Exhaustivity in Dynamics Semantics; Referential and Descriptive pronouns *

Robert van Rooy
ILLC/University of Amsterdam
vanrooy@hum.uva.nl

Abstract. In this paper I argue that *anaphoric pronouns* should always be interpreted *exhaustively*. I propose that pronouns are either used *referentially* and refer to the speaker's referents of their antecedent indefinites, or *descriptively* and go proxy for the description recoverable from its antecedent clause. I show how this view can be implemented within a dynamic semantics, and how it can account for various examples that seemed to be problematic for the view that for all unbound pronouns there always should be a notion of exhaustivity/uniqueness involved. The uniqueness assumption for the use of singular pronouns is also shown to be important to explain what the discourse referents used in dynamic semantics represent.

Keywords: pronouns, dynamic semantics, indefinites, referential expressions

1. Introduction

Anaphora seem to be *definite* expressions. To quote Quine (1960, p. 113), " 'He', 'she' and 'it' are definite singular terms on a par with 'that lion' and 'the lion'." As a result it seems natural that personal pronouns should always receive an *exhaustive* interpretation. For singular pronouns this means that there should always be a notion of *uniqueness* involved. The most obvious way to account for this uniqueness implication is to say that the denotation of a pronoun is determined by the *definite* or *universal* description recoverable from its antecedent clause, and thus should be treated as E- or D-type pronouns (Evans, 1977, Neale 1990). Thus, in a sequence of the form *Some S are P. They are Q*, the denotation/reference of the pronoun *they* is determined by the description *(all) the S such that P*.[1] However, in discourses like (1), (2), and (3),

(1) Yesterday, *some men* came to the door. *They* were strangers

(2) *A man* is walking in the park. *He* is whistling,

(3) I bought *a sage plant*. I bought eight others along with *it*.

* This paper is based on parts of chapter 2 of my Ph.D thesis (Van Rooy, 1997).

[1] Evans (1977) claimed that the pronoun *rigidly refers to* (all) the S such that P, while Neale (1990) argues that the pronoun *goes proxy* for the description.

we don't want to say that the pronouns denote, or refer to, *all* the men who came at the door, in (1), the *unique* man who is walking in the park, in (2), and certainly not the *unique* sage plant I bought, in (3), respectively. In recent theories of discourse analysis like Discourse Representation Theory (DRT; Kamp, 1981, and Kamp & Reyle, 1993), File Change Semantics (FCS; Heim, 1982), Dynamic Predicate Logic (DPL; Groenendijk & Stokhof, 1991), Dynamic Montague Grammar (DMG; Groenendijk & Stokhof, 1990, and Chierchia, 1996), and most explicitly in Van der Does (1994), the assumption that pronouns always receive an exhaustive interpretation is given up for reasons like this. Accepting that pronouns should be used descriptively, in these theories it is assumed that the pronouns in (1), (2) and (3) are going proxy for *indefinite* descriptions.

In this paper, however, I want to defend the view that pronouns should always be interpreted exhaustively, and to suggest how this can be accounted for. To account for the exhaustive interpretation of pronouns I will use a two-way strategy; pronouns are either *referentially* used, referring to the *speaker's referents* of their antecedent indefinites, or they can be used *descriptively*, 'referring' to the *exhaustive* set of individuals denoted by the description recovered from the clause in which the antecedent occurs. Descriptive pronouns are clearly interpreted exhaustively according to this analysis. Also for referentially used pronouns, however, a notion of exhaustivity is involved: the latter will refer to the unique (set of all) speaker's referent(s) of their antecedent indefinites.

In sections 2 and 3 of this paper I will defend this position by rather straightforward empirical arguments and formal implementation. Section 2 deals with referential pronouns and section 3 with descriptive ones. The fourth section will be about a more *conceptual* issue; I will argue that my treatment improves on standard systems of dynamic semantics because it can give a more satisfying account of the status of discourse referents in information states.

2. Referential pronouns

2.1. Speaker's reference

Is it relevant to semantics whether the speaker has a certain individual 'in mind' by his use of the indefinite in a discourse like

(4) a. *A man* is walking in the park.

b. *He* is whistling.

and if so, how? According to Chastain (1975) and Donnellan (1978), among others, it is relevant both to the proposition expressed by the sentence in which the indefinite occurs, e.g. (4a), *and* to the propositions expressed by sentences with pronouns that take this indefinite as its syntactic antecedent, e.g. (4b). Kripke (1977), Lewis (1979), and also Stalnaker (1998) have argued that speaker's reference is relevant to semantics, but only through pronominalization. That is, it is truth-conditionally irrelevant for (4a), the proposition expressed by the sentence (or clause) in which the indefinite occurs, but is truth-conditionally relevant for (4b), the proposition expressed by a later sentence with a pronoun that takes the indefinite as its syntactic antecedent. According to proponents of standard dynamic semantics like Kamp (1981), Heim (1982), and Groenendijk & Stokhof (1991), however, the notion of 'speaker's reference' is not (dynamic) semantically relevant at all: the object the speaker has in mind is at most important for pragmatics, where pragmatics should be something built 'on top' of (dynamic) semantics. According to these theories, pronouns are always used descriptively, and, in cases like (4a)-(4b), go proxy for the *indefinite* description recoverable from their antecedents.[2]

Although the recent theories of discourse representation are quite successful in accounting for anaphoric dependencies across sentential boundaries, these theories face, I think, some problems, both empirically and conceptually. First, it is unclear what the *discourse referents* that are used in these treatments stand for; it seems that they are *only used* to account for cross sentential anaphora, but their existence is not explained independently. I will argue that such an independent motivation can be given when a unique speaker's reference is associated with the indefinite that introduced the discourse referent to the discourse. Second, in these recent theories no justice is done to the fact that pronouns are *definite* expressions, i.e., that in the case of singular pronouns always a notion of *uniqueness* seems to be involved (cf. Kadmon, 1990). This second point can be illustrated by some data that suggest that unbound anaphoric pronouns should in general have a more specific interpretation than the standard dynamic theories can offer.[3] I will discuss the first point only in a later section and concentrate in this section on the second one.

It is commonly assumed (e.g. Strawson, 1952) that the phenomenon of *pronominal contradiction* shows that anaphoric pronouns can at least

[2] To be a bit more precise, in the discourse *An S is P. He is Q. He is R*, the first occurrence of *He* goes proxy for the indefinite description *An S who is P*, while the second occurrence goes proxy for *An S who is P and Q*.

[3] See Dekker & van Rooy (1998) and van Rooy (1997, 2000) for a discussion of similar data involving anaphoric relations across belief attributions.

sometimes be used referentially. When John asserts (5a), Mary may react by saying (5b):

(5) a. John: *A man* jumped off the bridge,

b. Mary: *He* didn't jump, he was pushed.

In these cases the pronoun appears to be used referentially, referring to the *speaker's referent* of John's use of the indefinite.

The following example,[4] which illustrates what I will call the *specificity problem*, suggests that pronouns are more generally used referentially. If John says (6a), it would be odd for him to reply to Mary's question (6b) by uttering (6c) if two men called John up yesterday and he knows this.

(6) a. John: *A man* called me up yesterday.

b. Mary: Did he have a gravel voice?

c. John: That depends: if *he* called in the morning *he* did,
but if *he* called in the afternoon, *he* did not.

It not easy to see how this phenomenon can be explained if it is assumed that pronouns should simply be treated as variables bound by dynamic existential quantifiers. It also seems clear that the phenomenon cannot be explained by just assuming that, by Gricean Quality, speakers have to believe what they say, and account for this in terms of classical entailment; (6a) and (6c) are wrongly predicted to be acceptable given that John knows that two men called him up yesterday, one in the morning and one in the afternoon. Dekker (1997) argues that to explain why (6c) cannot be used appropriately in its most straightforward reading in such a context, a *more specific* relation than classical entailment between what is believed and what is said is needed to account for the intuition that John just wants to talk about one of the two men. A natural explanation can be given if it is assumed that for the use of the pronoun the speaker must have a specific object 'in mind'.

