Focus

Lecture 3: Alternative Semantics 2.0

Yasutada Sudo
UCL
y.sudo@ucl.ac.uk

OVA 2026 at Debreceni Egyetem

Alternative Semantics with the squiggle:

(1) a. $\left[\!\left[\sim \; \texttt{C} \quad \texttt{XP} \right]\!\right]^g_o = \lambda w_s\colon\ g(\texttt{C}) \subseteq [\![\texttt{XP}]\!]^g_f \text{ and } [\![\texttt{XP}]\!]^g_o \in g(\texttt{C}) \text{ and } |g(\texttt{C})| \geqslant 2.\ [\![\texttt{XP}]\!]^g_o$

b. $\left[\!\left[\sim \; \texttt{C} \quad \texttt{XP} \right]\!\right]^g_f = \left\{ \left[\!\left[\sim \; \texttt{C} \quad \texttt{XP} \right]\!\right]^g_o \right\}$

(2) $\left[\!\left[\texttt{only}_{\texttt{VP}} \;\; \texttt{C} \;\; \texttt{VP} \;\; (\sim \; \texttt{C}) \right]\!\right]^g_o$ (VP = only$_{\texttt{VP}}$ C [VP …] ∼ C)

$$= \lambda w_s.\lambda x_e\colon \left[\begin{array}{l} g(\texttt{C}) \subseteq [\![\texttt{VP}]\!]^g_f \text{ and} \\ [\![\texttt{VP}]\!]^g_o \in g(\texttt{C}) \text{ and } |g(\texttt{C})| \geqslant 2 \text{ and} \\ [\![\texttt{VP}]\!]^g_o(w)(x) = 1 \text{ and} \\ \text{for each } f \in g(\texttt{C}),\ w \in \mathrm{dom}(f) \text{ and} \\ \quad x \in \mathrm{dom}(f(w)) \end{array} \right] \cdot \left[\begin{array}{l} \text{for each } f \in g(\texttt{C}), \\ \text{if } f \neq [\![\texttt{VP}]\!]^g_o, \text{ then } f(w)(x) = 0 \end{array} \right]$$

NB: According to this analysis, 'only' presupposes that the prejacent is true, and also projects the presuppositions of the alternatives. Both of these are potentially controversial. For instance, you might want to only project the presuppositions of the negated alternatives (see the discussion at the end of Lecture 2).

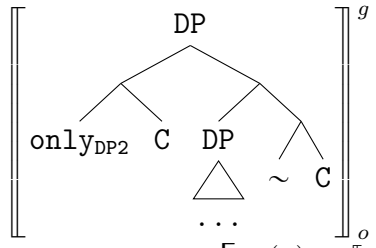
Pre-DP *only* can be analyzed as a generalized quantifier (recall *only DP* takes scope like quantifiers). Assuming that the modified DP is a referring expression (of type $\langle s, e \rangle$):

(3) $\left[\!\left[\texttt{only}_{\texttt{DP1}} \;\; \texttt{C} \;\; \texttt{DP} \;\; (\sim \; \texttt{C}) \right]\!\right]^g_o$ (DP = […])

$$= \lambda w_s.\lambda P_{\langle e,t\rangle}\colon \left[\begin{array}{l} g(\mathtt{C}) \subseteq [\![\mathtt{DP}]\!]^g_f \text{ and} \\ [\![\mathtt{DP}]\!]^g_o \in g(\mathtt{C}) \text{ and } |g(\mathtt{C})| \geqslant 2 \text{ and} \\ P([\![\mathtt{DP}]\!]^g_o(w)) = 1 \text{ and} \\ \text{for each } f \in g(\mathtt{C}),\ w \in \mathrm{dom}(f) \text{ and} \\ \quad f(w) \in \mathrm{dom}(P) \end{array}\right] . \left[\begin{array}{l} \text{for each } f \in g(\mathtt{C}), \\ \text{if } f \neq [\![\mathtt{DP}]\!]^g_o, \text{ then } P(f(w)) = 0 \end{array}\right]$$

If the modified DP is a quantifier (of type $\langle s, \langle\langle e,t\rangle, t\rangle\rangle$):

(4)



$$= \lambda w_s.\lambda P_{\langle e,t\rangle}\colon \left[\begin{array}{l} g(\mathtt{C}) \subseteq [\![\mathtt{DP}]\!]^g_f \text{ and} \\ [\![\mathtt{DP}]\!]^g_o \in g(\mathtt{C}) \text{ and } |g(\mathtt{C})| \geqslant 2 \text{ and} \\ [\![\mathtt{DP}]\!]^g_o(w)(P) = 1 \text{ and} \\ \text{for each } f \in g(\mathtt{C}),\ w \in \mathrm{dom}(f) \text{ and} \\ \quad P \in \mathrm{dom}(f(w)) \end{array}\right] . \left[\begin{array}{l} \text{for each } f \in g(\mathtt{C}), \\ \text{if } f \neq [\![\mathtt{DP}]\!]^g_o, \text{ then } f(w)(P) = 0 \end{array}\right]$$

We've added assignment functions and presuppositions, so let's also restate the compositional rule with these additional devices:

(5) *Extensional Functional Application*:
If α is a constituent consisting of β and γ such that $[\![\beta]\!]^g_o \in D_{\langle s,\langle\sigma,\tau\rangle\rangle}$ and $[\![\gamma]\!]^g_o \in D_{\langle s,\sigma\rangle}$, then

$$[\![\alpha]\!]^g_o = \lambda w_s\colon \begin{array}{l} w \in \mathrm{dom}([\![\beta]\!]^g_o) \text{ and } w \in \mathrm{dom}([\![\gamma]\!]^g_o) \text{ and} \\ [\![\gamma]\!]^g_o(w) \in \mathrm{dom}([\![\beta]\!]^g_o(w)) \end{array} .[\![\beta]\!]^g_o(w)([\![\gamma]\!]^g_o(w))$$

We will discuss four problems of this theory, all of which are discussed in existing work.

- The problem of unnatural functions
- The problem of over-focussing
- The problem of co-varying foci ('Tanglewood')
- The problem of multiple foci

1 The problem of unnatural functions

An F-marked phrase has as its alternatives everything that has the same semantic type. E.g.,

- If a transitive verb is focused, its alternatives are all functions of type $\langle s, \langle e, \langle e, t\rangle\rangle\rangle$.

(6) a. $[\![\mathtt{saw}_F]\!]^g_o = \lambda w_s.\lambda x_e.\lambda y_e.\ y \text{ saw } x \text{ in } w$
b. $[\![\mathtt{saw}_F]\!]^g_f = D_{\langle s,\langle e,\langle e,t\rangle\rangle\rangle}$

- Some of the alternative functions are 'natural', e.g. (7), but others are not, e.g. (9)–(11).

(7) a. $\lambda w_s.\lambda x_e.\lambda y_e.\ y$ met x in w
b. $\lambda w_s.\lambda x_e.\lambda y_e.\ y$ made small talk with x in w
c. $\lambda w_s.\lambda x_e.\lambda y_e.\ y$ discussed linguistics with x in w

(8) a. $\lambda w_s.\lambda x_e.\lambda y_e.\ y$ saw Berit in w
b. $\lambda w_s.\lambda x_e.\lambda y_e.\ y$ saw Quine in w
c. $\lambda w_s.\lambda x_e.\lambda y_e.\ y$ saw Chomsky in w

(9) a. $\lambda w_s.\lambda x_e.\lambda y_e.\ y$ sneezed in w
b. $\lambda w_s.\lambda x_e.\lambda y_e.\ y$ jumped in w
c. $\lambda w_s.\lambda x_e.\lambda y_e.\ y$ met Berit in w
d. $\lambda w_s.\lambda x_e.\lambda y_e.\ y$ ate pasta in w
e. $\lambda w_s.\lambda x_e.\lambda y_e.\ y$ walked to the park in w

(10) a. $\lambda w_s.\lambda x_e.\lambda y_e.$ it was raining in w
b. $\lambda w_s.\lambda x_e.\lambda y_e.$ the tallest person in the world was having lunch in w

(11) a. $\lambda w_s.\lambda x_e.\lambda y_e.\ x$ sneezed in w
b. $\lambda w_s.\lambda x_e.\lambda y_e.\ y$ jumped in w and x didn't jump in w
c. $\lambda w_s.\lambda x_e.\lambda y_e.\ x$ saw y in w

These unnatural functions lead to overgeneration problems (as previously discussed by Kadmon 2001: §15.3, Büring 2015: fn. 1 and Büring 2016: §5.4.1).

