_w(j, y)] \land \forall x \in D_e [\exists y \in D' [P_w(y) \land R_w(x, y)] \rightarrow x \leq j]$

   (John and no one else read a paper in the total domain $D$.)

   iii. $[O_D(O_F(S))] = \bot$

b. $O_{D,F} [S \text{ JOHN}_F, \text{ read any}_D \text{ papers }][S]$

   i. $[[S]] = \lambda w. \exists y \in D [P_w(y) \land R_w(j, y)]$

   ii. $\text{ALT}(S) = \{\lambda w. \exists y \in D' [P_w(y) \land R_w(x, y)] | D' \subseteq D, x \in D_e\}$

   iii. $[O_{D,F}(S)] = \bot$
component (57c) is DE with respect to the domain variable D. Therefore, any is licensed in (57), as it would be in any DE environments.

It is worthy noting that the G-view conventionally assumes that exhaustification operators operate on propositional alternatives. Therefore, for both DP-only and VP-only, this convention requires the exhaustification domain to be a set of propositions, represented as “ALT(p)” in (58).

(58) \[\text{only}(p) = \lambda w[p(w) = 1. \forall q \in \text{ALT}(p)[q(w) = 1 \rightarrow p \subseteq q]]\]

Following this convention strictly, the G-view of NPI-licensing would instead schematize the assertion of S as in (59). Here the F-alternatives are propositions of the form “y read some papers in the total domain D” where y is a contextually relevant individual.

(59) \[\lambda w. \forall q \in \text{F-ALT}(p)[q(w) = 1 \rightarrow \lambda w. \exists x \in D[P_w(x) \land R_w(j, x)] \subseteq q]\]

where \(\text{F-ALT}(p) = \{\lambda w. \exists x \in D[P_w(x) \land R_w(y, x)] \mid y \in D_e\}\)

This move does not change the polarity pattern of the assertion with respect to the D variable. It might not be easy to see the DE-ness of (59) with respect to D, but it is easy to prove it from the semantically equivalent formula (60), which is the conjunction of negated excludable F-alternatives: an F-alternative q is UE with respect to D, and thus its negation is DE with respect to D, and thus the conjunction of negated F-alternatives is DE with respect to D.\(^{14}\)

(60) \[\bigcap\{\lambda w. \neg q(w) : q \in \text{F-ALT}(p) \land \lambda w. \exists x \in D[P_w(x) \land R_w(j, x)] \not\subseteq q\},\]

where \(\text{F-ALT}(p) = \{\lambda w. \exists x \in D[P_w(x) \land R_w(y, x)] \mid y \in D_e\}\)

The G-view of NPI-licensing can be summarized as follows: the [D] feature of an NPI evokes the mandatory application of an OD-operator; if the constituent that OD attaches to is non-DE with respect to this NPI, employing OD yields a G-triviality. As for the case of only, the G-view shows that the asserted component of an only-clause is DE with respect to its unfocused part, which therefore can license NPIs.

### 3.4. Advantages of the G-view

Superior to previous studies on NPI-licensing, the G-view explicitly explains why NPIs cannot appear in non-DE environments. Moreover, this view is compatible with the strict DE condition and therefore is free from the problems that the Strawson DE analysis of NPI-licensing faces.

As pointed out by Lahiri (1998) and Gajewski (2011), for example, the Strawson DE condition cannot account for the distributional pattern of NPIs under a definite description of the form “the+NP\(_{\text{singular}}\)” or “both+NP”: the left argument is Strawson DE but cannot license NPIs. Given these problematic cases, Wagner (2006) admits that other conditions, such as not being Strawson-UE (Lahiri 1998, Cable 2002, Guerzoni & Sharvit 2007), are also required for licensing NPIs.

\(^{14}\)I thank an anonymous reviewer of Journal of Semantics for pointing out a mistake in an older version of this article (in proceedings of Sinn und Bedeutung 19), where I claimed that (59) is non-DE with respect to the variable D.
(61) * The student who had any linguistics did well.
   a. Presupposition: \(|\text{student}'_w| = 1\)
   b. Assertion: \(\text{student}'_w \subseteq \text{did-well}'_w\)

(62) * Both students who had any linguistics did well.
   a. Presupposition: \(|\text{students}'_w| = 2\)
   b. Assertion: \(\text{student}'_w \subseteq \text{did-well}'_w\)

Following the G-view, Gajewski (2011) provides a simpler explanation of the distributional pattern of NPIs under definite descriptions. He proposes that the cardinality inferences in (61a) and (62a) are not only presupposed but also asserted, as exemplified below. Hence the overall assertions of the sentences (61-62) are non-monotonic with respect to any, which therefore explains why the NPI any is not licensed.

(63) Assertion of (61) due to Gajewski (2011):
\[|\text{student}'_w| = 1 \land [\text{student}'_w \subseteq \text{did-well}'_w]\]

### 3.5. Problems with the G-view

Previous works on the G-view have not yet discussed the phenomenon of Licensing Asymmetry. In this section, I will first show that the current G-view is indeed sufficient for analyzing the basic cases in (64), where only is associated with any or anyP. But this explanation is highly restricted; it cannot extend to cases like (65), where only is associated with any across an existentially quantificational expression, nor to cases like (66) that are subject to Licensed-Nonfocal Mismatch.

(64) a. * John only read ANY\(_F\) papers.
   b. * John only read [any PAPERS]\(_F\), (he didn’t read every book).