On the assumption that pronouns are normally used referentially, we can also explain the frequently observed distinction between the discourse (7a) and the single sentence (7b).

(7) a. There is a doctor in London. *He* is Welsh.

b. There is a doctor in London *who* is Welsh.

[4] This example came up in a discussion with Paul Dekker and Ede Zimmermann.

According to the standard account the two are predicted to be equivalent. There does, however, seem to be a distinction and it is this: for the use of the personal pronoun in the discourse (7a), the speaker must have a specific individual 'in mind' that the second sentence with the personal pronoun is about; whereas no individual need be 'in mind' to ensure the acceptability of the single sentence (7b), in which a *relative* pronoun is used.

I want to argue that we can account for a wide range of phenomena that dynamic semantics accounts for if we assume that a lot of anaphoric pronouns in the original fragment of DRT/FCS/DPL are used referentially, because they pick up the relevant speakers' referents of their antecedent indefinites, the object, or set of objects, the speaker had 'in mind'.[5] Such an approach immediately explains the definiteness of referentially used pronouns: a singular pronoun refers to the unique individual that is the speaker's referent of its antecedent indefinite.

Notice that this proposed analysis has much in common with the analyses of Evans (1977), Cooper (1979), Neale (1990), Heim (1990) and many others, according to which the denotation/reference of an anaphoric pronoun should in general be (uniquely) determined by a contextually given description. I have followed those who have rejected the most straightforward implementation of this suggestion, i.e. that this contextually given description is recoverable from the clause in which its antecedent indefinite occurs. According to my analysis of referentially used pronouns, the denotation/reference of the pronoun might also be said to be determined by a description, but only a description of the form *the individual, if there is one, that is the speaker's referent of its antecedent indefinite*. Thus, just like Kadmon (1990), I assume that the relevant uniqueness should be explained from the *speaker's* point of view.

Heim (1982) and Kamp (1988) argue, partly on the basis of the asymmetry in acceptability between (8a) and (8b), against the proposal to determine the denotation/reference of pronouns by contextually given descriptions:[6]

(8) a. Exactly *one* of the ten balls is not in the bag. *It* is under the sofa.

b. Exactly *nine* of the ten balls are in the back. *?*It* is under the sofa.

[5] I will assume that object d is the speaker's referent of John's use of an indefinite iff (i) John has a representation of an individual; (ii) this representation was caused by d; and (iii) this representation was responsible for John's use of the indefinite (see also Stalnaker, 1998).

[6] The example is attributed to Partee.

They argue that this asymmetry cannot be predicted on the basis of the truth conditions of the first sentence in each discourse and the surrounding circumstances alone, because what seems crucial is how each sentence is worded. It seems that the difference cannot be accounted for by means of the existence and absence of speaker's reference in the first and second discourses, respectively, either; even if the speaker had a specific ball in mind in the second discourse, the use of the personal pronoun would still be odd. Heim and Kamp observe that if pronouns are treated as variables bound by 'text-scope' existential quantifiers associated with explicitly mentioned indefinites, the asymmetry can be explained; and they argue that this latter approach is in fact the way to go. So, all that counts for the appropriate use of a pronoun is whether an indefinite has been explicitly used in the previous discourse.

In discussing examples (6a)-(6c), we have already suggested that the use of an indefinite in the preceding discourse is not a *sufficient* condition for the appropriate use of a pronoun.[7] But it also doesn't seem to be a *necessary* condition. The explicit use of an indefinite is also not a necessary condition for the appropriate use of an anaphoric pronoun because, as has been observed by many authors, the pronoun *it* can be used appropriately in (8b) when the speaker makes it clear that he is interested in the tenth ball (by looking for it for a moment), or that he has the tenth ball in mind.[8] What this suggests is that it is not so much the explicit use of an indefinite that counts, but rather that the speaker has made it clear to the hearer(s) that he has a specific individual 'in mind'.

But why, then, is it at least normally the case that the speaker can use a personal pronoun to 'refer' back to an explicitly used indefinite? Why is (8b) normally so much worse than (8a)? The reason, I wish to suggest, is that it is a *speech convention* among language users, and thus known to be a speech convention, that when a speaker uses an indefinite

[7] See also several examples discussed by Kadmon (1990). She argues that in a story, hearers do become reluctant to accept a definite, if it seems unlikely that the speaker has a specific referent in mind. The following example (her example (20)), for instance, is odd: *Once upon a time Leif bought a chair. In fact, he bought several identical chairs that time. He got the chair/it at a tag sale.*

[8] See, for instance, the following scenario sketched by Neale (1990, p. 209):

> Suppose I have ten pet mice, one of whom is called 'Hector'. Hector is always getting out of the cage in which I keep all ten mice, and whenever he does so he goes and hides under the sofa. I open up the cage and begin counting mice: "One, two, three,..." When I reach 'nine' I turn to you and with a knowing look I say, *I put all ten mice in the cage an hour ago, and there are only nine here now.* Knowing Hector's habits, you might then reply by saying *I bet he's under the sofa again.*

explicitly, he *normally* has a specific individual in mind. It follows that the hearer will assume that the speaker has a specific individual in mind when he used an indefinite. As a result, thinking of pronouns as referential expressions can explain the above asymmetry. The reason is that in the first sentence of (8a) but not of (8b), an indefinite is explicitly used, and thus only in the former case can a (referential) pronoun normally be used appropriately.

This proposal that with the use of an indefinite a specific individual is associated that can be picked up later by an anaphoric pronoun sounds very much like what has been proposed by Chastain (1975), Donnellan (1978), and Fodor & Sag (1982). Although I wish to claim that specifically-used indefinites come with a speaker's referent, I want to follow Kripke (1977), Lewis (1979), and Stalnaker (1998), however, in taking the speaker's referent of the indefinite to be semantically irrelevant to the interpretation of the clause in which the indefinite itself occurs. That is, for the truth of a sentence of the form *An S is P*, only the *existential* information counts. Just as Kripke (1977) argued, speaker's reference is relevant to semantics, but only through *pronominalization*. But now, of course, the following question arises: How could we account for this *referential* analysis of pronouns, on the one hand, and the *existential* interpretation of indefinites, on the other? This question will be addressed in the next section.

2.2. Referential pronouns in dynamic semantics

Let me now sketch how things can be done formally, by defining a dynamic semantics for a syntax that is meant to represent a natural language. The syntax is very much like first-order predicate logic, with the only difference that not only variables will be treated as terms, but also the added *discourse markers* and *theta terms*. I will say that $\theta r P$ is a theta term, if P is a one-place predicate, and r a discourse marker.[9] Indefinites will be represented by theta terms, and unbound pronouns by discourse markers. Variables and discourse markers will, as usual, be interpreted with respect to partial *assignment functions*, but because discourse referents will be interpreted as individual concepts, we also need a world to evaluate them. To account for the intuition that speakers have specific individuals in mind for their use of indefinites that can be picked up by anaphoric pronouns, we have to guarantee that a *unique* individual (if there is one) is assigned to the theta term by the possibility with respect to which it is interpreted. For this reason

[9] Notice that r is a discourse marker and θ is *not* a variable binder that binds occurrences of r in P; P doesn't contain r. In the beginning I will neglect complex one-place predicates, but they will be introduced later.

we will assume that a possibility also contains a *reference function* that assigns to a theta term in each world the specific individual that intuitively is the unique speaker's reference of the indefinite (if there is one) represented by the term.[10] Limiting ourselves to singularly used indefinites and pronouns, we can say that each reference function ϕ is a function from properties, i.e. functions from worlds to subsets of our domain D of all possible individuals, to individual concepts, i.e. functions from W to $D \cup \{*\}$, where $*$ is the impossible object, no element of D. Thus, terms will be interpreted as follows:

- $[[t]]^{w,\phi,g}$ = $g(t)$, if t is a variable,
  = $g(t)(w)$, if t is a discourse referent,
  = $\phi(I(P))(w)$, if $t = \theta rP$ and $\phi(I(P))(w) \in I_w(P)$,
  = $*$ otherwise.