(12) Patrick only [saw]$_F$ Chomsky.

The focus semantic value of the VP is:

(13) $$\left[\!\!\left[\begin{array}{c} \texttt{VP} \\ \underline{\texttt{saw}}_F \quad \texttt{Chomsky} \end{array} \right]\!\!\right]^g_f = \{\, \lambda w_s.\lambda x_e.\ R(w)([\![\texttt{Chomsky}]\!]^g_o(w))(x) \mid R \in D_{\langle s,\langle e,\langle e,t\rangle\rangle\rangle} \,\}$$

If $g(\texttt{C})$ only contains 'natural' functions like (7), we derive sensible readings. However:

- If $g(\texttt{C})$ consists of 'unnatural' functions in (8), the reading will be identical to the object focus reading.
- Similarly, those in (9) will yield what is identical to the VP focus reading.
- Those in (10) will yield what is identical to the sentential focus reading.
- The theory allows $g(\texttt{C})$ to have a mixture of these functions too.

It is clear that the sentence cannot mean the readings that are predicted by the unnatural alternatives.
We need a way to restrict the focus alternatives to natural functions.

- Since *only* negates only contextually relevant alternatives, we could try to construct a pragmatic theory of contextual relevant in such a way that unnatural functions are always made contextually irrelevant. As far as I know, no specific proposal along these lines has been put forward.

- Flóra Lili Donáti and I claim that in a recent paper that this problem dissolves, at least partly, if we assume that alternatives are not semantic objects, but (grammatical) linguistic expressions (as previously proposed by Fox 1999, Fox & Katzir 2011). For example, natural language bans DPs without theta roles (cf. Chomsky's 1981 Theta Criterion). We'll discuss this idea in the next lecture.

Let us assume for now that there is some way of solving the problem of unnatural functions.

2 The problem of over-focussing

Büring (2016: p. 44) discusses the issue of over-focussing as another issue of Alternative Semantics.

- Fact: A broader focus will always have as its focus semantic value a superset of the focus semantic value of a narrower focus, for example, (14a) is a superset of (14b).

(14) a. $\left[\!\left[\begin{array}{c} \texttt{VP}_F \\ \texttt{saw}\ \ \underline{\texttt{Chomsky}} \end{array} \right]\!\right]^g_f = D_{\langle s,\langle e,t\rangle\rangle}$

b. $\left[\!\left[\begin{array}{c} \texttt{VP} \\ \underline{\texttt{saw}}_F\ \ \texttt{Chomsky} \end{array} \right]\!\right]^g_f = \{\, \lambda w_s.\lambda x_e.\ R(w)([\![\texttt{Chomsky}]\!]^g_o(w))(x) \mid R \in D_{\langle s,\langle e,\langle e,t\rangle\rangle\rangle} \,\}$

- The focus semantic value is only used to restrict $g(\texttt{C})$ to be a subset of it: $g(\texttt{C}) \subseteq [\![\texttt{XP}]\!]^g_f$.
- So, if $g(\texttt{C}) \subseteq$ (14b), then $g(\texttt{C}) \subseteq$ (14a).
- Then one and the same value of C should be compatible with both focus structures.

But this is wrong.

(15) a. #Patrick read my dissertation in its entirety, but only [opened Jacopo's]$_F$.
b. Patrick read my dissertation in its entirety, but only [opened]$_F$ Jacopo's.

The heart of this problem is that Alternative Semantics does not penalize unnecessarily broad focus. Importantly, however, Alternative Semantics is not incompatible with a mechanism that does (cf. AVOIDF of Schwarzschild 1999).

- E.g., *Maximize Presupposition! (MP)* (Heim 1991) forces the use of the presuppositionally strongest form (among a certain set of candidates), whenever its presupposition is satisfied, and this has been given various empirical support (see also related works like Percus 2006, Singh 2011, Spector & Sudo 2017, Marty 2017; see Bade 2021 for an overview).
- If MP compares different F-marking possibilities, it'll ban all focus structures except the one that yields the strongest presupposition compatible with the value of $g(\texttt{C})$.

I will put aside how to formally state MP, but I think it's fair to say that the problem of over-focussing does not threaten Alternative Semantics.

3 The problem of co-varying foci

3.1 The Tanglewood example

Kratzer 1991 argues that Roothian Alternative Semantics is not expressive enough to account for sentences like (16).

> Imagine now you are angry at me and start voicing the following accusations. "What a copy cat you are! You went Block Island because I did. You went to Elk Lake Lodge because I did. And you went to Tanglewood because I did." I feel you exaggerate and reply:
>
> (15) I only $_{\text{VP}}$[went to $_{\text{F}}$[Tanglewood]] because you did $_{\text{VP}}$[e]
>
> (Kratzer 1991: p. 830)

3.2 The problem

To simplify the discussion, let us change the pronouns to proper names, and assume $[\![\texttt{Alex}]\!]^g_o = \lambda w_s.\ a$ and $[\![\texttt{Bob}]\!]^g_o = \lambda w_s.\ b$.

(16) Alex only [$_{\text{VP}}$ went to Tanglewood$_F$ because Bob did]

3.2.1 Possibility 1

Let us assume that the ellipsis site contains a copy of its antecedent. If it contains no F feature:

(17) Alex only [$_{\text{VP}}$ went to Tanglewood$_F$ because Bob did $\langle$go to Tanglewood$\rangle$]

The focus alternatives that *only* will operate on will be:

(18) $$\left\{ \begin{array}{l} \lambda w_s.\lambda x_e.\ x \text{ went to Tanglewood because } b \text{ went to Tanglewood in } w, \\ \lambda w_s.\lambda x_e.\ x \text{ went to Block Island because } b \text{ went to Tanglewood in } w, \\ \lambda w_s.\lambda x_e.\ x \text{ went to Elk Lake Lodge because } b \text{ went to Tanglewood in } w \end{array} \right\}$$

We fail to account for the intuitive meaning.

3.2.2 Possibility 2

If the elided material is also F-marked:

(19) Alex only [$_{\text{VP}}$ went to Tanglewood$_F$ because Bob did $\langle$go to Tanglewood$_F\rangle$]

The focus alternatives that *only* will operate on will be:

(20)
$$\left\{\begin{array}{l}
\lambda w_s.\lambda x_e.\ x \text{ went to Tanglewood because } b \text{ went to Tanglewood in } w,\\
\lambda w_s.\lambda x_e.\ x \text{ went to Block Island because } b \text{ went to Block Island in } w,\\
\lambda w_s.\lambda x_e.\ x \text{ went to Elk Lake Lodge because } b \text{ went to Elk Lake Lodge in } w,\\
\\
\lambda w_s.\lambda x_e.\ x \text{ went to Tanglewood because } b \text{ went to Block Island in } w,\\
\lambda w_s.\lambda x_e.\ x \text{ went to Tanglewood because } b \text{ went to Elk Lake Lodge in } w,\\
\lambda w_s.\lambda x_e.\ x \text{ went to Block Island because } b \text{ went to Tanglewood in } w,\\
\lambda w_s.\lambda x_e.\ x \text{ went to Block Island because } b \text{ went to Elk Lake Lodge in } w,\\
\lambda w_s.\lambda x_e.\ x \text{ went to Elk Lake Lodge because } b \text{ went to Tanglewood in } w,\\
\lambda w_s.\lambda x_e.\ x \text{ went to Elk Lake Lodge because } b \text{ went to Block Island in } w
\end{array}\right\}$$

On the reading we are after, only the second and third alternatives in (20) are negated.

One could argue that in the context Kratzer 1991 describes, only the first three alternatives are relevant.