(65) a. * Mary only invited some students who read ANY\(_F\) papers.
   b. * Only some students who read ANY\(_F\) papers passed the test.

(66) a. * John only read any PAPERS\(_F\).
   b. * Only any BOYS\(_F\) arrived.

#### 3.5.1. A potential solution for Licensing Asymmetry

If we follow the G-view of NPI-licensing strictly, we would structure the LF of (64a) as (67a): a covert \(O_D\)-operator embeds the entire only-clause; the grammatical features \([F]\) and \([D]\) agree with only and \(O_D\), respectively. Following Chierchia’s (2006, 2013) assumption that the polarity item any is a variant of the existential indefinite some, we conjecture that its default F-alternatives are simply the scalar alternatives, as schematized in (67b). The semantics of (67a) thus proceeds as follows: first, applying only negates the stronger scalar/focus alternative, yielding the assertion (67c), which is UE with respect to the \(D\) variable; next, exercising \(O_D\) to check off the \([D]\) feature
leads to the semantic consequence of negating the proper D-alternatives, yielding a G-triviality. Crucially, (67e) contradicts the asserted scalar inference (67c), yielding a G-triviality.

(67) * John only read ANY \( F \) papers.
   a. \( O_D[S_2 \text{ only } [s_1 \text{ John read ANY } D, F \text{ papers}]] \)
   b. \( F\text{-ALT}(S1) = \left\{ \begin{array}{l} \exists x \in D[P(x) \land R(j,x)] \quad \text{(John read some papers in } D) \end{array} \right\} \)
   c. Assertion of S2: \( \neg \forall x \in D[P(x) \to R(j,x)] \)
      (John didn’t read all the papers in the total domain \( D \))
   d. D-ALT(S2) = \{ \neg \forall x \in D'[P(x) \to R(j,x)] : D' \subseteq D \}
   e. \( \forall D'[D' \subset D \to \forall x \in D'[P(x) \to R(j,x)]] \)
      (For each proper subdomain \( D' \), John read all the papers in \( D' \).)
   f. (67e) contradicts (67c). For example, let \( D = \{a, b\} \), then
      \( D\text{-ALT}(S2) = \{ \neg[R(j,a) \land R(j,b)], \neg R(j,a), \neg R(j,b) \} \)
      (67c) = \( \neg[R(j,a) \land R(j,b)] \)  (not [John read both \( a \) and \( b \)])
      (67e) = \( R(j,a) \land R(j,b) \)  (John read both \( a \) and \( b \))

3.5.2. Problem 1: Licensing Asymmetry with existentials

Nevertheless, the solution used above is highly restricted. If \( only \) is associated with \( any \) across an existential quantifier, the inference derived by negating the proper D-alternatives might do not contradict the asserted scalar inference.

Consider (65a) for illustration. Due to the Complex NP Constraint, the NPI \( any \) must be interpreted within the relative clause of the existentially quantificational phrase \( some \ students \). The following derivation shows that the meaning of the LF in (68a) is not contradictory.

(68) * Mary only invited some students who read ANY \( F \) papers.
   a. \( O_D[S_2 \text{ only } [s_1 \text{ Mary invited some students who read ANY } F \text{ papers}]] \)
   b. \( F\text{-ALT}(S1) = \left\{ \begin{array}{l} \exists x[S(x) \land \forall y \in D[P(x) \land R(x,y)] \land I(m,x)] \\
      \quad \text{(Mary invited some students who read some papers in } D) \end{array} \right\} \)
   c. Assertion of S2: \( \neg \exists x[S(x) \land \forall y \in D[P(x) \to R(x,y)] \land I(m,x)] \)
      (Mary didn’t invite any students who read all the papers in \( D \))
   d. D-ALT(S2) = \{ \neg \exists x[S(x) \land \forall y \in D'[P(x) \to R(x,y)] \land I(m,x) ] : D' \subseteq D \}
   e. \( \forall D' \subset D \exists x[S(x) \land \forall y \in D'[P(x) \to R(x,y)] \land I(m,x)] \)
      (For every proper subdomain \( D' \), Mary invited some students who read all the papers in \( D' \).)
f. (68c) does not contradict (68e). For example, let $D = \{a, b\}$, then:

$\neg \exists x[S(x) \land R(x, a \oplus b) \land I(m, x)]$

(Mary didn’t invite any students who read both $a$ and $b$.)

$\exists x[S(x) \land R(x, a) \land I(m, x)] \land \exists x[S(x) \land R(x, b) \land I(m, x)]$

(Mary invited some students who read $a$ and some students who read $b$.)

To understand this problem from a more general perspective, recall how the G-view explains the NPI-licensing effect of negation:

Claim 1: Assessing the [D] feature of any $D$ over negation is semantically vacuous.

Claim 2: In absence of negation, assessing the [D] feature of any $D$ yields a G-triviality.

Claim 1 can easily extend to other DE environments, such as only, as we have seen in section 3.3. This is so because the D-alternatives will be non-excludable as long as the [D] feature is assessed in a DE context. Nevertheless, Claim 2 does not necessarily hold in other non-DE environments, such as the one seen in (68): even though the proper D-alternatives are excludable, negating them does not necessarily yield a G-triviality.