Notice that the indefinite θrP denotes $*$ in w in case $\phi(I(P))(w)$ is an object that is no P in w.

Although, as we will see, our analysis will be *truth-conditional*, it will be very much *dynamic* in the sense that we allow some expressions to introduce discourse referents to the discourse. To account for the *referential* treatment of pronouns, and the *existential* treatment of indefinites, we will define the notion of *truth of a sentence*. But to define this notion we first have to determine how theta terms can introduce discourse markers to the discourse and when a sentence is *rigidly true*. We will determine both separately,[11] beginning with the definition of $Upd(E, \langle w, \phi, g\rangle)$, which tells us how the partial assignment function g is updated after the interpretation of expression E. Just to sketch the idea, I will assume for the moment that the language contains only one-place predicate constants.

- $Upd(P(t), \langle w, \phi, g\rangle)$ = $Upd(t, \langle w, \phi, g\rangle)$

[10] There exists also another approach to 'secure' uniqueness, adopted by Kadmon (1990); just allow additional predicates to 'enter' into the representation of the sentence. On this approach, however, there seems to be no general recipe available, or it would be to always add a predicate like 'individual that the speaker associates with his use of the indefinite'. But there are two problems with this alternative approach. First, by just *adding* predicates to the representation, pronominal contradiction still cannot be accounted for, and second, it remains unclear how such a purely *descriptive* approach would account for *occurrences* of indefinites. An additional problem is that to assure that a sentence of the form *An S is P* is true iff the intersection of the denotations of S and P is non-empty, Kadmon has to assume that the extra predicates are only added to the representation once a pronoun is used, although this extra information seems to be associated with the antecedent indefinite.

[11] But in the end the two notions have to be defined *simultaneously*.

- $Upd(t, \langle w, \phi, g\rangle)$ = g, if t is a variable or discourse referent,
  = $g[^{r}/_{\phi(I(P))}]$, if t is of form $\theta r P$[12]
- $Upd(A \wedge B, \langle w, \phi, g\rangle)$ = $Upd(B, \langle w, \phi, Upd(A, \langle w, \phi, g\rangle)\rangle)$

Where $g[^{r}/_{\phi(I(P))}]$ is g extended with an assignment of the speaker's referent of the indefinite to the discourse referent r.

Now we can define the notion of *rigid truth*, where we determine the truth of the sentence with respect to a possibility in which no reference functions are existentially quantified over. I will limit myself here to the two most important clauses.

- $[[P(t)]]^{w,\phi,g} = 1$ iff $[[t]]^{w,\phi,g} \in I_w(P)$
- $[[A \wedge B]]^{w,\phi,g} = 1$ iff $[[A]]^{w,\phi,g} = 1$ and $[[B]]^{w,\phi,h} = 1$, where $h = Upd(A, \langle w, \phi, g\rangle)$

Finally, we define the notion of *truth of sentence A with respect to* $\langle w, \phi, g\rangle$, $\langle w, \phi, g\rangle \models A$, by existentially quantifying over the set, Φ, of reference functions:

- $\langle w, \phi, g\rangle \models A$ iff $\exists \psi \in \Phi$ such that $[[A]]^{w,\psi,g} = 1$.

Notice that when A is interpreted after the sequence of sentences S_1 to S_n, A will be true with respect to w and ϕ iff $\exists \psi \in \Phi$ such that $[[A]]^{w,\psi,h} = 1$, where $h = Upd(S_1 \wedge ... \wedge S_n, \langle w, \phi, g\rangle)$ for any assignment function g. As a result, the above definition assures that an indefinite is always interpreted existentially, and a referentially-used anaphoric pronoun always refers back to the speaker's referent of its antecedent indefinite.

It is easy to see that both sentences of the discourse *A man jumped off the bridge. He died*, as represented by $JoB(\theta r(Man))$. $Died(r)$, are now predicted to be true with respect to w and ϕ iff there exists a man who jumped off the bridge: $I_w(Man) \cap I_w(JoB) \neq \emptyset$, and the speaker's referent of the indefinite died: $\phi(I(Man))(w) \in I_w(Died)$. Notice that because the speaker's referent of the indefinite of the first sentence need not have actually *jumped* off the bridge, a second speaker might truly react by saying that *he* didn't jump, but was pushed, which shows that

[12] For simplicity I have assumed, and will continue to assume, that the reference function remains the same. But we might also say that the reference function should change after the interpretation of theta term $\theta r P$. One way one might go (cf. Van Rooy, 1997) is to say that the new reference function, ψ, will be exactly like ϕ, except that $\psi(I(P)) = \phi(I(P) - \{\phi(I(P))\})$ to force a different choice from $I(P)$, where '$-$' has the obvious interpretation.

we can also account for pronominal contradiction examples. Notice also that we predict that the second sentence of discourse (7a), *There is a doctor in London. He is Welsh*, represented by $E(\theta rDL).\ Wr$ (where 'E' is the existence predicate), might be false even if the sentence (7b), *There is a doctor in London who is Welsh*, represented by $E(\theta rDL) \wedge Wr$,[13] is true, while this is not predicted to be the case in standard dynamic semantics. We don't predict the two to be equivalent when the speaker's referent of the indefinite *a doctor* is not Welsh, although another doctor in London is.

The approach sketched here can also account for the *specificity problem* discussed above, because it is assumed that, on its most straightforward reading, John must have a specific object 'in mind' for his use of the indefinite and pronoun. When John asserts (6a), *A man called me up yesterday*, we predict that John makes the specific individual that he had in mind for his use of the indefinite *a man* available for reference for pronouns and (other) short descriptions. It would normally be odd for him to answer the question (6b), *Did he have a gravel voice?* by saying (6c), *That depends: if he called me up in the morning he did, and if he called me up in the afternoon, he did not*, because it can be assumed that the individual that the speaker had in mind as antecedent was either the one who called him up in the morning or the one who called him up in the afternoon.

2.3. Speaker's reference and semantics/pragmatics

We have accounted for the intuition that indefinites have speakers' referents in terms of reference functions; even if there are more objects with property S in the world of $\langle w, \phi, g\rangle$, it still might be that in this possibility an indefinite of the form *An S* chooses only a specific individual, $\phi(I(S))(w)$. Actually, a closer look at what I call a reference function reveals that this individual need not even be an S. The reason is that we have not yet put any constraint on the individual concepts that reference functions assign to properties.[14] Although to account for pronominal contradiction examples like

(9) a. A: *A man* is running through the park.

b. B: *He* is not a man, he's just a boy.

[13] Later we will also introduce complex predicates of the form $\hat{x}A$, if A is a sentence. Then we can also represent (7b) by $E(\theta r\hat{x}[DLx \wedge Wx])$.

[14] Remember that this is no problem for the definition of rigid truth, because for the *interpretation* of theta terms the selected individual should have the corresponding property.

we should not demand that the chosen individual that figures as the speaker's referent of the indefinite *A man* be a man, it does seem reasonable to assume that the *speaker* should *believe* that it is a man. That is, the speaker must have a representation of this individual, and believe of him that he is a man.

When we make possibilities finer-grained than is usually assumed in dynamic semantics, by assuming that possibilities also need something like reference functions, such an intuition can be implemented. Let's see how.