However, we still predict that in the right context, you should be able to negate the other alternatives, but such a reading does not seem to be available.

3.3 More examples

(21) a. You always see more Edsels than I do.
b. No, I only saw more [pink]$_F$ Edsels than you did.
(adapted from Kratzer 1991: p. 831)

If all colors are relevant, the sentence should mean:

(22) I saw more pink Edsels than you saw pink Edsels, and;
a. ¬(I saw more pink Edsels than you saw red Edsels), and;
b. ¬(I saw more pink Edsels than you saw blue Edsels), and;
c. ¬(I saw more pink Edsels than you saw black Edsels), and;
d. ¬(I saw more pink Edsels than you saw white Edsels), and;
⋮
e. ¬(I saw more red Edsels than you saw red Edsels), and;
f. ¬(I saw more red Edsels than you saw pink Edsels), and;
g. ¬(I saw more red Edsels than you saw black Edsels), and;
h. ¬(I saw more red Edsels than you saw white Edsels), and;
⋮

Schematically,

(23) $$\text{U-red/blue/black/white} \gg \begin{array}{l}\text{I-pink} > \text{U-pink}\\ \text{I-red/blue/black/white}\end{array}$$

So you should conclude that you saw fewer pink Edsels than Edsels of other colors. The sentence doesn't have a reading that implies that.

(24) shows that this phenomenon is island insensitive.

(24) a. You always contact every responsible person before me.
b. No, I only contacted the person who chairs [the Zoning Board]$_F$ before you did.
(adapted from Kratzer 1991: p. 831)

3.4 Copy Theory of Movement

A similar issue arises under the copy theory of movement, when a moved phrase contains a focus (Erlewine 2014).

(25) The professor only wants the $[\text{MPhil}]_F$ students to ~~the $[\text{MPhil}]_F$ students~~ ask questions.

3.5 Focus variables

Kratzer 1991 proposes a binding mechanism for focus (inspired by Rooth 1985: p. 12; see also Erlewine 2014).[1]

- F-marking comes with an index n and a type τ, as in $\alpha_{F^{\tau}_{n}}$.
- Focus assignment function h maps the index to some member of D_{τ}.
- The set of denotations with respect to all possible focus functions is the alternative set.
- If two focussed phrases share an index, they co-vary.

3.5.1 Ordinary semantic values

(26)
a. $[\![\texttt{boy}]\!]^{g,h}_{o} = \lambda w_s.\lambda x_e.\ x \text{ is a boy in } w$
b. $[\![\texttt{run}]\!]^{g,h}_{o} = \lambda w_s.\lambda x_e.\ x \text{ runs in } w$
c. $[\![\texttt{saw}]\!]^{g,h}_{o} = \lambda w_s.\lambda x_e.\lambda y_e.\ y \text{ saw } x \text{ in } w$
d. $[\![\texttt{every}]\!]^{g,h}_{o} = \lambda w_s.\lambda P_{\langle e,t\rangle}\colon\ .\lambda Q_{\langle e,t\rangle}.\ \text{for each } x \in D_e, \text{ if } P(x) = 1, \text{ then } Q(x) = 1$

(NB: (26d) does not account for presupposition projection!)

3.5.2 Focus semantic values

Without F-marking

(27)
a. $[\![\texttt{boy}]\!]^{g,h}_{f} = [\![\texttt{boy}]\!]^{g,h}_{o}$
b. $[\![\texttt{run}]\!]^{g,h}_{f} = [\![\texttt{run}]\!]^{g,h}_{o}$
c. $[\![\texttt{saw}]\!]^{g,h}_{f} = [\![\texttt{saw}]\!]^{g,h}_{o}$
d. $[\![\texttt{every}]\!]^{g,h}_{f} = [\![\texttt{every}]\!]^{g,h}_{o}$

With F-marking

(28)
a. $\left[\!\left[\texttt{boy}_{F^{\langle s,\langle e,t\rangle\rangle}_{n}}\right]\!\right]^{g,h}_{f} = h(n)$
b. $\left[\!\left[\texttt{run}_{F^{\langle s,\langle e,t\rangle\rangle}_{n}}\right]\!\right]^{g,h}_{f} = h(n)$
c. $\left[\!\left[\texttt{saw}_{F^{\langle s,\langle e,\langle e,t\rangle\rangle\rangle}_{n}}\right]\!\right]^{g,h}_{f} = h(n)$
d. $\left[\!\left[\texttt{every}_{F^{\langle s,\langle\langle e,t\rangle,\langle e,t\rangle\rangle\rangle}_{n}}\right]\!\right]^{g,h}_{f} = h(n)$
e. $\left[\!\left[\texttt{no}_{F^{\langle s,\langle\langle e,t\rangle,\langle e,t\rangle\rangle\rangle}_{n}}\right]\!\right]^{g,h}_{f} = h(n)$

- Each focus assignment yields one specific focus alternative = a model-theoretic object of the same semantic type as the ordinary semantic value.

[1] The formulation here is not faithful to Kratzer 1991 in some superficial details, but the essence is not lost. Similar analyses have been put forward in the framework of Structured Meaning by Jäger 1999 and Krifka 1991, 2006.

- The focus semantic values compose via ordinary compositional rules.

(29) a. $\left[\!\left[\begin{array}{c} \\ \sim \quad \texttt{C} \end{array} \right.\right.$ XP $\left.\!\left]\right]^{g,h}_{o} = \lambda w_s\colon \begin{array}{l} g(\texttt{C}) \subseteq \{\, [\![\texttt{XP}]\!]^{g,h'}_{f} \mid h' \text{ is a focus assignment} \,\} \text{ and} \\ [\![\texttt{XP}]\!]^{g,h}_{o} \in g(\texttt{C}) \text{ and } |g(\texttt{C})| \geqslant 2 \end{array} . [\![\texttt{XP}]\!]^{g,h}_{o}$

b. $\left[\!\left[\begin{array}{c} \\ \sim \quad \texttt{C} \end{array} \right.\right.$ XP $\left.\!\left]\right]^{g,h}_{f} = \left[\!\left[\begin{array}{c} \\ \sim \quad \texttt{C} \end{array} \right.\right.$ XP $\left.\!\left]\right]^{g,h}_{o}$

The meaning of *only* is the same as before. Pre-VP *only* looks like:

(30) $[\![\texttt{only}_{\texttt{VP}}\ \texttt{C}]\!]^{g,h}_{o} = \lambda w_s.\lambda P_{\langle e,t\rangle}\lambda x_e\colon \begin{array}{l} \text{for each } f \in g(\texttt{C}), \\ w \in \mathrm{dom}(f) \text{ and } x \in \mathrm{dom}(f(w)) \text{ and} \\ P(x) = 1 \end{array} . \begin{array}{l} \text{for each } f \in g(\texttt{C}), \\ f(w)(x) = 0 \end{array}$

If there are two focussed phrases bearing the same index, they will co-vary.

(31) Alex only [$_{\text{VP}}$ went to Tanglewood$_{F^{\langle s,e\rangle}_5}$ in w,
because Bob did $\langle$go to Tanglewood$_{F^{\langle s,e\rangle}_5}\rangle$]

Each focus alternative looks like (32) (for different focus assignment h):

(32) λw_s. Alex went to $h(5)(w)$, because Bob went to $h(5)(w)$ in w

3.6 De-accenting?

Charlow 2008 puts forward an alternative analysis of the Tanglewood sentence under a pronominal analysis of VP ellipsis where the ellipsis site contains no structure, so no F-marking.
However, a co-varying interpretation seems to be available without VP ellipsis, with de-accenting. If this is correct, his proposal is arguably not general enough.

(33) I only went to Tanglewood$_F$, because you went to Tanglewood.

3.7 Expressive power

This is a proper extension of Roothian Alternative Semantics.

- Kratzer 1991 assumes that every focus index on an overt material has to be distinct ('Novelty Condition'), so co-varying foci are only possible under deletion/ellipsis.
- But if (33) has a co-varying interpretation, this condition needs to be relaxed.
- In the next lecture, we will consider an analysis of co-varying foci that does not require a binding mechanism.

