Focusing Particles & Ellipsis Resolution

Johan Bos

Universität des Saarlandes

Report 61
December 1994
Gehört zum Antragsabschnitt: 8.6.1, 8.6.3, 8.7.1, 8.7.2, 8.9.6, 8.9.7

Abstract

We present a semantic framework which integrates a compositional version of Discourse Representation Theory, Van der Sandt’s presupposition theory, and a treatment of focus in the style of Rooth’s Alternative Semantics. We will discuss the semantics of focusing particles like too and only within this framework. The function of these particles is maintaining coherence in discourse or dialogue. This explicitly allows them to introduce contrast between phrases by means of presupposition. Of our interest is the interaction between focusing particles and elliptical phrases. In particular, we pay attention to cases of VP-ellipsis in English. It turns out that the interpretation of focusing particles naturally accounts for the occurrences of sloppy and strict readings in VP-ellipsis. This is because their presupposition adds contrast between the source and target clause. This feature distinguishes the approach sketched in this paper from known approaches to ellipsis, which disregard the function of focusing particles.

Thanks

I would like to thank Claire Gardent, Julia Heine, Andrea Kowalski, Manfred Pinkal, and the participants of the Focus & NLP workshop for their valuable comments and contributing discussions. All remaining errors are mine.
## Contents

1  Introduction  
2  Rooth’s Alternative Semantics  
3  DRT-framework  
4  Presupposition as Anaphora  
5  Focus in DRT  
6  VP-ellipsis  
7  Conclusion  

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Introduction</td>
<td>3</td>
</tr>
<tr>
<td>Rooth’s Alternative Semantics</td>
<td>4</td>
</tr>
<tr>
<td>DRT-framework</td>
<td>5</td>
</tr>
<tr>
<td>Presupposition as Anaphora</td>
<td>6</td>
</tr>
<tr>
<td>Focus in DRT</td>
<td>7</td>
</tr>
<tr>
<td>VP-ellipsis</td>
<td>9</td>
</tr>
<tr>
<td>Conclusion</td>
<td>11</td>
</tr>
</tbody>
</table>
1 Introduction

The presence of focus-sensitive particles often facilitates interpretation of VP-ellipsis in English (and e.g. gapping in German). Consider the difference in acceptability between cases where these particles appear (1, 2, and 3) and cases where they don’t (4). As their name suggests, focus-sensitive particles (or generally: focalizers) associate with focus. They normally do not add anything to the meaning of the sentence, but rather “judge” whether the sentence in which they appear is acceptable in a given context or not. Focusing particles do this by means of presupposition. They introduce a presupposition sensitive to their foci. A focussed constituent introduces an alternative set relative to the context and is normally prosodically stressed.¹

The particle too in the second sentence of (1) presupposes that someone distinct from Bill did whatever Bill did. Similar presuppositions are obeyed in the second sentences of (2) and (3). To satisfy these presuppositions we are restricted in the interpretation of the elided VP. We can only succeed in this process when firstly, revised his paper is chosen as antecedent, and secondly, the pronoun in the target phrase is allowed to have either John or Bill as its antecedent, even if there are more suitable antecedents available in discourse.

It is argued that it is presupposition that makes (1) sound natural, and that the absence of a similar presupposition causes (4) to sound infelicitous.

(1) John revised his paper. [Bill]$_f$ did, too.
(2) John revised his paper. [Bill]$_f$ did not.
(3) John revised his paper. Only [John]$_f$ did.
(4) ? John revised his paper. Bill did.

In the examples above contrast is explicitly introduced by particles. However, contrasting phrases in discourses like (5) do not necessarily require particles.² Contrast here can be defined as a structural discourse relation between two phrases, which meanings only differ with respect to the constituent in focus. This phenomenon, although related to the cases in (1-3), will not be discussed here.

(5) [John]$_f$ revised his paper, and [Bill]$_f$ did.