First, I will make the assumption that belief states are structured around 'belief objects'. This means that a belief state of an individual in world w is not simply modeled by a set of possible worlds, but rather by a pair, $\langle K, S\rangle$ consisting of (i) a set of worlds K, and (ii) a set, S, of partial functions from worlds to individuals, i.e. individual concepts, such that for each $s \in S$, all worlds in K are in the domain of s. The individual concepts in these belief states are supposed to model the 'belief objects', and the belief objects are representations of what the agent takes to be real individuals.[15]

Second, we have to relate the reference function of the possibility to the belief state of the speaker by putting a constraint on models. Considering only cases where one speaker introduces the relevant discourse markers to the discourse, we can continue assuming that a possibility contains only one reference function. Let us now suppose that in world w the belief state of our speaker can be represented by the pair $\langle K, S\rangle$. To implement our above intuitions we can now put a constraint on the models; we can say that $\langle w, \phi, g\rangle$ is an appropriate possibility only if for each property $P \in dom(\phi)$ there is at most one 'belief object' s in S of the speaker's belief state $\langle K, S\rangle$ in w such that for each $v \in K$ it is the case that $\phi(P)(v) = s(v) \in I(P)$. Although associated with each world and our speaker there might be several possible reference functions that obey this constraint, for each finer-grained possibility that also contains a reference function, the speaker will associate with each indefinite at most one specific belief object that satisfies its descriptive content in the worlds compatible with what the speaker believes. Thus, just like Kaplan (1989) requires that a quadruple $\langle a, p, t, w\rangle$ can function as a proper context only in case the agent, a, is a person who is at the place, p, at the time, t, in the world, w, of the possibility, we demand that

[15] Thus, if w is the actual world, such a concept s of the belief state $\langle K, S\rangle$ might, but need not, be defined for w. If so, this 'belief object' s is a representation of a real individual, $s(w)$, if not, not.

a reference function can be 'part' of a context only if it fits with the beliefs and intentions of the speaker in the context.[16]

Now we go back to our example (6a) - (6c), where we might assume that John associates two belief objects with the information that they are men who called him up yesterday. However, and crucially, only one of those two belief objects will be associated with (the relevant occurrence of) the indefinite *A man* in each possibility that also contains a reference function that obeys the above discussed constraint. Thus, this extra information contained in a possibility models, or is supposed to model, the intuition that John has, and knows he has, a specific belief object 'in mind' with his use of the indefinite.[17] Now we have to make sure that he would also associate the same 'belief object' with the pronoun that he uses to refer back to this indefinite. This can be guaranteed when we assume, by Gricean Quality, that except for our constraint for the use of specific indefinites, speakers also have to believe what they say. Combining both constraints, we demand that if John utters A in possibility $\langle w, \phi, g\rangle$, where ϕ is the reference function that belongs to his own belief state, $\langle K(j,w), S(j,w)\rangle$, and g the assignment function, he can do this appropriately only if $\forall v \in K(j,w): \ [[A]]^{v,\phi,g} = 1$. This explains why it would be odd for him to respond to Mary's question (6b) by saying (6c); it would mean that he associated a specific 'belief object' with his use of the indefinite, but did not associate with this belief object the information regarding *when* he called, i.e., that he called in the morning, or that he called in

[16] It is important to note that we only have to put such a constraint on the models because we use a *type*-analysis. If we would have used a *token*-analysis, and assumed that the speaker's reference of an indefinite depends on the relevant *token* of the indefinite, the speaker's reference of an indefinite would not depend on an additional reference function, but rather on *facts* about the *world*. Notice that when speaker's reference depends on tokens of indefinites, possibilities will not be finer-grained than worlds. Assuming that a pronoun that does not go proxy for the definite description recoverable from its antecedent indefinite should refer to the speaker's referent of the *token* of its antecedent, Stalnaker (1998) argued that facts about anaphora do not rule out a purely possible world analysis of contexts. We have not made use of a token-analysis in this paper, for one thing because this would make the analysis in some respects rather complicated (or so I fear). In our type-analysis we have made a formal distinction between facts about the subject matter of conversation, and facts about the conversational situation itself. This is conceptually somewhat unfortunate, because intuitively also the latter facts are 'worldly' facts, although they are not treated that way.

[17] If the individual concept that models this belief object that the speaker has 'in mind' also has a value in the world where the discourse is taking place, this value will be the *speaker's referent* that I talked about above.

the afternoon.[18] Notice that the combination of both constraints also explains why for the appropriate use of a sequence of the form *A man is P. He is Q* by a single speaker, she should believe that there is a man who has both property P and Q. In particular, it explains why for pronominal contradiction sentences like (5a) - (5b) at least *two* speakers should be involved.

Of course, some proponents of standard dynamic semantics might be sceptical of my use of the notion of speaker's reference within a *semantic* theory. They might think that the notion of speaker's reference is irrelevant to semantics, even for determining the *truth conditions* of later sentences in which anaphoric pronouns occur. On this view, the notion might at best be relevant to pragmatics. Already at this moment two responses can be given to such sceptics:[19] (i) The phenomenon of *pronominal contradiction* suggests that the notion of speaker's reference is relevant for determining the truth conditions of a later sentence in which an anaphoric pronoun occurs, and (ii) even if the notion of speaker's reference were relevant to pragmatics only, the resulting pragmatic analysis would be, I claim, very close to my *semantic* analysis.[20]

What should an integrated semantic/pragmatic theory of anaphoric relations look like? That depends, of course, on how we think of the semantic/pragmatic interface. The perhaps most popular view on pragmatics is to think of it as something that should be built 'on top' of truth-conditional semantics. According to another view, defended, for instance, by Van Fraassen (1967), and assumed also by most proponents of the satisfaction theory of presuppositions,[21] we should rather think of semantics as an *abstraction* from pragmatics. If we adopt the latter view on the semantic/pragmatic interface for the analysis of anaphora, we can say that the analysis I have given above is just a pragmatic account, and think of a semantic analysis of anaphoric relations as an abstraction of this pragmatic analysis.

In the theory that I have just formulated, possibilities contain more information than the possibilities used in standard dynamic theories; they also contain the information of/about what the speaker's referents are of particular indefinites. This extra information is responsible for the main differences between my proposal and the standard accounts

[18] Of course, such a reading does exist too (and is predicted to exist by my analysis), but it is not the reading that I wanted to account for.

[19] An additional response concerning the status of discourse referents will be given in section 3.3.2.

[20] And note that, at least as far as I know, also in pragmatics no such a detailed analysis has ever been given (but see Dekker (1997) for a related proposal.)

[21] See Stalnaker (1970).

that I pointed out above. But, of course, it is possible to abstract away from this extra information; and if we do so, what results is (truth-conditionally equivalent to) the standard dynamic theory. That is, when we don't want to say that pronouns should refer back to speaker's referents, we can define the truth of sentence A after sequence S as follows: A is true* after sequence S in $\langle w, \phi, g\rangle$, $\langle w, \phi, g\rangle \models^*_S A$, iff there is a $\psi \in \Phi$ such that $[[S_1 \wedge ... \wedge S_n \wedge A]]^{w,\psi,g} = 1$. As a result, the so-called (conjunctive) donkey-equivalence between *There is a doctor in London. He is Welsh* and *There is a doctor in London who is Welsh* holds again, just as in standard dynamic systems. Thus, I am not claiming that proponents of standard dynamic systems are saying anything wrong, but only that they aren't saying enough: they should take the notion of speaker's reference more seriously than they actually do. Or better, perhaps, they should not think of pragmatics as something built on top of semantics, but should rather think of semantics as an abstraction from (this relevant part of) pragmatics.

2.4. Shifting reference functions

The above treatment of pronouns as referential expressions is very much in line with what Kripke (1977) had proposed. But then it is only to be expected that my analysis, just like Kripke's, has problems dealing with a *donkey sentence* like (10):

(10) If *a donkey* walks in the park, *it* brays.

The problem is that in the above sentence the pronoun doesn't refer to a particular donkey, while the pronoun can arguably also not be treated as an abbreviation for the *definite* description *the unique donkey that walks in the park*, because it seems that the sentence can also be true in case more than one donkey is walking in the park.