3.5.3. Problem 2: Licensed-Nonfocal Mismatch

We have seen from section 3.3 that the asserted component of an only-sentence is DE with respect to any unfocused position of this sentence. Since the G-view takes the DEness of the asserted component as the only requirement for licensing weak NPIs, it cannot capture the phenomenon of Licensed-Nonfocal Mismatch.

For instance, following the G-view, we would assume that the sentence in (66a) takes the LF in (69a). Applying only negates all the excludable F-alternatives and returns an assertion that is DE with respect to the $D$ variable. Next, since the D-alternatives are all entailed by the assertion, using $O_D$ to check off the [D] feature of the NPI any does not yield a G-triviality. Therefore, if the LF in (69a) were possible, the NPI any should have been licensed in (66a), contra fact.

(69) * John only read any PAPERS$_F$.

a. $O_D [S_2 \text{ only } [S_1 \text{ John read any}_D \text{ PAPERS}_F]]$

b. F-ALT(S1) = \{ \exists x \in D[P(x) \land R(j, x)] \} (John read some papers in $D$.)

\{ \exists x \in D[B(x) \land R(j, x)] \} (John read some books in $D$.)

c. Assertion of S2: $\neg \exists x \in D[B(x) \land R(j, x)]$ (John didn’t read any books in $D$.)

d. D-ALT(S2) = \{ \neg \exists x \in D'[B(x) \land R(j, x)] : D' \subseteq D \}

e. $[O_D(S2)] = (69c) = \neg \exists x \in D[B(x) \land R(j, x)]$ (John didn’t read any books in $D$.)

4. A new analysis

We have seen that the F-movement operation is needed to unify the distributional patterns of NPIs under DP-only and VP-only, especially for the facts related to Licensed-Nonfocal Mismatch.
But, this operation has not yet been well-motivated. A natural move would be to incorporate F-movement into the G-view of NPI-licensing and motivate F-movement based on the requirement of avoiding G-trivialities.

Why is it that a logical inference can motivate a syntactic operation? Chierchia (2013: 444) indicates that the structure-building apparatus (e.g., Merge, Move, Agree) and the inferential one are not radically different: “grammar only sees functional/logical material; logic sees functional/logical material and whether the lexical material is the same or different.” G-triviality, in particular, is the type of L-triviality that takes effect in grammar.

Nevertheless, we have also seen that neither Wagner (2006) nor Chierchia (2006, 2013) can properly capture the distributional pattern of NPIs under only. Both theories predict that NPIs can appear in any Strawson DE or DE environment. And both theories predict that the asserted component of an only-sentence is DE with respect to any unfocused position. The phenomenon of Licensed-Nonfocal Mismatch, however, suggests that the licensing status of an NPI in an only-sentence is not just determined by the polarity pattern of the environment where this NPI gets interpreted. For instance, compare the minimal pair in (70): despite the fact that the asserted component of the entire only-sentence is DE with respect to any in (70a) and UE with respect to any in (70b), both occurrences of any are not licensed. We can conclude that a pure monotonicity-based account cannot account for the distributional pattern of NPIs under only.

(70)  

a. * Only any BOYS\textsubscript{F} arrived.  
b. * Only ANY\textsubscript{F} boys arrived.

I argue that only is not just an NPI-licenser but also an “NPI-unlicenser.” It is an NPI-licenser because it creates a DE environment in its unfocused part. It is also an NPI-unlicenser because interpreting an NPI within its syntactic complement makes this NPI not licensed. In the G-view of NPI-licensing, an NPI is unlicensed if and only if the assessment of its [D] feature yields a G-triviality. To this extent, we can say that the O\textsubscript{D}-operator can trigger an NPI-unlicensing effect, which occurs when O\textsubscript{D} is applied to a constituent that is non-DE with respect to an NPI.

To capture NPI-unlicensing effect of only, I propose that only, just like the covert O\textsubscript{D}-operator, can check off the grammatical feature [D]. A more general assumption is stated as follows:

(71)  

a. Only can check off any alternative-generating features (e.g., [F], [D]).  
b. An alternative-generating feature that always takes “+” value (e.g., the [D] feature of any) must agree with the closest probe.

The locality constraint (71b) is exclusive to alternative-generating features that always take “+” value.\textsuperscript{15} When an alternative-generating feature takes “+” value, it activates the corresponding

\textsuperscript{15}For instance, it is clearly not the case that a focused element has to be associated with the closest focus-sensitive operator. In (1), the second focused item BOBBY is not associated with the closest focus-sensitive particle only.

(1)  

We only\textsubscript{1} recovered the diary entries that MARILYN\textsubscript{F1} made about John.
alternatives; otherwise it does not. In Chierchia’s (2013) system, the domain feature [D] and the scalar feature [σ] of polarity items (including NPIs, free choice items (FCIs), and minimizers), if syntactically visible, always take “+” value, obligatorily activating D-alternatives and scalar alternatives. In contrast, the [σ] feature of a regular scalar item (e.g., some) can take either “+” or “–” value, and hence the activation of the corresponding scalar alternatives is optional.

According to the assumptions in (71), contra the traditional G-view, the occurrence of DP-only in (70a) checks off not only the [F] feature of the focused item BOYS but also the [D] feature of the NPI any, as illustrated in (72b). There is thus no need to posit a covert OD-operator in the LF. In particular, the feature-checking operation on [D], as we will see in section 4.1.3, gives rise to a G-triviality and “unlicenses” the NPI any: only negates the proper D-alternatives, returning an assertion that contradicts the prejacent presupposition of only.