We will come back to focus indices in §4.2.

4 The problem of multiple foci

Roothian Alternative Semantics has no problem accounting for the following types of multiple foci (with or without focus assignments):

- One focus sensitive operator associating with multiple foci.

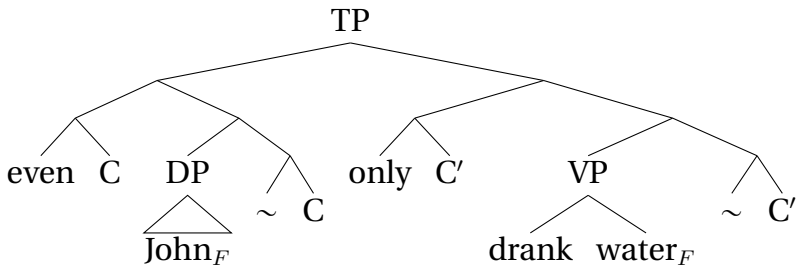
(34) John only introduced Bill$_F$ to Sue$_F$. (Krifka 1991: p. 130)

(35) $\left[\!\left[\begin{array}{c} \texttt{VP} \\ \texttt{introduced}\ \ \texttt{Bill}_F \quad \texttt{Sue}_F \end{array} \right]\!\right]^g_f = \{\, \lambda w_s.\lambda x_e.\ x \text{ introduced } y \text{ to } z \text{ in } w \mid y, z \in D_e \,\}$

- Two focus sensitive operators, each associating with one focus, with no crossing dependency.

(36) Even John$_F$ only drank water$_F$. (Krifka 1991: p. 131)

(37)



- Two focus sensitive operators, each associating with one focus, with nested dependencies.

(38) John even [$_{\text{VP}}$ [drank only water$_F$] [on his birthday]$_F$].

Recall:

(39) a. $\left[\!\left[\begin{array}{c} \texttt{XP} \\ \sim\ \texttt{C} \end{array} \right]\!\right]^g_o = \lambda w_s\colon\ g(\texttt{C}) \subseteq [\![\texttt{XP}]\!]^g_f \text{ and } [\![\texttt{XP}]\!]^g_o \in g(\texttt{C}) \text{ and } |g(\texttt{C})| \geqslant 2.\ [\![\texttt{XP}]\!]^g_o$

b. $\left[\!\left[\begin{array}{c} \texttt{XP} \\ \sim\ \texttt{C} \end{array} \right]\!\right]^g_f = \left\{ \left[\!\left[\begin{array}{c} \texttt{XP} \\ \sim\ \texttt{C} \end{array} \right]\!\right]^g_o \right\}$

So *even* will not see the alternatives of *water*$_F$.

However, the following types of multiple foci pose issues.

1. Two focus sensitive operators associating with the same focus (or overlapping foci).

(40) John even only drank water$_F$. (Krifka 1991: p. 131)

(41) John once only drank wine$_F$. John also once only drank water$_F$.

(Wold 1996: p. 320)

2. Two focus sensitive operators associating with two distinct foci in the same domain:

(42) John only introduced Sue$_F$ to Fred.

John also only [$_{VP}$ introduced Sue$_F$ to Bill$_F$]. (Wold 1996: p. 320)

The first case is easier to solve than the second. But there is controversy over the judgment of (42) (Beck & Vasishth 2009, Erlewine 2025).

4.1 Same focus (or overlapping foci)

4.1.1 A closer look at the examples

(43) John even only drank water$_F$. (Krifka 1991: p. 131)

This example might have multiple readings with respect to the interpretation of *even*.

(44) 'only drank water' is more noteworthy than
   a. 'only drank beer'
   b. 'drank alcohol (and possibly other things)'
   c. 'not only drank water' (e.g., 'drank water and alcohol')
   d. other relevant actions

The following example also has multiple interpretations with respect to *also*.

(45) John once only drank wine$_F$. John also once only drank water$_F$.

(Wold 1996: p. 320)

The additive particle triggers a presupposition that a focus alternative distinct (or more precisely, logically independent) from the prejacent to be true (more on this in Lecture 7). Grammatically, there are some candidates for the focus alternative that is presupposed to be true.

(46) a. 'only drank wine'
   b. 'drank wine'
   c. 'drank alcohol'
   d. 'not only drank water' (e.g., 'drank water and alcohol')
   e. some other relevant action

For the argument here, it's crucial that the reading based on the alternative that includes *only* in (a) is available as a distinct reading. It's actually not so clear if that is the case.

The crucial reading is perhaps easier to detect for an example like (47).

(47) Mary also posted a picture where only Chomsky$_F$ is smiling.

This example seems to have a reading that presupposes about someone who is not Chomsky that Mary posted a picture where only they are smiling.

4.1.2 The problem

The problem is that once ∼ applies for the lower focus sensitive operator (e.g., *only*), the focus alternatives will be lost for the the higher focus sensitive operator (e.g., *also*).

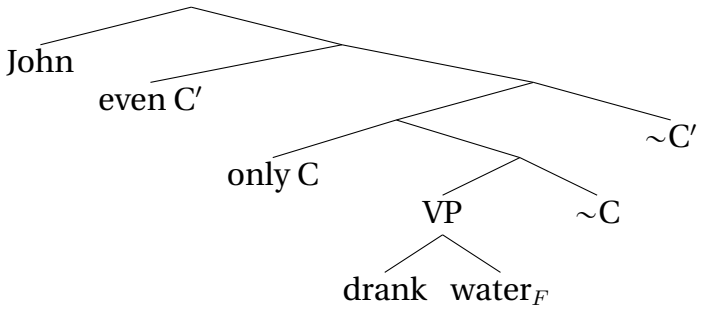
(48) a. $[\![\sim \ \texttt{C} \ \ \texttt{XP}]\!]^g_o = \lambda w_s\colon\ g(\texttt{C}) \subseteq [\![\texttt{XP}]\!]^g_f \text{ and } [\![\texttt{XP}]\!]^g_o \in g(\texttt{C}) \text{ and } |g(\texttt{C})| \geqslant 2.\ [\![\texttt{XP}]\!]^g_o$

b. $[\![\sim \ \texttt{C} \ \ \texttt{XP}]\!]^g_f = \{\ [\![\sim \ \texttt{C} \ \ \texttt{XP}]\!]^g_o\ \}$

(49) $[\![\sim\texttt{C} \ \ \texttt{VP}\ (\texttt{drank} \ \ \texttt{water}_F)]\!]^g_f = \{\ [\![\texttt{VP}\ (\texttt{drank} \ \ \texttt{water}_F)]\!]^g_o\ \}$

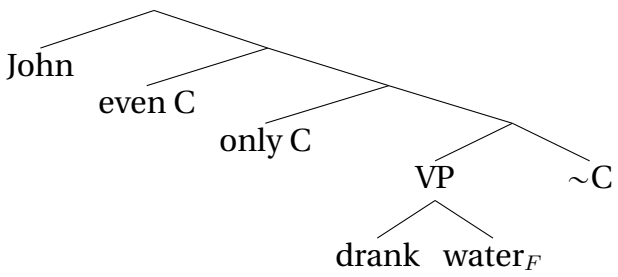
Consequently, the presupposition of (50) won't be satisfiable (as $g(\texttt{C}')$ will be presupposed to be a subset of a singleton and also to have two or more members at the same time).

(50)



For this example, one can consider positing just one ∼ and using the same C:

(51)



(52) $\left[\!\left[\begin{array}{c} \texttt{VP} \\ \texttt{drank} \quad \texttt{water}_F \end{array} \right]\!\right]^g_f = \{\, \lambda w_s.\lambda x_e.\ x \text{ drank some } y \text{ such that } P(y) = 1 \text{ in } w \mid P \in D_{\langle e,t\rangle} \,\}$

According to this analysis, $g(\texttt{C})$ will be a subset of (52), so it won't contain 'only'. If there's a reading where the alternatives for the higher focus sensitive operator *even* include 'only', it will fail to capture it. However, if this is available—and there does not seem to be reason to assume it is not—it's not so clear if we need to postulate a different reading.