One way to solve the problem is to assume that indicative conditionals are analyzed in terms of conjunction and negation (like in DPL), and that a logical operator like *negation* is treated as an *intensional* operator, in that it allows the contextual reference function to shift:[22,23]

[22] Note that Kaplan (1989) would call such an 'intensional' treatment of negation a *monster*. Still, it is well known that for some context dependent expressions, contexts can be shifted. Consider for instance the following example (Rossdeutscher, pc.): *The salt is to the right of the pepper. If I was sitting on the other side of the table, the salt would be to the left of the pepper.* More relevant, perhaps, is Partee's (1989) *Whenever we want to stop for the night, the nearest motel is five miles down the road.*

[23] In the previous section I have argued that the belief state of the speaker puts a constraint on the reference function of the possibility. To account for donkey

• $[[\neg A]]^{w,\phi,g} = 1 \quad$ iff $\quad \neg\exists\psi \in \Phi :\ [[A]]^{w,\psi,g} = 1$

It is easy to see that we can now account for the universal reading of (10), when we interpret the conditional, like in DPL, in terms of negation and conjunction as follows: $\neg(WiP(\theta rD) \wedge \neg Bray(r))$.[24] If we would say that $Upd(\neg A, \langle w, \phi, g\rangle) = g$, we would moreover guarantee that negation figures as a plug with respect to anaphoric binding, just like in original DRT/FCS. We also might represent the donkey sentences like in FCS as $ALWAYS(WiP(\theta rD), Bray(r))$. Such formulae can be interpreted as follows (where ADV stands for any kind of adverb of quantification, and $[ADV]$ stand for its usual interpretation):

• $[[ADV(A,\ B)]]^{w,\phi,g} = 1 \quad$ iff
$[ADV](\{Upd(A, \langle w, \psi, g\rangle) :\ \psi \in \Phi\ \&\ [[A]]^{w,\psi,g} = 1\},$
$\{Upd(A, \langle w, \psi, g\rangle) :\ \psi \in \Phi\ \&\ [[A \wedge B]]^{w,\psi,g} = 1\}),$[25]

giving rise to the *unselective* interpretation. The *selective*, or *asymmetric* reading can, as usual, be handled by quantifying over *equivalence classes* of cases (cf. Chierchia, 1992, and Dekker, 1993).[26] Quantified sentences can be analyzed similarly, although they are normally interpreted selectively, quantifying only over the values of variables:

• $[[Det_x(A, B)]]^{w,\phi,g} = 1 \quad$ iff
$[Det](\{d \in D :\ \exists\psi \in \Phi\ \&\ [[A]]^{w,\psi,g[^x/_d]} = 1\},$
$\{d \in D : \exists\psi \in \Phi\ \&\ [[A \wedge B]]^{w,\psi,g[^x/_d]} = 1\})$

Although I believe that most conditional donkey sentences should be treated as I suggested above, I don't think that the analysis of asymmetric readings should be accounted for in this way; for it would lead to the unwanted prediction that some pronouns should be treated

sentences in terms of shifting reference functions, we should not demand that also the shifted reference functions have to obey this constraint.

[24] But we have to guarantee, of course, that there are enough reference functions in the model, one for each donkey.

[25] In fact, by assuming that indefinites do not introduce individuals, but individual *concepts*, we should not simply quantify over assignments, but rather over equivalence classes of assignments. I will leave the (here not so crucial) tedious formal details to the reader, however.

[26] Selective and unselective interpretations differ from each other only if we allow for more-place relations. We haven't defined yet, however, how in these cases discourse referents are introduced. This can be done as follows:

• $Upd(R(t_1, .., t_n), \langle w, \phi, g\rangle) = Upd(t_n, .., Upd(t_1, \langle w, \phi, g\rangle)..)$

• $[[R(t_1, ..., t_n)]]^{w,\phi,g} = 1 \quad$ iff $\quad \langle [[t_1]]^{w,\phi,g}, ..., [[t_n]]^{w,\phi,g}\rangle \in I_w(R)$

as abbreviations for (dependent) *indefinite* descriptions after all. From now on I will assume that a sentence of the form $Det_x(A, B)$ should be interpreted in the following traditional way,

- $[[Det_x(A,B)]]^{w,\phi,g} = 1$ iff $[Det](\{d \in D:\ [[A]]^{w,\phi,g[^x/_d]} = 1\}, \{d \in D:\ [[A \wedge B]]^{w,\phi,g[^x/_d]} = 1\})$,

with no shifting of reference-contexts involved. It should be obvious that this treatment of quantification gives rise to the question of how to account for quantificational donkey sentences like *Every farmer who owns a donkey beats it.* I will come back to this issue below.

According to Fodor & Sag (1981), indefinites distinguish themselves from ordinary quantificational noun-phrases by being able to 'outscope' so-called syntactic islands. Indefinites in the antecedent of a conditional, for instance, might 'outscope' the whole conditional, while this is claimed to be impossible for ordinary quantificational NPs. This can be illustrated by example (11), where the pronoun *He* in the following sentence can take the indefinite *a friend of mine* as its antecedent.

(11) If *a friend of mine* leaves, I will be unhappy. *He* is a nice guy.

Fodor & Sag propose that this should be accounted for by assuming that indefinites can also be used referentially, in which case they automatically receive highest 'scope'.

One unfortunate consequence of our analysis of donkey sentences in terms of shifted reference-contexts is that in this way the natural connection between Fodor & Sag's (1981) analysis and mine is lost; indefinites occurring in, for instance, the antecedent of a conditional, would always be interpreted in terms of *shifted* reference functions, and it is not clear anymore how to account for the anaphoric dependency in (11).

If we don't want to allow for the possibility that indefinites can outscope conditionals,[27] it is still possible to solve this problem; we simply have to find a mechanism that enables us to somehow *remember* the original reference function. One way to do this is by making use of *indexed actuality operators*.[28] We can assume (i) that operators like *negation* and quantifiers will be *indexed* by variables that refer to the reference function[29] of the possibility with respect to which this

[27] In the next section we will also introduce *complex predicates* to the language. With this extra machinery we can assure that not only quantifiers, but also *terms* can stand in non-trivial *scope* relations with other expressions.

[28] These operators were already used by Forbes (1985) to account for the various ways a *definite* description can be interpreted *in situ* when it is embedded under modal operators.

[29] Or perhaps to the world/reference function pair.

operator itself is interpreted, and (ii) that (theta) terms can be fronted by an *indexed actuality operator*, $@_i$, that says that the theta term has to be evaluated with respect to the reference function that is the value of the index:

- $[[\neg_i A]]^{w,\phi,g} = 1 \quad \text{iff} \quad \neg\exists\psi \in \Phi : \ [[A]]^{w,\psi,g[^i/_\phi]} = 1$
- $[[@_i t]]^{w,\phi,g} \quad = \quad [[t]]^{w,g(i),g}$

If we now would represent my utterance of (11) by $\neg_1(Leave(@_1\theta r(F)) \wedge \neg Unhappy(I))$, the rigid truth or falsity of the sentence will depend on the specific friend of mine that I have in mind.[30] But now we also have to take care of the anaphoric pronoun *He* of (11). To do this we have to give up the assumption that negations are absolute plugs with respect to anaphoric binding. That is, we have to admit that pronouns can pick up discourse referents introduced under the scope of a negation. However, we should allow for this only if such a discourse referent has a *unique* denotation, i.e., when the discourse referent is introduced by an indefinite interpreted by the original reference function.[31] Otherwise it would even for the speaker be unclear to which object(s) he would refer by his use of the anaphoric pronoun. Both conditions can be met when we (re)define $Upd(\neg_i A, \langle w, \phi, g\rangle)$ as follows (Where $k \setminus g$ is the subtraction of g from k):

- $Upd(\neg_i A, \langle w, \phi, g\rangle) \ = \ g \cup \{\langle r, o\rangle \in Upd(A, \langle w, \psi, g[^i/_\phi]\rangle) \setminus g : \ \psi \in \Phi \ \& \ \forall\chi : \ Upd(A, \langle w, \chi, g[^i/_\phi]\rangle)(r) = o\}$

Notice that if we assume that names and definite descriptions also introduce discourse referents to the discourse, we could account for the fact that when such expressions are used under the scope of (operators defined in terms of) a negation, we can always refer back to them in the ongoing conversation.[32] In fact, we don't even require the actuality operator to account for this; the (re)definition of $Upd(\neg_i A, \langle w, \phi, g\rangle)$ is enough.