(72) *Only any BOYSF arrived.
   a. Traditional G-view
   b. New analysis

In the case of VP-only association, F-movement is motivated so as to avoid this G-triviality. The following lists out all the syntactically well-formed LFs for the sentence in (73). If the focused item JOHN is interpreted in situ, as in (73a), the [D] feature of any will be checked off by the closest probe, which can be the particle only, as in (73a-i), or a local OD-operator, as in (73a-ii). Both of these LFs yield a G-triviality and thus cannot make any licensed. Conversely, if the focused item undergoes F-movement as in (73b), only c-commands the focused item but not the NPI any; therefore, the [D] feature of any is left to be checked off by a global covert OD-operator. Moreover, the scope of only creates a DE environment; therefore, using OD to assess the [D] feature does not yield a G-triviality, making the NPI any licensed.

We also recovered [the diary entries that MARILYN made about BOBBY].
(73) Mary only gave any books to JOHN\(_F\).

a. #Without F-movement

   (i) Without \(O_D\)

   \[
   \text{only} \quad \text{VP} \\
   \text{Mary gave any}\_D \text{ books to JOHNF} \\
   \]

   (ii) With \(O_D\)

   \[
   \text{only} \quad O_D \quad \text{VP} \\
   \text{Mary gave any}\_D \text{ books to JOHNF} \\
   \]

b. \(\text{OK}\) With F-movement

\[
\text{OD} \quad \text{only} \quad \text{JOHNF} \quad \text{VP} \\
\text{Mary gave any}\_D \text{ books to } t_i \\
\]

4.1. DP-*only* association

4.1.1. Semantics of DP-*only*

I define the semantics of DP-*only* cross-categorically as in (74), where \(\alpha\) and \(P\) stand for the syntactic restrictor (i.e., the left argument) and the scope (i.e., the right argument), respectively. The letter \(\tau\) stands for an arbitrary semantic type. This definition follows the convention initiated by Horn (1969): *only* presupposes the truth of the prejacent and asserts an exhaustivity inference.\(^{16}\)

(74) **Semantics of DP-*only***

\[
\begin{align*}
\text{a. } [\text{only}(\alpha_\tau)(P_{(\tau,\tau)})] &= \lambda w. \forall a \in [\alpha]^{f,d}[[P](a)(w) = 1 \rightarrow [P][[\alpha]^0] \subseteq [P](a)] \\
\text{b. Presupposition: } [P][[\alpha]^0]
\end{align*}
\]

Following the Roothean convention of Alternative Semantics (Rooth 1985, 1992, 1996), I use \([\alpha]^0\) and \([\alpha]^{f,d}\) to represent the *ordinary value* and the *focus-and-domain value* of \(\alpha\), respectively. The

\(^{16}\)For the proposed analysis, all we need is that *only* obligatorily triggers a prejacent inference; it isn’t crucial how this prejacent inference is generated. It could be an independent presupposition, as assumes in (74b). It could also be derived as in (42), where a scalar implicature together with Ippolito’s (2008) conditional presupposition entails a prejacent inference. Here I go for the former option for simplicity.
When an NPI appears, its [D] feature will be assessed by a covert $O_D$-operator that embeds the entire *only*-clause (labeled “S”). The asserted component of the *only*-clause is DE with respect to the scope. Therefore, assessing the [D] feature of *any* via an $O_D$ is

17There are also other presuppositions triggered by "only" (such as the additive presupposition, see fn. 8). I ignore these presuppositions because they are not crucial for the purpose of this paper.
semantically vacuous and would not yield a G-triviality. An example with schematized derivation is given in (77). The application of \( O_D \) is semantically vacuous, because all the D-alternatives (i.e., members in \([S]^d\)) are entailed by the asserted component of the only-sentence (i.e., \([S]^0\)). The only difference between this analysis from the traditional G-view is that here only c-commands the focused item \( \text{JOHN}_F \) but not the the NPI \textit{any}.

\[(77) \text{Only } \text{JOHN}_F \text{ read any papers.}\]

\[
\begin{array}{c}
\text{JOHN}_F \\
\downarrow \\
\text{read any } D \text{ papers}
\end{array}
\]

- a. \([\text{JOHN}_F]^0 = j\)
- b. \([\text{JOHN}_F]^f = D_e\)
- c. \([\textit{read any papers}] = \lambda y\lambda w. \exists x \in D[P_w(x) \land R_w(y, x)]\)
- d. \([S]^0 = \lambda w. \forall y \in [\text{JOHN}_F]^f[[\textit{read any papers}](y)(w) \rightarrow [[\textit{read any papers}]([[\text{JOHN}_F]^0]) \subseteq [[\textit{read any papers}](y)]

\[
\begin{align*}
&= \lambda w. \forall y \in D_e[\exists x \in D[P_w(x) \land R_w(y, x)] \rightarrow \\
&\quad \lambda w. \exists x \in D[P_w(x) \land R_w(j, x)] \subseteq \lambda w. \exists x \in D[P_w(x) \land R_w(y, x)]
\end{align*}
\]

(Only John read any papers in the total domain \(D.)\)
- e. \([S]^d = \{\lambda w. \forall y \in D_e[\exists x \in D'[P_w(x) \land R_w(y, x)] \rightarrow y \leq j] : D' \subseteq D\}\)
- f. \([O_D(S)] = \lambda w [[S]^0](w) = 1 \land \forall p \in [S]^d[p(w) = 1 \rightarrow [S]^0 \subseteq p]

\[
\begin{align*}
&= [S]^0
\end{align*}
\]

(Only John read any papers in the total domain \(D.)\)
- g. \([S]^0\) entails every member in \([S]^d\).