The following example cannot be analyzed with a common $g(\texttt{C})$, because the alternatives for the two focus sensitive operators differ in semantic type: *also* wants alternatives of type $\langle s, \langle e, t\rangle\rangle$, *only* wants alternatives of type $\langle s, e\rangle$.

(53) Mary also$_{C'}$ [posted a picture where only$_C$ [Chomsky$_F$]$_{\sim C}$ is smiling]$_{\sim C'}$

The only thing that is F-marked is *Chomsky*, whose alternatives are not accessible to $\sim$ C′, so this example is predicted to have an unsatisfiable presupposition, contrary to fact.

4.1.3 A solution

The assumption that $\sim$ eats up focus alternatives is necessary to account for multiple foci with nested dependencies like (38). But we can also have a version of $\sim$ that passes the focus alternatives (Panizza & Sudo 2020, Erlewine 2025, etc.; see Erlewine 2025 for more references). Call it $\approx$.

(54) a. $\left[\!\left[\begin{array}{c} \texttt{XP} \\ \approx \quad C \end{array} \right]\!\right]^g_o = \lambda w_s\colon\ g(C) \subseteq [\![\texttt{XP}]\!]^g_f \text{ and } [\![\texttt{XP}]\!]^g_o \in g(C) \text{ and } |g(C)| \geqslant 2.\ [\![\texttt{XP}]\!]^g_o$

b. $\left[\!\left[\begin{array}{c} \texttt{XP} \\ \approx \quad C \end{array} \right]\!\right]^g_f = [\![\texttt{XP}]\!]^g_f$

The following parses are interpretable:

(55) a. Mary also$_{C'}$ [posted a picture where only$_C$ [Chomsky$_F$]$_{\approx C}$ is smiling]$_{\sim C'}$
b. Mary also$_{C'}$ [posted a picture where only$_C$ [Chomsky$_F$]$_{\approx C}$ is smiling]$_{\approx C'}$

4.2 Distinct foci in the same domain

(56) John only introduced Sue$_F$ to Fred.

John also only [$_{VP}$ introduced Sue$_F$ to Bill$_F$]. (Wold 1996: p. 320)

Let's assume for now that the sentence is acceptable under the intended focus associations.

4.2.1 The problem

$\sim$ and $\approx$ are *unselective* binders, and associate with all F-marked phrases in their scope.

(57) a. [$_{VP}$ introduced Sue$_F$ to Bill$_F$]$_{\sim C}$

b. $[_{\text{VP}}$ introduced Sue$_F$ to Bill$_F]_{\approx \text{C}}$

In these representations, C will be required to be a subset of (58).

(58) $[\![[_{\text{VP}}$ `introduced Sue`$_F$ `to Bill`$_F$`]` $]\!]^g_f$
$= \{ \lambda w_s.\lambda x_e.\ x \text{ introduced } y \text{ to } z \text{ in } w \mid y, z \in D_e \}$

This itself is not a huge issue, but it matters how we solve the over-focussing problem mentioned in §2:

- If the relevant alternatives for *only* all contain 'to Bill', *Bill* shouldn't be F-marked.
- But how the upper focus operator affects this needs to be considered carefully.
- We don't have time to introduce ideas about how anti-presuppositions project now, so let's not dwell on this potential problem (see Singh 2011, Spector & Sudo 2014 for relevant discussion).

The example might have a reading where both focus operators quantify over the same set of alternatives. This reading can be accounted for. And maybe that's all you need to say about this example.
In order to see the problem more clearly, let us consider a case where the two focus operators must operate on different sets of alternatives.

(59) John has read a textbook that only [cites one$_F$ paper by Chomsky].
He has also [read a textbook that only [cites one$_F$ paper by Jackendoff$_F$]].

As before, let's worry about the acceptability of the example later.

- We want the alternatives for *only* to be a subset of (60):

  (60) $\{ \lambda w_s.\lambda x_e.\ x \text{ cites } n\text{-many papers by Jackendoff in } w \mid n \text{ is a number} \}$

- And the alternatives for *also* to be a subset of (61):

  (61) $\{ \lambda w_s.\ \lambda x_e.\ x \text{ read in } w \text{ a textbook that cites exactly one paper by } y \text{ in } w \mid y \in D_e \}$

The problem is that $\sim$ and $\approx$ are unselective binders. In both of (62), $g(\mathtt{C})$ will be a subset of the focus value of the VP, which contains two F-marked elements.

(62) a. only$_{\text{C}}$ $[_{\text{VP}}$ cites one$_F$ paper by Jackendoff$_F$ $]_{\sim \text{C}}$
b. only$_{\text{C}}$ $[_{\text{VP}}$ cites one$_F$ paper by Jackendoff$_F$ $]_{\approx \text{C}}$

Similarly, in the following parses, C′ will also be a subset of the focus semantic value of the VP that contains two F-marked elements.

(63) a. also$_{\text{C}'}$ $[_{\text{VP}}$ read a textbook that only$_{\text{C}}$ $[_{\text{VP}}$ cites one$_F$ paper by Jackendoff$_F$ $]_{\approx \text{C}}$ $]_{\sim \text{C}'}$
b. also$_{\text{C}'}$ $[_{\text{VP}}$ read a book that only$_{\text{C}}$ $[_{\text{VP}}$ cites one$_F$ paper by Jackendoff$_F$ $]_{\approx \text{C}}]_{\approx \text{C}'}$

In this example, *Jackendoff*$_F$ can perhaps move outside the lower VP, but this phrase could be in an island as in (64).

(64) He has also read a textbook that only [cites one$_F$ paper [that criticizes Jackendoff$_F$]].

4.2.2 *Focus indices*

Kratzer's (1991) refinement of Alternative Semantics comes with focus indices, so one might wonder if they help us solve the problem.
In particular $\sim$ can be made selective, rather than unselective.

(65) $\left[\!\left[\sim^{\tau}_{i}\ \texttt{C}\ \ \texttt{XP} \right]\!\right]^{g,h}_{o} = \lambda w_s\colon \begin{array}{l} g(\texttt{C}) \subseteq \left\{ [\![\texttt{XP}]\!]^{g,h[i\mapsto x]}_{f} \;\middle|\; x \in D_\tau \right\} \text{ and} \\ [\![\texttt{XP}]\!]^{g,h}_{o} \in g(\texttt{C}) \text{ and } |g(\texttt{C})| \geqslant 2 \end{array} .\ [\![\texttt{XP}]\!]^{g,h}_{o}$