[30] Observe that although it matters truth-conditionally whether we determine the truth or falsity of the first sentence of (11) with respect to a *contextually given* reference function, or that we *existentially* quantifying over reference functions, both approaches can account for the unusual 'scope' of the indefinite in an equal way.

[31] Or it must be presupposed that the indefinite can only have a unique interpretation.

[32] We would not need to follow Kamp & Reyle (1993), for instance, and assume a special proper name rule to account for this. We also wouldn't need to account for it in terms of scope.

3. Descriptive and functional pronouns

It's a fact about anaphora that indefinites occurring under the scope of two negations normally cannot be picked up by a later pronoun. It is also very unusual that a pronoun occurring in one disjunct can pick up an indefinite of the other disjunct. In standard dynamic semantics this is accounted for by means of *accessibility* constraints. Unfortunately, there are well known counterexamples to these constraints on anaphoric binding; pronouns can sometimes take an indefinite as antecedent, although the anaphoric accessibility constraints predicted by DRT/FCS are violated:

(12) Either John does not own *a donkey*, or he keeps *it* very quiet. (Evans, 1977)

(13) It is not true that John didn't bring *an umbrella*. *It* was purple and *it* stood in the hallway. (Muskens & Krahmer, 1995)

It is well known that standard dynamic semantics has problems with such sentences. The reason is that negation is treated here as a plug with respect to anaphoric binding. Note that contrary to the standard dynamic approach, negations don't have this property according to the E-type account of Evans (1977) and his followers. Proponents of the standard account argue that negations *should* be treated as plugs to account for the unacceptability of (14).

(14) There is *no guest* at this wedding. *He* is standing right behind you.

The unacceptability of (14) can be accounted for by *structural* means: an object 'introduced' under the scope of a negation cannot be picked up by anaphoric means in further discourse. But the E-type approach has, of course, no problem with the unacceptability of (14). The sequence (14) is out, not for structural but for *semantic* reasons. If the pronoun *he* of the second sentence would stand for *the guest at this wedding*, the second sentence would be trivially false, in case the first sentence is true. That is quite a natural reasoning, I would say. And does the acceptability of the sentences (12) and (13) not justify this reasoning?[33]

[33] Although the case for the existence of E-type pronouns is usually found most convincing with respect to plural pronouns, (quantificational) subordination, and with pronouns that take *definite* descriptions as antecedent, I will limit myself in this section to singular pronouns going back to indefinites that are not subordinate to quantifiers.

Not so, say Krahmer & Muskens (1995). Negation is a syntactic plug with respect to anaphoric binding, and the reason why (12) - (13) are acceptable is that a double negation is a plug unplugged. A clause of the form '$\neg\neg A$' is not only truth-conditionally, but also *dynamically* equivalent with 'A'. They can account for this claimed equivalence in a way that is not completely ad hoc by using techniques from partial logic.[34]

There are some worries with their approach, however. First, intuitively there seems to be no difference between (12) and a sentence like (15):

(15) a. It is possible that John does not own *a donkey*,

b. but it is also possible that he keeps *it* very quiet.

It would be nice if both could be handled by the same mechanism. But it is rather doubtful that this mechanism will be that '$\neg\neg A$' is equivalent to 'A'. Second, if an indefinite occurs under the scope of two negations, it seems that a singular pronoun can take it as its syntactic antecedent only if there is exactly one object that could be the referent of the indefinite. For (12) and (13), for instance, the uses of the pronoun *it* in the second disjuncts can pick up the *unique* donkey that John owns, and the *unique* umbrella John brought, respectively, only. If it is presupposed that possibly John owns more donkeys, and if there are maybe more umbrella's John brought, the uses of *it* in the second disjunct of (12) and the second sentence of (13) would be, I think, inappropriate.

It seems that Krahmer & Muskens agree. Discussing the contrast in acceptability between (16) and (17),

(16) It is not true that there is no guest at this wedding.
?*He* is standing right behind you.

(17) It is not true that there is no bride at this wedding.
She is standing right behind you.

they say that the distinction is due to a uniqueness effect.

> Given some highly unlikely context in which it is understood between speaker and hearer that at most one guest can be present at this particular wedding (16) would be fine. We feel that it is precisely the unlikelihood of such a context which explains the markedness of (16). (Krahmer & Muskens, 1995, p. 359)

[34] Groenendijk & Stokhof (1990) and Dekker (1993) reach a similar result by using lifting, instead of partiality.

I completely agree. But then they make the following claim about these problematic cases:

> Since such apparent counterexamples on closer examination turn out to be no counterexamples at all, it seems we can take it as a general rule that as far as truth conditions and the possibility of anaphora are concerned double negations in standard English behave as if no negation at all were present. (Krahmer & Muskens, 1995, p. 359)

I'm afraid that I don't understand this. That you can explain why a counterexample to your approach *is* a counterexample, doesn't mean that on closer examination it 'turns out to be no counterexample at all'.

I propose to take the counterexample seriously. This can be done by assuming in the cases discussed above that the pronouns go proxy for the definite descriptions recoverable from their antecedents, i.e. as *descriptive* pronouns. This suggests that the division of labor between referential and descriptive pronouns should be taken seriously in the following way: Where referential pronouns take specifically used indefinites as antecedents, descriptive pronouns can take indefinites as antecedents for which the speaker had no specific individual in mind, and which were used unspecifically. Moreover, a *singular* descriptive pronoun can be appropriately used only if it is presupposed that the associated description has a *unique* instantiation. Kripke's (1977) following example shows, however, that descriptively used pronouns can also pick up at least definites as antecedents that are used *specifically*. If John says *Her husband is kind to her*, and he has a specific individual in mind, Mary might still react by saying *No he isn't. The man you are referring to isn't her husband*, where the pronoun is used descriptively. This suggests that descriptive pronouns can also refer back to indefinites that are used specifically. Notice that by adopting the existence of E-type pronouns, we, contrary to proponents of standard dynamic semantics who also want to make use of E-type pronouns, don't have to give up an (in one sense) uniform analysis of pronouns. Only on our account it will be the case that there is *always* a notion of uniqueness/exhaustivity involved with pronouns.