This paper specially contributes to the understanding of the role focusing particles play in ellipsis resolution. Our aim is to combine Focus theory (Rooth, 1992a) with

¹Focussed constituents are indicated by a lower f subsctipt in examples.
²Some participants of the Focus and NLP conference pointed out to me that is is somehow strange that contrasting phrases require a discourse connector like and or but. However, as Nicholas Asher pointed out to me, this observation is violated when a sequence of elliptical utterances is put into contrast: John revised his paper. [Tom]$_f$ did, and [Bill]$_f$ did.
Presupposition theory (Van der Sandt, 1992) in a dynamic semantic framework (Groenendijk and Stokhof, 1991; Kamp and Reyle, 1993). After briefly reviewing Alternative Semantics a compositional version of Discourse Representation Theory is introduced. We will give an analysis of focusing particles within this framework, and then give an account to the interpretation of VP-ellipsis.

2 Rooth’s Alternative Semantics

This section briefly summarizes Rooth’s theory of association with focus (Rooth, 1992a). Rooth’s focus interpretation theory (Restricted) Alternative Semantics is characterized by adding a “focus semantic value” (\([\phi]^f\)) to the normal semantic value (\([\phi]^0\)). For example:

\[
\begin{align*}
\text{[John likes [Mary]}_f{]}^0 & = \text{like(john,mary)} \\
\text{[John likes [Mary]}_f{]}^f & = \text{the set of propositions: John likes x}
\end{align*}
\]

The function of focus is to generate alternatives. To relate a focused constituent to its alternatives, a focus-sensitive operator \(\sim\) is introduced which has the following contraints:

**set case:** \(\phi \sim C\) presupposes that \(C\) is a subset of \([\phi]^f\) and contains both \([\phi]^0\) and an element distinct from \([\phi]^0\).

**individual case:** \(\phi \sim \psi\) presupposes that \(\psi\) is an element of \([\phi]^f\) and distinct from \([\phi]^0\).

More informally, \(\sim\) handles the identification of a focused phrase with a particular semantic object. In some cases this “antecedent” is given by discourse, in other cases it is pragmatically constructed. Rooth uses the *individual case* for the treatment of contrasting phrases, and the *set case* for a number of other phenomena, including focusing adverbs.

The representation of (1) in Rooth’s framework is \([\text{Bill}_f{]} did \sim C\), where focus is interpreted at the sentence level. From this logical form follows that it is required that \(C\) is of the form “x did”. This requirement is met by the antecedent sentence *John revised his paper*.

Summarizing, what focus does according to Alternative Semantics is introducing a variable which is anaphoric to a contrasting element, or a set of elements. What we will do in the following sections is settle a version of Rooth’s theory in a dynamic semantic framework, which enables us to put alternative semantics together with a theory of anaphora and presupposition.
3 DRT-framework

In this section we describe a compositional DRT-like framework. Unlike traditional DRT (Kamp, 1981), no distinction is made between DRSs and conditions (as for instance in (Van Eijck and Kamp, 1994)). This yields a system that is more akin to dynamic predicate logic (Groenendijk and Stokhof, 1991) on the one hand, but still favours the use of discourse markers on the other. The syntax is defined as follows:

DRS Syntax

Syn1: If x is a variable, x is a DRS

Syn2: If P is an n-place predicate, x₁, ..., xₙ are variables, P(x₁,...,xₙ) is a DRS

Syn3: If Φ and Ψ are DRSs, then Φ ∧ Ψ, Φ ∨ Ψ, Φ ⇒ Ψ and ¬Φ are DRSs.