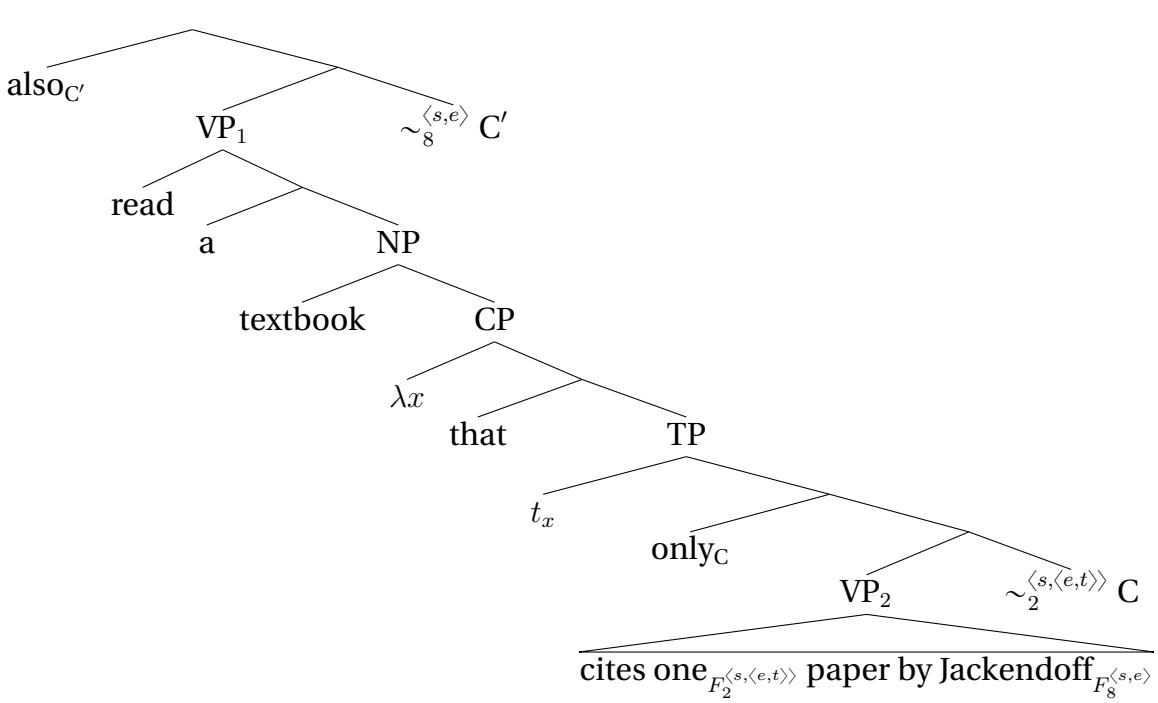
(66) $[\![\texttt{[only}_{\texttt{VP}}\ \texttt{C]}\ \ \texttt{VP}]\!]^{g,h}_{o} = \lambda w_s.\lambda x_e\colon \begin{array}{l} \text{for each } f \in g(\texttt{C}), \\ w \in \mathrm{dom}(f) \text{ and } x \in \mathrm{dom}(f(w)) \text{ and} \\ [\![\texttt{VP}]\!]^{g}_{o}(w)(x) = 1 \end{array} .\ \begin{array}{l} \text{for each } f \in g(\texttt{C}), \\ f(w)(x) = 0 \end{array}$

Instead of postulating $\approx$, this theory achieves the same effect by double indexing.

(67) Mary also$_{C'}$ [posted a picture where only$_{C}$ [[Chomsky$_{F_3^{\langle s,e\rangle}}$]$_{F_5^{\langle s,e\rangle}}$]$_{\sim_3^{\langle s,e\rangle}\text{C}}$ is smiling]$_{\sim_5^{\langle s,e\rangle}C'}$

But Wold 1996 points out that this won't be enough.
The example in question will be analyzed as (68).[2]

(68)



This is problematic in two ways:

1. First, the ordinary semantic value of VP$_1$ is asserted to be true of the subject in this sentence. We want $[\![\texttt{VP}_1]\!]^{g,h}_{o}$ to be (69).[3]

[2] I tentatively analyze numerals as intersective NP modifiers; nothing hinges on this assumption, as far as I can tell.

[3] We assume existential projection of the presupposition. This is wrong but is a good compromise in a simple static framework like the present one.

(69) $\lambda w_s.\lambda x_e$: [PRESUPPS on $g(\mathtt{C})$] and for some textbook t, $[\![\mathtt{VP}_2]\!]^{g,h}_o(\mathrm{w})(\mathrm{t})=1$. x read a textbook t in w s.t. for each $f \in g(\mathtt{C})$ t does not cite papers p in w s.t. p is by Jackendoff in w and $f(w)(p) = 1$

But we will fail to get the asserted meaning of (69) right. This has to do with the values of C. Because the higher $\sim$ does not do anything to the ordinary semantic value of VP_1, the computation will always be relative to g and h, up to $\sim_2$, and C will be presupposed to be a subset of (70).

(70) $$\left\{ \left[\!\left[[_{\mathtt{VP}_2}\ \mathtt{cites\ one}_{F_2^{\langle s,\langle e,t\rangle\rangle}}\ \mathtt{paper\ by\ Jackendoff}_{F_8^{\langle s,e\rangle}}] \right]\!\right]^{g,h[2\mapsto f]}_f \;\middle|\; f \in D_{\langle s,\langle e,t\rangle\rangle} \right\}$$
$$= \left\{ \lambda w_s.\lambda x_e. \begin{array}{l} x \text{ cites some papers } p \text{ such that} \\ p \text{ is by } h[2 \mapsto f](8)(w) \text{ and } f(w)(p) = 1 \end{array} \right\}$$

But then we've lost Jackendoff! $h[2 \mapsto f](8)(w)$, which is identical to $h(8)(w)$. This value may be Jackendoff, but could well be someone else, depending on what h is (if $8 \in \mathrm{dom}(h)$ at all; otherwise it will be uninterpretable).

So the predicted assertive component of $[\![\mathtt{VP}_1]\!]^{g,h}$ will be dependent on h, as in (71), which is not correct.

(71) x read a textbook t in w s.t. for each $f \in g(\mathtt{C})$ t does not cite papers p in w s.t. p is by $h(8)(w)$ in w and $f(w)(p) = 1$

(roughly, x read a textbook in w that does not cite multiple papers by $h(8)(w)$)

2. C' will be required to be a subset of the following set:

(72) $$\left\{ [\![\mathtt{VP}_1]\!]^{g,h[8\mapsto x]}_f \;\middle|\; x \in D_{\langle s,e\rangle} \right\}$$

Let us consider one of these alternatives where $x = \lambda w_s.$ Chomsky.

We want this alternative to be (73).

(73) $\lambda w_s.\lambda x_e$: [PRESUPPS on $g(\mathtt{C})$] and some textbook t cites one paper by Chomsky in w . x read a textbook t in w s.t. for each $f \in g(\mathtt{C})$ t does not cite papers p in w s.t. p is by Chomsky in w and $f(w)(p) = 1$

This time, we correctly interpret $\textit{Jackendoff}_{F_8^{\langle s,e\rangle}}$ in the computation of focus alternatives for *only*.

(74) $$\left[\!\left[\mathtt{Jackendoff}_{F_8^{\langle s,e\rangle}} \right]\!\right]^{g,h[8\mapsto \lambda w_s.\ \mathrm{Chomsky}][2\mapsto f]}_f = \lambda w_s.\ \mathrm{Chomsky}$$

But we failed to derive the correct prejacent presupposition for *only*. Let t be the relevant textbook. The prejacent presupposition of *only* will be:

(75) $$\left[\!\left[[_{\mathtt{VP}_2}\ \mathtt{cites\ one}_{F_2^{\langle s,\langle e,t\rangle\rangle}}\ \mathtt{paper\ by\ Jackendoff}_{F_8^{\langle s,e\rangle}}] \right]\!\right]^{g,h[8\mapsto \lambda w_s.\ \mathrm{Chomsky}][2\mapsto f]}_o (w)(t)$$
$= 1$ iff t cites (at least) one paper by Jackendoff in w

4.2.3 *Local selectivity*

Wold 1996 identifies the cause of the problem in the bi-dimensionality of the system.

- To solve the first problem, we don't want to compute the focus semantic value of the entire VP_2, but just a part of it. The bi-dimensional system is too crude to do so.
- To solve the second problem, we want to change the ordinary semantic value of VP_2, but by assumption the ordinary semantic value cannot vary.

Wold proposes to get rid of bi-dimensionality.[4]

- A focus semantic index is dormant, when h is not defined for it.[5]

(76) $$[\![\alpha_{F^\tau_i}]\!]^{g,h} = \begin{cases} h(i) & \text{if } i \in \text{dom}(h) \\ [\![\alpha]\!]^{g,h} & \text{otherwise} \end{cases}$$

- $\sim$ only makes i defined for h temporarily when it applies, so that this does not carry over to further computation.

(77) $$\left[\!\!\left[\begin{array}{c} \text{XP} \\ \sim^\tau_i \ \ \text{C} \end{array} \right]\!\!\right]^{g,h} = \lambda w_s : \begin{array}{l} g(\text{C}) \subseteq \left\{ [\![\text{XP}]\!]^{g,h[i \mapsto x]} \;\middle|\; x \in D_\tau \right\} \text{ and} \\ [\![\text{XP}]\!]^{g,h} \in g(\text{C}) \text{ and } |g(\text{C})| \geqslant 2 \end{array} . [\![\text{XP}]\!]^{g,h}$$

This solves the above two problems.