For example, let \(D = \{a, b\}\) and \([\text{JOHN}_F]^f = \{j, m\}\), then:

\[
\begin{align*}
[\textit{read any papers}] &= \lambda x[R(x, a) \lor R(x, b)] \\
[S]^0 &= \neg[R(m, a) \lor R(m, b)] \\
[S]^d &= \{\neg[R(m, a) \lor R(m, b)], \neg R(m, a), \neg R(m, b)\}
\end{align*}
\]

Hence, \([O_D(S)] = [S]^0 = \neg[R(m, a) \lor R(m, b)]\) (Mary read neither \(a\) nor \(b\)).

4.1.3. The NPI-unlicenser use of DP-only

When an NPI \textit{any} appears within the left argument of DP-only, \textit{only} checks off both \([F]\) and \([D]\) features, and therefore its exhaustification domain consists of not only the focus value of its left
argument but also the domain-value of its left argument. In the asserted component, applying *only* negates all the excludable alternatives, including the proper D-alternatives, yielding the inference in (78d). Crucially, (78d) contradicts the prejacent presupposition of *only*, as in (78e). Hence, the meaning of (78) is always undefined, which therefore explains why the NPI *any* is not licensed in this sentence.

(78) *Only anyD BOYSF left.

\[
S \quad \leftarrow \text{left} \\
\quad \downarrow \text{only} \\
\quad \downarrow \text{anyD BOYSF}
\]

a. \([\text{anyD BOYSF}]^0 = \lambda_f \exists x \in D[B(x) \land f(x)]\)
b. \([\text{anyD BOYSF}]^f = \{\lambda_f \exists x \in D[g(x) \land f(x)] : g \in D_{(e, st)}\}\)
c. \([\text{anyD BOYSF}]^d = \{\lambda_f \exists x \in D'[B(x) \land f(x)] : D' \subseteq D\}\)
d. \([S] = \lambda_w \forall \mathcal{P} \in [\text{anyD BOYSF}]^f \cdot \exists x \in D[B(x) \land f(x)] \rightarrow [\text{anyD BOYSF}]^0([\text{left}]) \subseteq \mathcal{P}([\text{left}])\]

e. \[\lambda_w \forall \mathcal{P} \in [\text{anyD BOYSF}]^d \cdot \exists x \in D'[B(x) \land f(x)] \rightarrow [\text{anyD BOYSF}]^0([\text{left}]) \subseteq \mathcal{P}([\text{left}])\]

(For any proper subdomain D', no boy in D' left.)
f. Presupposition of S: \(\exists x \in D[B(x) \land L(x)]\) (Some boys in the total domain D left.)
g. (78e) contradicts (78f). For example, let D = \{a, b\}, then:

(78e) = \neg L(a) \land \neg L(b) \quad \text{(Neither a nor b left.)}

(78f) = L(a) \lor L(b) \quad \text{(a or b left.)}

4.2. VP-*only* association

4.2.1. Semantics of VP-*only*

In the case of VP-*only* association, if focus is interpreted in situ, the alternatives are all propositional. Following Rooth (1985, 1992), I schematize the semantics of VP-*only* as in (79). The letter S stands for the VP complement of only.

(79) **Semantics of VP-*only* (without F-movement)**

a. \([\text{only}(S)] = \lambda_w \forall q \in [S]^{f \cdot d} \cdot q(w) = 1 \rightarrow [S]^0 \subseteq q\]
b. Presupposition: \([S]^0\)
When F-movement takes place, VP-*only* is defined cross-categorically the same as DP-*only*. Now the two arguments $\alpha$ and $P$ correspond to the F-moved phrase and the remnant VP, respectively.

(80) **Semantics of VP-only (with F-movement)**

a. $[\text{only}(\alpha_t)(P_{(x,s)})] = \lambda w. \forall a \in [\alpha]^f \cdot [[P](a)](w) = 1 \rightarrow [P](\alpha^0) \subseteq [P](a)$

b. Presupposition: $[P](\alpha^0)$

In the absence of NPIs, F-movement is not motivated and therefore the exhaustification domain of *only* would be a set of propositions. A schematized example is given in the following.

(81) Sue only invited JOHN F’s advisors.

Let $A_w(x) = \exists y [\text{advisor}_w'(y, x) \land \forall z [\text{advisor}_w'(z, x) \rightarrow z \leq y]]$

(a) $[[\text{Sue invited JOHN F’s advisors}}] = \lambda w. I_w(s, A_w(j))$

(b) $[[\text{JOHN F}}] = D_e$

(c) $[\text{Sue invited JOHN F’s advisors}] = \{ \lambda w. I_w(s, A_w(x)) \mid x \in D_e \}$

(d) Assertion: $\lambda w. \forall q \in \{ \lambda w. I_w(s, A_w(x)) \mid x \in D_e \}[q(w) = 1 \rightarrow \lambda w. I_w(s, A_w(j)) \subseteq q]$

(Any true proposition of the form “Sue invited x’s advisors” is entailed by the prejacent that “Sue invited John’s advisors.”)

e. Prejacent presupposition: $\lambda w. I_w(s, A_w(x))$

### 4.2.2. Case 1: F-movement is not motivated

Assuming that F-movement is used to avoid G-trivialities, we conjecture that interpreting focus in situ is allowed as long as it does not yield a G-triviality. For instance, in (82), the NPI *any* can be licensed by negation: assessing the [D] feature of *any* by the overt particle *only* or a covert $O_D$-operator that immediately c-commands negation does not yield a G-triviality; therefore, F-movement is not motivated in (82).