The first rule describes the syntax of discourse markers, the second rule basic conditions, and the third rule complex conditions. DRSs are interpreted in extensional first order models, consisting of the set of individuals D, and an interpretation function F that assigns sets of n-tuples to the n-place predicates.³ An assignment g is a total function that assigns an individual to each variable. h[x]g means that h differs from g at most in the value h assigns to x. The interpretation of a DRS is a set of pairs of assignments: a pair ⟨g, h⟩ can be viewed as an “input-output” pair, and is part of the interpretation of a DRS when h is a possible result of the evaluation of this DRS with respect to g. Formalized, the semantics is defined according to the following clauses:

DRS Semantics

Sem1: [ x ] = {⟨g, h⟩| h[x]g}

Sem2: [ P(x₁,...,xₙ) ] = {⟨g, h⟩| g = h & < h(x₁), ..., h(xₙ) > ∈ F(P) }

Sem3: [ Φ ∧ Ψ ] = {⟨g, h⟩| there is a k such that ⟨g, k⟩ ∈ [Φ] & ⟨k, h⟩ ∈ [Ψ] }

Sem4: [ Φ ∨ Ψ ] = {⟨g, h⟩| g = h & there is a k such that ⟨h, k⟩ ∈ [Φ] or there is a k such that ⟨h, k⟩ ∈ [Ψ] }

Sem5: [ Φ ⇒ Ψ ] = {⟨g, h⟩| g = h & for every k: ⟨h, k⟩ ∈ [Φ] it is the case that there is a j: ⟨k, j⟩ ∈ [Ψ] }

Sem6: [ ¬Φ ] = {⟨g, h⟩| g = h & there is no k: ⟨h, k⟩ ∈ [Φ] }

Note that some of these rules are interpreted as tests (the pairs of input and output assignments are required to be identical) whereas others perform a context update.⁴

So far we didn’t define any syntax and semantics for anaphora and presupposition. Obviously, we require means to represent anaphoric links and presuppositions.

³To make things not too complicated, eventualities are interpreted as individuals.

⁴The interested reader is referred to Groenendijk & Stokhof (1991) for an introduction to dynamic semantics.
in order to interpret focusing phenomena. In the next section we will integrate a version of Van der Sandt's theory of presupposition, and introduce a single operator to handle both anaphora and presupposition.

4 Presupposition as Anaphora

Presuppositions are claimed to be anaphoric expressions (Van der Sandt, 1992). They behave essentially the same as pronouns, but they possess more descriptive content which restricts the choice of antecedent. It is this semantic content that allows accommodation of an antecedent if none are available.

To simulate Van der Sandt's A-Structures, we introduce the "alfa-operator" $\alpha(x_1,x_2)$: $\Phi$, where $\Phi$ is a DRS and contains the anaphoric (or presuppositional) material.\(^5\) The operator $\alpha$ has two additional arguments: $x_1$ which is the principle anaphoric variable, and $x_2$, a discourse marker which plays the role of antecedent. The second argument slot is optional and obtained after resolution, so we end up with two different syntactic representations:

**Syn4:** If $x_1, x_2$ are variables, $\Phi$ is a DRS, $\alpha(x_1):\Phi$ and $\alpha(x_1, x_2):\Phi$ are DRSs

A simple example serves to show how DRSs actually look like and how anaphora and presupposition are represented. Consider (7), in which the definite description in the second sentence anaphorically presupposes "a man" in the first sentence.

\begin{equation}
\begin{aligned}
x_1 \land betty(x_1) \land x_2 \land man(x_2) \land like(x_1, x_2) \land \alpha(x_1, x_2) : (\text{man}(x_4)) \land se(x_4)
\end{aligned}
\end{equation}

The semantic interpretation rules of $\alpha$ are consequently defined as:

**Sem7:** $[\alpha(x):\Phi] = \{\langle g, h \rangle|\exists k: k[x]g \& \langle k, h \rangle \in [\Phi]\}$

**Sem8:** $[\alpha(x_1, x_2):\Phi] = \{\langle g, h \rangle| g = h \& [x_1]_g = [x_2]_g\}$

The first interpretation rule of $\alpha$ (where there is no antecedent specified) can be viewed as a means to *underspecify* anaphoric expressions. In terms of dynamic semantics, this interpretation is externally dynamic. Clause **Sem8** is the usual way of interpreting anaphoric links. The anaphoric content in $\Phi$ does not really matter anymore and a static interpretation is satisfactory.