1. In the computation of the prejacent of *also*, $[\![\text{VP}_1]\!]^{g,h}$, h will stay unmodified. Assuming that h is not defined for the index on 'Jackentoff$_{F^{\langle s,e\rangle}_8}$', $\left[\!\!\left[\text{Jackentoff}_{F^{\langle s,e\rangle}_8} \right]\!\!\right]^{g,h}$ will always be λw_s. Jackendoff.

(78) $$\left[\!\!\left[[_{\text{VP}_2} \text{ cites one}_{F^{\langle s,\langle e,t\rangle\rangle}_2} \text{ paper by Jackendoff}_{F^{\langle s,e\rangle}_8}] \right]\!\!\right]^{g,h[2 \mapsto f]}(w)(t) = 1$$
iff t cites some papers p by Jackendoff in w such that $f(w)(p) = 1$

2. In an alternative to VP_1, e.g., $[\![\text{VP}_1]\!]^{g,h[1 \mapsto \lambda w_s.\ \text{Chomsky}]}$, 'Jackentoff$_{F^{\langle s,e\rangle}_8}$' will always denote λw_s. Chomsky (assuming there's no rebinding), and we will derive the correct prejacent presupposition of *only* for this alternative:

(79) $$\left[\!\!\left[[_{\text{VP}_2} \text{ cites one}_{F^{\langle s,\langle e,t\rangle\rangle}_2} \text{ paper by Jackendoff}_{F^{\langle s,e\rangle}_8}] \right]\!\!\right]^{g,h[8 \mapsto \lambda w_s.\ \text{Chomsky}]}(w)(t) = 1$$
iff t cites (at least) one paper by Chomsky in w

4.3 Judgments of multiple foci

If sentences with multiple focus sensitive operators associating with multiple foci in the same domain are grammatical, it seems necessary to make rather substantial changes to the Roothian bi-dimensional framework.

[4] The following presentation of Wold's 1996 differs from his own in some technical details.
[5] Wold 1996 works with just one assignment, but we keep it closer to Kratzer 1991.

- The theory of Wold 1996 is selective and uni-dimensional.
- We'll discuss a less complicated alternative (although still selective) in the next lecture.

Beck & Vasishth 2009 cast doubt on the acceptability of relevant examples in the literature (see also Erlewine 2025).

(80) John only introduced Sue$_F$ to Fred.

John also only [$_{VP}$ introduced Sue$_F$ to Bill$_F$].

(Wold 1996: p. 320)

(81) We only recovered the diary entries that Marilyn$_F$ made about John.

We also only recovered the diary entries that Marilyn$_F$ to Bobby$_F$.

(Rooth 1996: (45))

As von Fintel 1994: fn. 44 points out, not all combinations of focus sensitive particles seem to be as acceptable.

(82) A: I know that John drank water at the party. What else did he drink?
B$_1$: Besides water, he only drank [carrot juice]$_F$.
B$_2$: #He only also drank [carrot juice]$_F$. (Beck & Vasishth 2009: p. 164)

They ran an acceptability judgment experiment:

- Three focus-sensitive operators (*only, also, even*), crossed with four operators (*only, also, even, nobody*). *Nobody* is not relevant for us, so I will omit the items with it.

(83) *also-only*
A: You only showed the photos to Carol.
B: I also only showed the photos to Robin.

(84) *also-even*
A: You showed the pictures to Carol. You even showed her the movie.
B: Yes. I also even showed the movie to Robin.

(85) *only-also*
A: You showed the photos to Carol. Did you also show the movie to her?
B: No. I only also showed the movie to Robin.

(86) *only-even*
A: You showed the photos to Carol's assistant. You even showed the movie to him.
B: No. I only even showed the movie to Carol.

(87) *even-only*
A: You only showed the photos to Carol.
B: Right. I even only showed the photos to the boss.

(88) *even-also*
A: You showed the photos to Carol.
B: Yes. I also showed the movie to her. I even also showed the movie to Carol's assistant.

- Control items had the same meaning as the targets but without multiple foci. They were judged in a separate session from the target items (half the participants did the targets first, the other half the controls first) and there was at least 4 weeks between the two sessions.

  (89) *also only*
  A: You only showed the photos to Carol.
  B: Another person I only showed the photos to is Robin.

  (90) *also-even*
  A: You showed the pictures to Carol. You even showed her the movie.
  B: Yes. Another person I even showed the movie to is Robin.

  (91) *only-also*
  A: You showed the photos to Carol. Did you also show the movie to her?
  B: No. The only person that I also showed the movie to is Robin.

  (92) *only-even*
  A: You showed the photos to Carol's assistant. You even showed the movie to him.
  B: No. The only person I even showed a movie to is Carol.

  (93) *even-only*
  A: You only showed the photos to Carol.
  B: Even the boss, I only showed the photo to.

  (94) *even-also*
  A: You showed the photos to Carol.
  B: Yes. I also showed the movie to her. Even her assistant I showed the photos to as well.

- 20 filler sentences, 4 of which were ungrammatical (which received low ratings, $1.07 - 1.50$).

- Written context description at the beginning.

  (95) A and B are detectives in the San Francisco police force. They are building a case against the well-known director of a local bank. Their main evidence consists of a set of photographs and a video showing the suspect in incriminating circumstances. Their assistants are trying to find eye witnesses to the events taking place in the pictures. Unfortunately, there has been a leak to the press. A and B are trying to find out how the information could have gotten out.

- The dialogues were pre-recorded and auditorily presented (no visual presentation, it seems).

- Acceptability on a 4-point Likert scale (1 = completely unacceptable, 4 = completely acceptable).

- 16 native speaker of English (in Berlin). 12 from the USA, 4 from the UK.

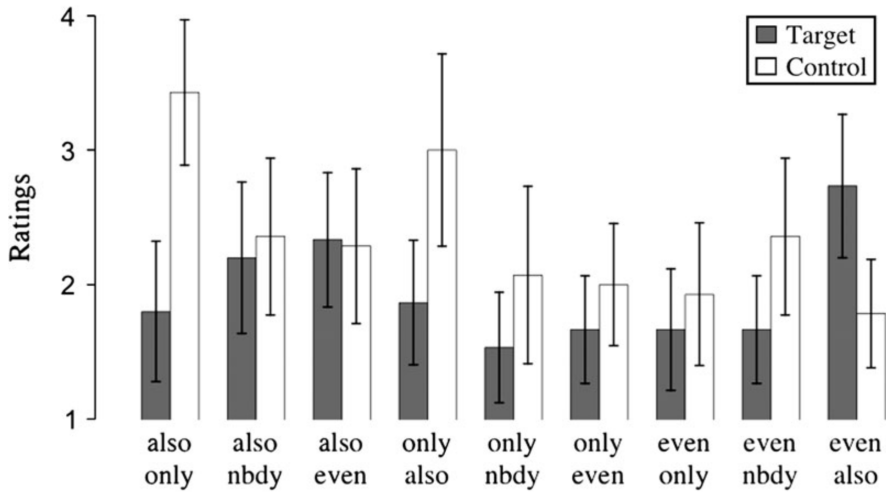


Figure 1 Mean ratings for target and control conditions for the nine pairs. A rating of 1 corresponds to the judgement "completely unacceptable" and the rating 4 corresponds to the judgement "perfectly acceptable".