To account for descriptive pronouns in our dynamic framework, I assume that indefinites can introduce not only specific individuals (or concepts) into the discourse, but also properties. The idea is that the theta term θrP used in formula $Q(\theta rP)$ and interpreted in possibility $\langle w, \phi, g\rangle$ will not only introduce the speaker's referent (if there is one), $\phi(I(P))(w)$, to the discourse under discourse referent r, but also the

property $I(P) \cap I(Q)$ under discourse referent r'.[35] To account for this in a compositional way, however, we have to define functions that determine (i) which discourse referent (if any) term t introduces to the discourse, and (ii) what the property is that corresponds to the theta term. These functions will be denoted by dr and pty, and $dr(t)$ and $pty(t)$ will have values r and P, respectively, when t is of the form $\theta r P$, and will denote dummies otherwise. Notice that when $dr(t)$ has value r, $dr(t)'$ will stand for discourse referent r' and denotes a function from worlds to sets of individuals. From now on I will assume that one-place predicates need not be simple, but can be *complex*, too. That is, if A is a sentence, and x a variable, I will say that $\hat{x}A$ is a complex one-place predicate. If we now interpret complex predicates and discourse referents that denote properties as follows, and redefine $Upd(P(t), \langle w, \phi, g\rangle)$ as below,

- $I_{w,\phi,g}(\hat{x}A) = \{d \in D : [[A]]^{w,\phi,g[^x/_d]} = 1\}$[36]
- $Upd(P(t), \langle w, \phi, g\rangle) = Upd(t, \langle w, \phi, g\rangle)[^{dr(t)'}/_f]$, where $\forall w \in W : f(w) = I_{w,\phi,g}(P) \cap I_{w,\phi,g}(pty(t))$.
- $[[r']]^{w,\phi,g} = d$, if $g(r')(w) = \{d\}$,
  $= *$ otherwise

[35] It follows that referential and descriptive singular pronouns should be represented by discourse referents of a different type. But this is obviously not crucial; we can simply assume that the 'objects' introduced that model speakers' referents should also be modelled by properties, but then properties that have in each world at most one instantiation (Van Rooy, 1997). Such properties correspond one-to-one to individual concepts. Notice that in that case unbound singular pronouns would formally *not* be ambiguous, although they intuitively can be used in two different ways. So, on our analysis *indefinites* are not ambiguous because, although not all of them come with an actual speaker's reference, they are all represented by theta terms, *pronouns* are not ambiguous because there is always a notion of uniqueness/exhaustivity involved.

[36] Once we assume that complex one-place predicates exist, we have to change some of the definitions we used above. First, we have to allow for the case that predicates introduce discourse referents. Thus, we have to say that $Upd(P(t), \langle w, \phi, g\rangle) = Upd(P, \langle w, \phi, Upd(t, \langle w, \phi, g\rangle)\rangle)$, where $Upd(P, \langle w, \phi, g\rangle) = g$, if P is a primitive one-place predicate, and $Upd(A, \langle w, \phi, g\rangle)$, if P is the complex predicate of the form $\hat{x}A$. Next we have to account for the fact that the predicate might also contain anaphora. We have to say that $P(t)$ is true in $\langle w, \phi, g\rangle$ iff the referent of t is an element of $I_{w,\phi,Upd(t,\langle w,\phi,g\rangle)}(P)$, which is simply $I_w(P)$ in case P is primitive. Notice that also theta terms can involve complex predicates. This means that instead of applying the reference function simply to $I(P)$, we have to apply it to $I_{\phi,g}(P)$, because the complex predicates might contain other indefinites and anaphora. Finally, we can no longer define the interpretation of terms and notions of rigid truth and $Upd(E, \langle w, \phi, g\rangle)$ completely independent of each other; from now on the notions have to be defined *simultaneously*.

we can easily account for pronouns 'referring' back to quantificationally used indefinites like (12) and (13), given our dynamic interpretation rule of (operators defined in terms of a) negation of the previous section.

We noticed in the previous section that due to our earlier (non-shifting) analysis of quantificational noun phrases, we could no longer account for the exhaustive interpretation of pronouns in quantificational donkey sentences like

(18) Every bishop who meets *another bishop*, greets *him*.

However, once we assume that indefinites can also introduce discourse referents to the discourse that can be picked up by descriptive pronouns, it turns out to be easy to account for the exhaustive interpretation of the pronoun in (18). We have assumed until now that an indefinite introduces a discourse referent that denotes (for each world) a *set* of individuals. But we might obviously loosen this assumption, and assume that some indefinites introduce discourse referents that denote for each world a *function* from individuals to sets of individuals. The easiest way to go about is to say that a formula like $Q(x, \theta r\hat{y}P(x,y))$, for instance, introduces into possibility $\langle w, \phi, g\rangle$ the function $f \in [(W \times D) \rightarrow \wp(D)]]$ such that for any $\langle w, d\rangle$ in the domain of f:

$$f(\langle w, d\rangle) \quad = \quad I_{w,\phi,g[^x/_d]}(\hat{y}[P(x,y) \wedge Q(x,y)])$$

Now we can introduce a distribution operator δs, that can front a sentence. Let us say that if t a term, and A a sentence, δtA is a sentence too. If we now assume that a sentence of the form δtA is interpreted as follows,

- $[[\delta tA]]^{w,\phi,g} = 1$ iff $\forall d \in [t]^{w,\phi,g}: \ [[A[^t/_x]]]^{w,\phi,g[^x/_d]} = 1$,[37]

we can represent (18) as (19),

(19) $\forall_x(B(x) \wedge Meet(x, \theta r\hat{y}[B(y) \wedge y \neq x]), \delta r'(x)Greet(x, r'(x)))$.

According to this analysis (18) is predicted to be true iff every bishop who meets another bishop, greets *all* bishops he meets.[38,39]

[37] Where $A[^t/_x]$ is A with every occurrence of t replaced by fresh x, and where $[t]^{w,\phi,g}$ is the same as $[[t]]^{w,\phi,g}$, except for the uniqueness condition for descriptive pronouns.

[38] This analysis closely resembles the approach of Neale (1990). Notice that the distribution operator 'δ' is really a universal quantifier taking as domain the individuals in $[t]^{w,\phi,g}$. Van der Does (1994) has generalized this kind of analysis by allowing this quantifier also to be *existential*. I obviously don't like this generalization, because by adopting it we would give up the assumption that pronouns should always have an exhaustive interpretation.

[39] Notice that if we wanted we could also account for conditional donkey sentences in terms of descriptive pronouns and our distribution operator.

Not only can the interpretation of *descriptive* pronouns be dependent on the interpretation of other terms, it appears that this can be the case for *referential* pronouns, too. Consider Heim's (1990) example (20a) - (20b), for instance, that corresponds with Strawson's (1952) original case (5a) - (5b) which was one of the prime motivations for a referential analysis of anaphoric pronouns:

(5a) John: *A man* jumped off the bridge

(5b) Mary: *He* didn't jump, but was pushed.

(20) a. John: Every time I was here, *a man* jumped off the bridge

b. Mary: I bet that in most cases *he* didn't jump, but was pushed.

Just like the pronoun *He* in (5b) cannot be treated as a descriptive pronoun, referring to the/a man who jumped off the bridge, for reasons of inconsistency, also the pronoun *he* in (20b) cannot be treated as a descriptive pronoun. We have argued that for (5b) the obvious solution was to assume that the pronoun refers to the *individual* that the first speaker had in mind with his use of the antecedent indefinite. It is clear, though, that for (20b) this doesn't work; John had no specific individual in mind to which Mary can refer back with her use of the anaphoric pronoun. However, John could have a specific individual in mind for *every time* he was there. But this means that with his use of the sentence (20a) he could also have something specific in mind related with the indefinite: namely a (partial) Skolem function from times (that he was here) to a specific man.

Let us now assume that if x is the free variable in P, terms like θrP introduce in possibility $\langle w, \phi, g\rangle$ the function $f \in [D \to [W \to D]]$ associated with r such that for any d in the domain of f: $f(d) = \phi(I_{\phi,g[^x/_d]}(P))$. Now the above intuition can be accounted for immediately when we make the evaluation of predicates time-dependent, and represent (20a) as follows (considering times as special kinds of individuals):

(21) $\forall_t(Here(i,t), Jump(\theta r\hat{x}Man(x,t),t))$

Because the theta term contains the free variable t, the Skolem function introduced by the theta term (if we fix the world) will then indeed be a function from times to men. It is then possible for Mary to pick up this

[39] Notice that this rule is indeed just a generalization of our definition of $Upd(\theta rP, \langle w, \phi, g\rangle)$ in section 2.2.