We will not pay any attention to accommodation here. A Van der Sandtian accommodation process can be adopted in this framework by allowing the content $\Phi$ of $\alpha$ to move (or alternatively, copy) to accessible discourse levels.

---

\(^5\)We deviate from Van der Sandt on two points. Firstly, anaphoric structures are embedded in the DRS as a so-called alfa-DRS. Therefore a DRS is not viewed as a triple of universe, conditions and anaphoric material. Secondly, $\alpha$ has a distinguished marker which is the principle anaphoric marker. This technical difference solves the problem Van der Sandt’s account has with indeterminates that occur in anaphoric structures.
5 Focus in DRT

In the previous section we introduced the operator $\alpha$ to cover both anaphora and presupposition. Recall that focused constituents typically refer indirectly to alternatives, which depend very much on the context. Clearly, we are not able to use the apparatus introduced for presupposition in the previous section for the representation of focused constituents. We are compelled to extend our framework with an operator that behaves quite similarly to Rooth's $\sim$ operator.

This is the operator $\pi$, which contextually determines alternatives for an object described by a DRS, excluding the focused expression itself. A DRS $\pi(x,Z) : \Phi$ is informally defined as $x$ being a member of the alternatives described by $\Phi$, $Z$ the set of these alternatives, and $\Phi$ being a DRS that restricts the choice of possible alternatives. Like $\alpha$, $\pi$ can also accommodate an antecedent if discourse does not provide one, although this certainly is a process quite distinct from presuppositional accommodation (we will not discuss it in this paper). The definition of $\pi$'s syntax is:

**Syn5:** If $x$ is a variable, $Z$ a non empty set of discourse markers, $\Phi$ is a DRS, $\pi(x) : \Phi$ and $\pi(x,Z) : \Phi$ are DRSs.

For the purposes of this paper we assume that $\pi$-DRSs are assigned to constituents that bear prominent stress. We also assume a typed lambda-calculus for DRSs (cf. (Muskens, 1993; Bos et al., 1994)). Semantic composition corresponds then to:

\[
\begin{align*}
8 & \quad [\text{Mary}]_f = \lambda P. \pi(x_4) : (\text{mary}(x_4)) \land P(x_4) \\
9 & \quad \text{likes} [\text{Mary}]_f = \lambda y. \lambda e. \pi(x_4) : (\text{mary}(x_4)) \land \text{like}(e,y,x_4) \\
10 & \quad \text{John likes} [\text{Mary}]_f = \lambda e. x_3 \land \text{john}(x_3) \land \pi(x_4) : (\text{mary}(x_4)) \land \text{like}(e,x_3,x_4)
\end{align*}
\]

The idea is now that focalizers can assign a focus feature to $\pi$ operators which cause a different semantic interpretation. To put it differently, we make the interpretation of $\pi$ either focus sensitive or not. In the former, $\pi$ is featured with an attribute $f$. Assigning a focus feature to one or more $\pi$ DRSs is notated as $\Phi^f$, indicating that some $\pi$-DRS $\in \Phi$ has assigned to it the focus feature. Hence, we end up with three different interpretation rules for the $\pi$ operator:

---

6This is an essential difference to the twiddle operator of Rooth’s theory. Mats Rooth pointed out to me that this feature of excluding alternatives could cause problems by the interpretation of multiple foci in examples like: *John only introduced [Bill]$^f$ to [Sue]$^f*.* Among the alternatives, the proposition “John introduced Tom to Sue” would be excluded in our framework.