(82) Mary only didn’t give any$_D$ books to JOHN$_F$

Only $[(O_D) \textbf{not} [\text{Mary gave any$_D$ books to JOHN$_F$}]]$

Assuming F-movement to be conditional better controls the use of this operation. For instance, contrary to Wagner’s (2006) predictions (seen in section 2.3.3), the new analysis predicts that F-movement is not motivated in (83) and (84), because interpreting focus in situ does not yield a G-triviality: (83) has no NPI, while the NPI *any* in (84) is licensed by negation.

(83) She only wanted to kiss [at most 3 students]$_F$.

(84) Mary only didn’t give [any books]$_F$ to John. She did her best to help him.
4.2.3. Case 2: F-movement is motivated

In the case of DP-only association, we have seen that interpreting an NPI within the immediate syntactic complement of only yields a G-triviality, making the NPI unlicensed. In the case of VP-only association, interpreting focus in situ would also cause a G-triviality, but this G-triviality can be avoided by F-movement under certain configurations.

Let us first see why the two un-F-moved forms in (73a), repeated below, yield a G-triviality.

(85) Mary only gave any books to JOHN. (Without F-movement)

a. Without OD

In the LF (85a), only immediately c-commands the entire prejacent VP (labeled “S”) and therefore checks off the [F] feature of JOHN as well as the [D] feature of anyD. Hence, the exhaustification domain of only consists of not only F-alternatives but also D-alternatives, as schematized in (86a). Applying only negates all the excludable alternatives of S and yields the exhaustivity assertion (86b), which further entails the inference (86c). Crucially, (86c) contradicts the prejacent presupposition (86d), as illustrated in (86e). Hence, the value of the LF in (85a) is always untrue.

(86) Interpreting the LF (85a)

[only [S Mary gave anyD books to JOHN]]

a. Exhaustification domain of only: ALT(S) = F-ALT(S) ∪ D-ALT(S)

i. F-ALT(S) = {∃x ∈ D[B(x) ∧ G(m,y,x)] | y ∈ De}

ii. D-ALT(S) = {∃x ∈ D'[B(x) ∧ G(m,j,x)] | D' ⊆ D}

b. Assertion: λw.∀q ∈ ALT(S)[q(w) = 1 → ∃x ∈ D[B(x) ∧ G(m,j,x)] ⊆ q]

↓

c. ∀D'[D' ⊆ D → ¬∃x ∈ D'[B(x) ∧ G(m,j,x)]]

(Mary didn’t give John any books in any proper subdomain D'.)

Alternatively, we can also assume a larger exhaustification domain for only, namely, the set of propositions of the form “Mary gave some books in D’ to y”, as schematized in (1). This alternative set is point-wise composed from the D-alternatives of anyD and the focus alternatives of JOHN. Since this set includes all the D-alternatives in (86a-ii), exhaustifying over this set also yields an assertion that entails the inference in (86c).

(1) ALT(S) = {∃x ∈ D'[B(x) ∧ G(m,y,x)] | y ∈ De, D' ⊆ D}
d. Presupposition: \( \exists x \in D [B(x) \land G(m, j, x)] \)
(Mary gave John some books in the total domain \( D \).)

e. (86c) contradicts (86d). For example, let \( D = \{a, b\} \), then:
\[
(86c) = \neg G(m, j, a) \land \neg G(m, j, b) \quad \text{(Mary didn’t gave John } a \text{ or } b.)
\]
\[
(86d) = G(m, j, a) \lor G(m, j, b) \quad \text{(Mary gave John } a \text{ or } b.)
\]

In the LF (85b), the [D] feature of any \( D \) is checked off by a local \( O_D \)-operator. From (87), it is easy to see that the prejacent of only, namely ‘\( O_D(S) \)’, is a contradiction: \( S \) is UE with respect to the \( D \) variable; affirming \( [S] \) and negating all the proper D-alternatives of \( S \) yields a contradiction. With a contradictory prejacent presupposition, the meaning of the LF in (85b) is always undefined.

(87) **Interpreting the LF (85b)**

[only \( [O_D [S \text{ Mary gave any}_D \text{ books to JOHN}_F]]] \]

a. \([S] = \exists x \in D [B(x) \land G(m, y, x)]\)

b. \(D-\text{ALT}(S) = \{\exists x \in D'[B(x) \land G(m, j, x)] \mid D' \subseteq D\}\)

c. Presupposition: \([O_D(S)] = \lambda w[[S](w) = 1 \land \forall p \in D-\text{ALT}(S)[[S] \not\subseteq p \rightarrow p(w) = 0]]\)

\(\vdash\)

For example, let \( D = \{a, b\} \), then:
\[
[S] = G(m, j, a) \lor G(m, j, b)
\]
\(D-\text{ALT}(S) = \{G(m, j, a), G(m, j, b), G(m, j, a) \lor G(m, j, b)\}\)
\([O_D(S)] = [G(m, j, a) \lor G(m, j, b)] \land \neg G(m, j, a) \land \neg G(m, j, b) = \bot\)

Alternatively, under the F-moved LF in (73b), repeated below, only does not c-command any, and hence the [D] feature of any\( D \) is assessed by a covert \( O_D \)-operator. As seen in section 3.3, this LF does not suffer G-triviality: only creates a DE environment with respect to the \( D \) variable in its scope; therefore, in this LF, assessing [D] by the \( O_D \)-operator is semantically vacuous.