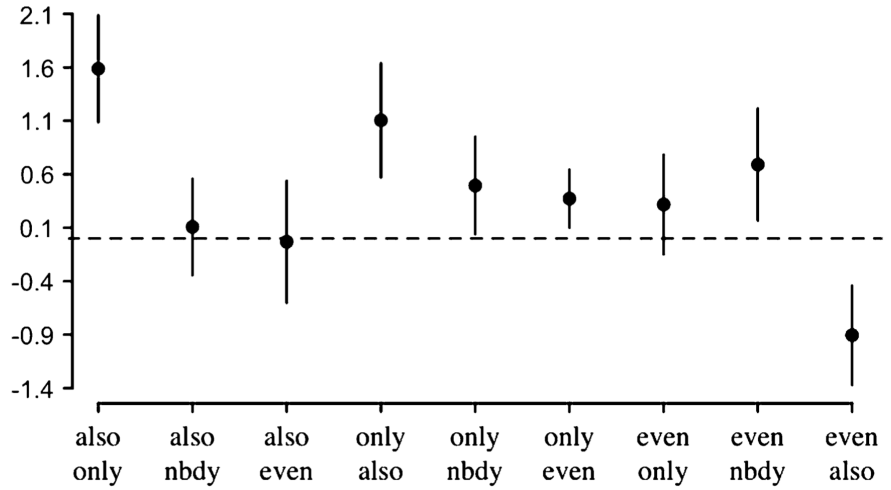


Figure 2 Estimated coefficients and 95% confidence intervals for the nine target–control comparisons; the regression was carried out using linear mixed-effectLMEs models, with the target–control as a fixed effect and participants as random effect. The interpretation of the graph is discussed in the text.

- For *even-also*, Beck & Vasishth 2009: p. 172 write: "upon closer examination, it was found that the control had an incorrect name in the critical sentence (rendering it inappropriate in the context in which it was presented)", but I think the mistake is that *the photos* should be *the movie*.
- They conclude that overall multiple foci in the relevant configuration are not acceptable.

Some potential issues

- It seems that no correction for multiple comparisons was performed, so we should take the 95% CIs with a grain of salt.
- *Also-even, only-even even-only* are not so different from the controls.
- The control of *even-also* might not be interpretable, but it's notable that the target is pretty good.

- The controls are generally bad too, so it's not so clear what low ratings mean. Potentially A's utterance and the non-target parts of B's utterance mattered.
- Part of the unacceptability might be prosodic in nature; it's unclear if the difference from the control items is about grammaticality/interpretability.
- Only one item per condition.

We might want to play with the structure a bit more.

(96) John read a textbook that only cites one$_F$ paper by Chomsky.
   a. He also [read a textbook that only [cites one$_F$ paper by Jackendoff$_F$]].
   b. He [read [a textbook that only cites [one$_F$ paper by Jackendoff$_F$]]] too/as well.

(97) He gets angry, if his students only cite recent$_F$ papers by Chomsky.
He even [gets angry, if his students only [cites recent$_F$ papers by Jackendoff$_F$]]

(98) Students are generally very active in class; they answer questions, and also ask questions. Except for one guy, who never speaks. In fact, he gets angry, when I tell him$_F$ to also answer questions in class.
He even [gets angry, when I tell him$_F$ to also ask$_F$ questions].

(99) You should call Uncle John, but don't invite him to our place this weekend. I like him but he's been really busy.
And I won't complain, if you also call your brother$_F$.
I will only [complain, if you also [invite$_F$ the bastard$_F$]].

Erlewine 2025 claims that the only possible interpretation for examples like these is the one predicted by the unselective Roothian system with $\approx$. E.g., according to him, (96), to the extent it is acceptable, should entail that the textbook in question cites only one paper in total, and that is by Jackendoff.

References

Bade, Nadine. 2021. On the scope and nature of *Maximize Presupposition*. *Language and Linguistics Compass* 15(6). e12416. https://doi.org/10.1111/lnc3.12416.

Beck, Sigrid & Shravan Vasishth. 2009. Multiple focus. *Journal of Semantics* 26(2). 159–184. https://doi.org/10.1093/jos/ffp001.

Büring, Daniel. 2015. Unalternative Semantics. In Sarah D'Antonio, Mary Moroney & Carol Rose Little (eds.), *Proceedings of SALT 25*, 550–575. https://doi.org/10.3765/salt.v25i0.3634.

Büring, Daniel. 2016. *Intonation and meaning*. Oxford: Oxford University Press. https://doi.org/10.1093/acprof:oso/9780199226269.001.0001.

Charlow, Simon. 2008. Free and bound pro-verbs: A unified treatment of anaphora. In Tova Friedman & Satoshi Ito (eds.), *Proceedings of SALT 18*, 198–211. https://doi.org/10.3765/salt.v18i0.2487.

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

Erlewine, Michael Yoshitaka. 2014. *Movement out of focus*. Massachusetts Institute of Technology dissertation.

Erlewine, Michael Yoshitaka. 2025. Focus intervention, multiple association, and the unity of focus and *wh* alternatives. *Linguistics and Philosophy* 48. 527–571. https://doi.org/10.1007/s10988-025-09435-x.

von Fintel, Kai. 1994. *Restrictions on Quantifier Domains*. University of Massachusetts, Amherst Ph.D. dissertation.
Fox, Danny. 1999. Reconstruction, binding theory, and the interpretation of chains. *Linguistic Inquiry* 30(2). 157–196.
Fox, Danny & Roni Katzir. 2011. On the characterization of alternatives. *Natural Language Semantics* 19(1). 87–107. https://doi.org/10.1007/s11050-010-9065-3.
Heim, Irene. 1991. Artikel und Definitheit. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantik: Ein internationales Handbuch der zeitgenössischen Forschung*, 487–535. Berlin: de Gruyter.
Jäger, Gerhard. 1999. Focus without variables: A multi-modal analysis. In Kimary N. Shahin, Susan Blake & Eun-Sook Kim (eds.), *Proceedings of WCCFL 17*, 317–331. Stanford, CA: CSLI.
Kadmon, Nirit. 2001. *Formal pragmatics*. Oxford: Blackwell.
Kratzer, Angelika. 1991. The representation of focus. In Arnim von Stechow & Dieter Wunderlich (eds.), *Semantik: ein internationales Handbuch der zeitgenössischen Forschung*, 825–834. Berlin: De Gruyter. https://doi.org/10.1515/9783110126969.10.825.
Krifka, Manfred. 1991. A compositonal semantics for multiple focus constructions. In Steven K. Moore & Adam Zachary Wyner (eds.), *Proceedings of SALT 1*, 127–158. https://doi.org/10.3765/salt.v1i0.2492.
Krifka, Manfred. 2006. Association with focus phrases. In Valéria Molnár & Susanne Winkler (eds.), *The architecture of focus*, 105–136. Berlin: Mouton de Gruyter. https://doi.org/10.1515/9783110922011.105.
Marty, Paul P. 2017. *Implicatures in the DP domain*. Massachusetts Institute of Technology dissertation.
Panizza, Daniele & Yasutada Sudo. 2020. Minimal sufficiency with covert *even*. *Glossa* 5(1). 1–25. https://doi.org/10.5334/gjgl.1118.
Percus, Orin. 2006. Antipresuppositions. In Ayumi Ueyama (ed.), *Theoretical and empirical studies of reference and anaphora: toward the establishment of generative grammar as an empirical science*, 52–73. Japan Society for the Promotion of Science.
Rooth, Mats. 1985. *Association with focus*. University of Massachusetts, Amherst dissertation.
Rooth, Mats. 1996. Focus. In Shalom Lappin (ed.), *The handbook of contemporary semantic theory*, 271–297. Blackwell.
Schwarzschild, Roger. 1999. GIVENness, AvoidF and other constraints on the placement of accent. *Natural Language Semantics* 7(2). 141–177.
Singh, Raj. 2011. *Maximize Presupposition!* and local contexts. *Natural Language Semantics* 19(2). 149–168. https://doi.org/10.1007/s11050-010-9066-2.
Spector, Benjamin & Yasutada Sudo. 2014. Presupposed ignorance and exhaustification: how scalar implicatures and presuppositions interact. Ms., Institut Jean Nicod and UCL (under review).
Spector, Benjamin & Yasutada Sudo. 2017. Presupposed ignorance and exhaustification: How scalar implicatures and presuppositions interact. *Linguistics and Philosophy* 40(5). 473–517. https://doi.org/https://doi.org/10.1007/s10988-017-9208-9.
Wold, Dag E. 1996. Long distance selective binding: the case of focus. In Teresa Galloway & Justin Spence (eds.), *Proceedings of SALT 6*, 311–328. https://doi.org/10.3765/salt.v6i0.2766.