7This wouldn’t help us to solve the focus projection problem, ambiguities in stress marking. Moreover, focalizers are able to have elliptical constituents as focus: I could ask my colleague, at the end of the day: “[Also]$^f$ tired?”, where the presupposition introduced by this sentence is that someone distinct from the hearer (for instance, the speaker) is tired. These problems are not addressed in this paper.
The first clause corresponds to underspecified interpretation. In case alternatives are determined in interpretation amounts to a test whether one of the alternatives equals the focused constituent. If the focus feature is not assigned interpretation is done in the ordinary way (Sem11).

The rest of this section will be concerned with the semantics of the particles too and only. They serve to exemplify the function of π-conditions in our framework.

The particle too in the example below (11) introduces a presupposition that John likes someone distinct from Mary. The semantics of too is stated in (12). It takes its argument (a DRS of type “eventuality”), turns it into a proposition and asserts it, and presupposes this argument with the focus feature assigned to it.

(11) John likes [Mary]f too.
(12) too: \( \lambda \Phi, e_1 \wedge \Phi(e_1) \land a(e_2) : \Phi^f(e_2) \)

In a discourse where it is asserted that John likes Betty, Betty corresponding to marker x8, the presuppositional part of (12) would look like:

(13) \( a(e_6) : (x_3 \wedge john(x_3) \wedge \pi^f(x_4, \{x_8\}) : mary(x_4) \wedge like(e_6, x_3, x_4)) \)

Note that too normally is associated with only one focused item. This probably explains the infelicity of too in John likes Mary, and [Bill]f likes [Sue]f too.8

In (14), the focusing adverb only introduces the presupposition John likes Mary, and asserts the proposition that there is no other person distinct from Mary that John likes.

(14) John only likes [Mary]f.
(15) only: \( \lambda \Phi, a(e_1) : \Phi(e_1) \land \neg(e_2 \wedge \Phi^f(e_2)) \)

In this case, without any context, it as not that clear what the alternative set might be.9 Assume that Nancy and Betty are salient in discourse, introduced with discourse markers x8 and x9 respectively, the assertional part introduced by only would be:

(16) \( \neg(e_6 \wedge x_3 \wedge john(x_3) \wedge \pi^f(x_4, \{x_8, x_9\}) : mary(x_4) \wedge like(e_6, x_3, x_4)) \)

Another option is a form of accommodation, in which the antecedent for the π-expression would be created in discourse, guided by domain specific information and world knowledge. A third option is to leave the antecedent underspecified.

8However, John likes Mary and Mary likes John too sounds acceptable to me.
9See Blok & Eberle, this volume, for addressing this problem.
6 VP-ellipsis

A lot of attention has been paid to the treatment of VP-ellipsis in the literature. Strikingly interesting is the appearance of sloppy and strict readings in discourses like (17), a standard example:

(17) John hates his dog. [Bill]$_T$ does, too.

Interpretation of (17), assuming that John and his corefer, results in two possible readings: Bill hates John’s dog, which is called the strict reading, and Bill hates Bill’s dog, the sloppy reading. To account for this observation, quite a number of analyses have been proposed that treat the source clause as ambiguous: either referential or bound interpretation of the pronoun. Resolution of the elided VP then boils down to choosing one of either interpretation. These approaches, also termed identity of relation approaches, nonetheless face serious difficulties because they have to postulate a “wild” ambiguity in the source clause when more than one pronoun is present (Dalrymple et al., 1991).

Our analysis abandons the assumption that the source clause is ambiguous in cases like (17). Instead, we are in favour of a two-level architecture for the interpretation of VP-ellipsis:

1. reconstruct the elided VP from a potential antecedent
2. resolve anaphoric relations in new context and perform a “presuppositional check”

Level one is a copying and renaming operation (for example by Asher’s “C-Abstraction” (Asher, 1993)). We will not consider this level in any detail here. Required is that reconstruction of elided material results in a target DRS with anaphoric material underspecified. Level two consists of anaphora resolution and the interpretation of pragmatic contrast or focusing particles (as for the purposes of this paper, we limit ourselves to focusing particles). Consequently, the analysis of (17) after reconstructing the elliptical material actually is the analysis of (18):

(18) John hates his dog. [Bill]$_T$ hates his dog, too.