(88) Mary only gave any books to JOHN\( F \). (With F-movement)

In conclusion, in the case of VP-only association, F-movement is motivated if only c-commands an NPI and this NPI is not licensed within the prejacent of only. Note that the discussion above is irrelevant to overt F-movement. Since G-triviality is assessed at LF, the requirement of avoiding G-trivialities can only motivate covert movement. Hence, this requirement is not the cause of overt F-movement in languages like Hungarian and Basque. (See Kiss 1995 for more instances of languages with overt F-movement.)
4.2.4. Case 3: F-movement is unhelpful

Let us move on to Licensing Asymmetry: *only* cannot license an NPI if this NPI appears within its F-associate or an island that contains its F-associate. Relevant examples mentioned in the previous sections are collected in (89). I will show that NPIs are not licensed in these sentences because their G-trivialities cannot be salvaged by F-movement.

(89)  
- a. * John read only ANY$_F$ papers.  
- c. * John read only any PAPERS$_F$, (he didn’t read any books).  
- d. * The president only met with a chair of any HUMANITIES$_F$ department.  
- e. * The president only met with a CHAIR$_F$ of any humanities department.  
- f. * Mary only gave a book to John [because BILL$_F$ gave any book to him].  
- g. * The president only met with [*anyone’s ADVISOR$_F$].  
- h. * Mary only invited some students who read ANY$_F$ papers.

Take (89c) for example, where *only* is directly associated with the NP complement of *any*. From (90), it is easy to see that all the possible LFs of (89c), with and without F-movement, yield a G-trivial meaning. The two un-F-moved forms in (90a) are interpreted in the same way as the two forms in (85). The F-moved form in (90b) is interpreted in the same way as the one in (78), where an NPI appears within the restrictor of DP-*only*). Crucially, F-movement is subject to island constraints, and hence the F-moved item has to be the entire any$_P$.

(90)  
*John read only any PAPERS$_F$, (he didn’t read any books).

a. # Without F-movement

![Diagram without F-movement]

or

b. # With F-movement

![Diagram with F-movement]
Moreover, under the new analysis, as long as the prejacent of only is non-DE with respect to an NPI, a contradiction arises between the prejacent presupposition and the negation of the proper D-alternatives. This generalization also applies to the case where only is associated with an NPI across an existential quantifier. For illustration, compare the following semantic derivation with the one in (68), which follows the traditional G-view.

(91) *Mary only invited some students who read ANY_F papers.
   a. only [S Mary invited some students who read ANY_F,D papers]
   b. Exhaustification domain of only: ALT(S) = F-ALT(S) ∪ D-ALT(S)
      i. F-ALT(S) = \{\exists x[S(x) \land \exists y \in D[P(x) \land R(x,y)] \land I(m,x)]\}
         (Mary invited some students who read some papers in D.)
      ii. D-ALT(S) = \{\exists y[S(x) \land \exists y \in D'[P(x) \land R(x,y)] \land I(m,x)] : D' \subseteq D\}
   c. Assertion: \lambda w.\forall q \in ALT(S)[q(w) \rightarrow \exists x[S(x) \land \exists y \in D[P(x) \land R(x,y)] \land I(m,x)] \subseteq q]

   d. \forall D'[D' \subset D \rightarrow \neg \exists x[S(x) \land \exists y \in D'[P(x) \land R(x,y)] \land I(m,x)]]
      (For any proper subdomain D', Mary didn’t any students who read some papers in D'.)
   e. Prejacent presupposition: \exists x[S(x) \land \exists y \in D[P(x) \land R(x,y)] \land I(m,x)]
      (Mary invited some students who read some papers in the total domain D.)
   f. (68d) contradicts (68e). For example, let D = \{a, b\}, then:
      (68d) = \neg \exists x[S(x) \land R(x,a) \land I(m,x)] \land \neg \exists x[S(x) \land R(x,b) \land I(m,x)]
      (Mary didn’t invite any students who read a or any students who read b.)
      (68e) = \exists x[S(x) \land R(x,a \cup b) \land I(m,x)]
      (Mary invited some students who read a or b.)

In sum, if the prejacent of only is non-DE with respect to the D variable of any, then regardless of whether F-movement takes place, using only to assess the [D] feature of any yields a contradiction between the prejacent presupposition and the negation of the proper D-alternatives, making the NPI any unlicensed.

4.3. Head Restriction

Recall that a focused head cannot undertake F-movement alone, because F-movement is phrasal movement. Hence, in (92), it is impossible to move the focused verb CUT_F alone to the restrictor of only while leaving the NPI any in the scope of only.

(92) *John only CUT_F any vegetables.
   [O_D [\{only CUT_F,i\} \land VP John t_i anyD veggies]]
Unlike the cases that are subject to island constraints, cases that are subject to the Head Restriction do have the option of using a global covert $O_D$-operator to assess the $[D]$ feature of $any_D$. For instance, in (93), as Gajewski points out, the $any_P$ can vacate the VP, stranding the remnant VP associated with $only$ (reported by Wagner 2006 in fn. 14).