That is, ellipsis in our view is nothing but extreme phonetic reduction. This is contra Fiengo & May (Fiengo and May, 1992) who put forward the thesis that the effect of ellipsis is eliminative. Also, Rooth shows some possible counterevidence against this thesis (Rooth, 1992b).

We will discuss resolution of example (17) in detail. The DRS of the source clause is (19). The DRS for the target clause Bill does, after step (1) of VP-ellipsis resolution is (20):

(19) $e_1 \land x_1 \land john(x_1) \land \alpha(x_3, x_1) : (x_2 \land dog(x_2) \land af(x_2, x_3)) \land hate(e_1, x_1, x_2)$
Anaphora resolution leaves two possible antecedents for $x_9$, corresponding to the pronoun his, i.e., the object in focus ($x_7$) or John. In the first case, the ordinary DRS paraphrases that Bill hates Bill’s dog, and it is presupposed that there is some $x$ distinct from Bill that hates $x$’s dog. This results in a bound interpretation and justifies a sloppy reading for (17). In the other case, where $x_9$ is linked to the marker corresponding to John, the presupposition would obey that there is someone distinct from Bill that hates John’s dog. This represents the referential interpretation and causes a strict reading for (17).

Consider the “coach-examples” (21) and (22). Note that our analysis allows a strict and sloppy reading for (21), but only a strict reading for (22). There is no sloppy reading for (22) because the presupposition stemming from the particle too should be of the form $x$ thinks that $x$ will win, which cannot be satisfied (there is no antecedent) by the given context.

(21) John’s coach thinks he will win, and [Bill’s]$_f$ coach does too.

(22) John’s coach thinks he will win, and [Bill]$_f$ does too.

A further consequence of our analysis, is that when he in (21) corefers with John, the reading where Bill’s coach thinks he will win is excluded. Again, this is due to the fact that the presupposition that follows from this interpretation cannot be anaphorically linked to an existing entity. It should be noted then that binding is strongly preferable to accommodation of these types of presupposition.

Compare the analysis proposed in this section with the Bound Variable Principle (Logical Form) in Kratzer’s framework (Kratzer, 1991):

the phonological content of a pronoun may optionally be deleted if it is c-commanded by a co-indexed empty pronoun.

(23) I only said that [Sue]$_f$ thinks she is funny.

Sentence (23) allows three different logical forms, and two different p-sets for: the set $x$ thinks Sue is funny and the set $x$ thinks $x$ is funny.

In our framework we will only get the bound interpretation, unless there is a discourse marker corresponding to Sue outside (23) available, which then will result in a referential interpretation as well.

This might create genuine problems considering (24). This sentence has certainly both a bound and referential reading. The assertion of only in this example is either that there is no one, distinct from John that likes his dog, or that there is no one but John that likes John’s dog. Maybe surprisingly, this referential reading does not show up when the pronoun corefers with an indefinite description (25).

(24) Only John likes his dog.
Only a farmer likes his dog.

This can be explained by postulating that proper names, in contrast to indefinites, always accommodate to the main DRS, and hence are available for anaphoric reference. Holding on to this observation, there is no need to introduce a special principle in our framework that accounts for referential readings in examples like (23) and (24).

7 Conclusion

In this paper we discussed a dynamic semantic framework that treats anaphora, presupposition and focus. We were interested in the treatment of focusing particles. We gave a semantics of too and only, particles that respectively presuppose and assert information strongly dependent on the constituent in focus. Finally we showed how these focusing particles affect the interpretation of ellipsis and gave an account for VP-ellipsis resolution without using any additional artificial constraints.

References