(93) *John only CUT$_F$ any vegetables.

\[ O_D [_{DP any_D \text{vegs}}], [\text{only } [_{VP \text{John CUT}_F t_i}]] \]

Nevertheless, the NPI $any$ still cannot be licensed. The reason is obvious: vacating the VP, the $any_P$ can be raised to the position sandwiched between $O_D$ and $only$, and this is the only position where it can be raised to. In this structure, the $[D]$ feature of $any$ is assessed under the immediate complement of $O_D$, yielding a G-triviality.

On the other hand, if the $any_P$ could be raised to a DE context, the NPI $any$ would be licensed. For instance, the conditional sentence (94), which takes (93) as its antecedent, is more acceptable. This is so for the following reason: the whole conditional is DE with respect to $any$; applying $O_D$ over the entire conditional does not yield a G-triviality.

(94) (?) If John only CUT any vegs (and didn’t STEAM any vegs), Mary would be unhappy.

\[ O_D [\text{if } [_{IP \text{any}_D \text{vegs}}], [\text{only } [_{VP \text{John CUT}_F t_i}]]] [_{IP \text{Mary would be happy}}] ] \]

Moreover, if the quantifier raising of the $any_P$ is blocked, the NPI $any$ would not be licensed. For instance, the NPI $any$ is not licensed in (95), a conditional where $only$ is associated with the NP complement of $any$. First, the determiner $any$ cannot take F-movement alone, ruling out the possibility in (95a). Second, since an $only$-associated focus cannot be moved from beneath $only$ (Tancredi 1990, Beaver and Clark 2003), the F-contained $any_P$ cannot raise over $only$, ruling out the possibility in (95b).

(95) *If John only invited [anyone’s ADVISORS$_F$], the students would be unhappy.

a. If John only invited anyone’s ADVISORS$_F$, ...

b. If John only invited anyone’s ADVISORS$_F$, ...

5. Conclusions

The goal of this paper has been to explain the distributional pattern of weak NPIs in $only$-sentences. The phenomenon of Licensed-Nonfocal Mismatch suggests that the licensing status of an NPI in
an *only*-sentence is constrained not only by the monotonicity pattern of the environment where this NPI gets interpreted, but also by syntactic factors in the LF of *only*-sentences.

I reviewed two representative analyses, namely, the F-movement theory by Wagner (2006) and the G-view of NPI-licensing by Chierchia (2006, 2013). I showed that both analyses have clear advantages but also make incorrect predictions. First, neither analysis can properly explain the phenomenon of Licensed-Nonfocal Mismatch. Both analyses assume that weak NPIs can be licensed in any Strawson DE or DE environment, and both analyses predict that the asserted component of an *only*-sentence is DE with respect to any unfocused position. Second, in Wagner’s treatment, the F-movement operation in VP-*only* association is not well-motivated or properly controlled. Third, Chierchia’s analysis cannot capture cases where *only* is associated with an NPI across an existential quantifier.

As an alternative to these approaches, I incorporated F-movement into the G-view by assuming that the requirement of avoiding G-trivialities can motivate covert F-movement. Moreover, I argued that *only* is not just an NPI-licenser but also an “NPI-unlicenser.” To capture the latter role, I proposed that *only*, just like the covert $O_D$-operator, can check off the grammatical feature [D] and operate on D-alternatives.

Compared with previous analyses, my proposal yields two new predictions. First, when an NPI is interpreted within the immediate complement of *only*, a contradiction arises between the exhaustivity assertion and the prejacent presupposition, making this NPI unlicensed. Second, in the case of VP-*only* association, this contradiction can be avoided via F-movement if the NPI does not appear within the F-moved phrase. These predictions explain why the distributional pattern of NPIs in *only*-clauses exhibits Licensed-Nonfocal Mismatch, as well as why VP-*only* association sometimes triggers F-movement.

**Acknowledgements**

For helpful discussions on the current and earlier versions of this paper, I thank Gennaro Chierchia, Noah Constant, Michael Erlewine, Danny Fox, Martin Hackl, I-Ta Chris Hsieh, C.-T. James Huang, Andreea Nicolae, Gary Thoms, Hedde Zeijlstra, as well as additional audiences at MIT LF Reading Group, the 37th GLOW colloquium, and the 19th Sinn und Bedeutung. I am grateful to the editor (Yael Sharvit), the proofreader (Nathan Klinedinst), and the anonymous reviewers of *Journal of Semantics*. Their comments have substantially shaped the way the ideas in this paper are presented. Special thanks to Zuzanna Fuchs for proofreading this paper. This paper expands significantly on Xiang (2015) in the *Proceedings of Sinn und Bedeutung 19*. All errors are mine.

YIMEI XIANG  
Department of Linguistics  
Harvard University  
Boylston Hall Floor 3  
Cambridge, MA 02138, USA  
Email: xiang.yimei@gmail.com
References


Fox, Danny (2007). Free choice and the theory of scalar implicatures. In Sauerland, Uli and

Fox, Danny (2013). Mention some readings of questions. MIT class notes.


McCawley, James D. (1993). *Everything that linguists have always wanted to know about logic... but were ashamed to ask* (2nd edition). University of Chicago Press.


Tancredi, Christopher (1990). Not only EVEN, but even ONLY. Unpublished MS.