Skolem function to determine the referent of the pronoun *he* in (20b), and the values of this function need not have jumped off the bridge at the relevant times.[40,41]

When we assume that the 'referents' of specifically used indefinites, and the pronouns that refer back to them, can depend on the values of certain variables, we can also account for some examples that were supposed to be problematic for the standard E-type analysis of pronouns. First, when we allow indefinites to introduce specific Skolem functions, we can easily account for certain *asymmetric* and *weak* readings of quantificational donkey sentences (cf. Von Heusinger, 1997). For instance, when we represent (22a) as (22b),

(22) a. Most of the time when I have *a dime* in my pocket, I put *it* in the parking meter.

b. $Most_t(Time(t) \wedge InP(\theta r \hat{x} D(x,t),t), PinM(r(t),t))$

the sentence is predicted to be true when most of the time when I have at least one dime in my pocket, I will put (at least) one of them in the meter. Notice that although for the sentence to be true at a particular moment not all dimes that I have in my pocket at that moment need to be involved, there is still a notion of uniqueness or exhaustivity involved: at each occasion the unique dime is determined by the relevant reference function.

By means of Skolem functions we can also account for some cases with 'indistinguishable antecedents' like

(23) Every woman who bought *a sage plant* <u>*she*</u> *liked*, bought eight others along with *it*.

We can analyze this sentence without giving up the assumption that with singular pronouns there is always a notion of uniqueness or exhaustivity involved, when we make use of Skolem functions, and represent the sentence as follows:

(24) $\forall_x(Wx \wedge B(x, \theta r(\hat{y}[S(y) \wedge L(x,y)])), \exists 8\hat{z}[S(z) \wedge B(x,z) \wedge z \neq r(x)])$

Sentence (23) is of course very close to Heim's (1982) original sage plant-example, (25), that was used as one of the prime arguments to show that the E-type analysis was problematic.

[40] The pronouns that involve such functions will be represented by something like $r(t)$, and are interpreted in $\langle w, \phi, g\rangle$ as $g(r)([[t]]^{w,\phi,g})(w)$.

[41] Heim (1990) proposes a closely related analysis, but in a somewhat more *ad hoc* fashion.

(25) Every woman who bought *a sage plant* here, bought eight others along with *it*.

Kadmon (1990) argued that the E-type analysis can be rescued, when we assume that for such examples we associate with the pronoun *it* a function from sage plant buyers to sage plants. In our approach this can be implemented systematically when we represent the noun 'woman who bought a sage plant' as $Wx \wedge E(\theta r(\hat{y}[S(y) \wedge B(x,y)]))$, where E is the existence predicate.[42]

4. Diagonalization and the status of discourse referents

In section 2.1 I have argued that we should take the notion of speaker's reference more seriously in dynamic semantics to account for some *empirical* phenomena that are problematic for the standard theories. In this section I will argue that once we do so, we also gain an important *conceptual* advantage; we will be able to give a natural account of what the discourse referents used in dynamic semantics stand for.

[42] Heim (1990) admitted that Skolem functions might help to account for some (apparently) problematic examples for the E-type analysis, but noted that such an analysis would not help to account for the right truth-conditions for a sentence like

(i) No parent with *a son still in high school* has ever lent *him* the car on a weekend.

For suppose that Mary has two sons in high school, and that she has lent one of the two boys, but not both, the car on a weekend. On the assumption that for the truth of the sentence we have to quantify over the reference functions, and thus indirectly over the Skolem functions, it will be predicted that the sentence can be true on our account, because Mary never lent the car to one of her two sons, and will neither be true nor false on the supervaluation account, although the sentence seems to be unambiguously false in such circumstances.

However, when we now assume that adverbials like *ever* quantify over *events*, and that events are special kind of individuals, we might represent our problematic (i) as (ii), on the assumption that e denotes the contextually given event:

(ii) $No_x(P(x,e) \wedge H(x,\theta r\hat{y}[Sof(y,x,e) \wedge HS(y,e)],e), \exists e'[Before(e',e) \wedge LentCar(x,r'(x,e'),e')])$

Notice that in the above formula, the discourse referent s will denote a function from an individual and an event to the sons of that individual who are still in high school in that event. The sentence will be predicted to be false in the circumstances sketched above, because there is a parent, i.e., Mary, with a son who is still at high school, such that there will be an event before the contextually given one where her unique son *at that event* lent at that event a car for the weekend. Of course, for such an analysis to work, it has to be the case that an event is not just a time-slice of a world, but rather something like a time-*place*-slice.

4.1. Unclarity of reference and successful communication

I have argued above that a lot of pronouns are used referentially, referring back to the speaker's referents of their antecedent indefinites. In this way it can be explained how a notion of uniqueness is involved for these pronouns. But the resulting treatment seems to have an unwelcome consequence which is avoided on standard accounts. A common assumption in the philosophy of language is that in determining the referents of referential expressions, one can represent a context by an n-tuple of objects, and that it is clear to both speaker and hearer what this context is. This latter assumption is based on the thought that speakers ought to assume that hearers have enough information to determine what proposition they have expressed. If the hearer fails to recognize what object is referred to by a referentially-used expression, then she cannot determine what proposition is expressed by the speaker, who thus violates the conversational maxim. It seems to follow that if some anaphorically-used pronoun is treated as a referential expression, the speaker has to presuppose that the hearer can recognize what object the speaker is intending to refer to by the use of the pronoun. Otherwise, the hearer will not understand what is meant by the sentence in which the pronoun occurs. Unfortunately, this is commonly not the case, and the hearer cannot tell which object the speaker has been intending to refer to with the indefinite or pronoun. But how, then, can communication be successful?

Below I will first show how we *can* account for successful communication on the assumption that many occurrences of anaphoric pronouns are being used referentially, and afterwards I will argue that we *should* explain these anaphoric relations in this way in order to explain the status of discourse referents in information states used to analyze discourses.

4.2. Bridging the gap by diagonalization

According to the causal/historical theory of reference, the referents of certain terms used by the speaker are determined by the 'causal' relations that the speaker bears to the world. Normally this is assumed only for proper names and demonstratives; in this paper, however, I have argued that this also holds for most (other) uses of anaphoric expressions. But just as agents might be unclear about what the referent of a proper name or demonstrative is, because they are unclear about the origin of the relevant referential chain, agents might also be unclear about the referent of a pronoun, because they are unclear about the causal origin of the relevant anaphoric chain. In Stalnaker (1978) it is argued that we can describe how successful communication is achieved despite the

uncertain reference of proper names and demonstratives by means of *diagonalization*. In this section I want to argue that the gap between the unclear reference for pronouns and successful communication should be bridged by diagonalization too.

Ideally, a referential expression is used only when it is clear to the hearer what the expression refers to. It is clear, however, that ideal conditions do not always obtain. If the speaker says something and the hearer disagrees, there might be *two* reasons for this disagreement. First, the hearer might have understood what the speaker has said, but he disagrees with the speaker about the facts that the discourse is about. Second, speaker and hearer might agree about these facts, but disagree because the hearer thinks that the speaker has said something different from what the latter has actually intended to say. The latter situation might obtain if the speaker uses a referential expression. These two different reasons for disagreement can be accounted for in the *two-dimensional* theory of reference proposed by Kaplan (1989) and Stalnaker (1970). The reason is that in this theory a conceptual distinction is made between two kinds of facts: (i) facts about the *subject matter* of conversation, and (ii) facts about the *conversational situation* itself. What is expressed by a sentence, then, might depend on the facts of the conversational situation. Suppose Hans says *I will see you at 10 o'clock tomorrow* in a conversation with Ede and Paul. Although Hans intends to refer to Ede by his use of the demonstrative pronoun *you*, Paul might react by saying *No, because I will take the train to Amsterdam this evening*. In this case, Paul need not disagree with Hans about the facts relevant to the conversation's subject matter, but can nevertheless have misunderstood what Hans has intended to say because Hans's use of *you* has been accompanied by an unclear pointing gesture. If we say that reference contexts represent facts about the conversational situation, we can think of a reference context in this simple situation as a possible referent of the demonstrative pronoun *you*. Clearly, there are two possible referents, Ede and Paul. In a two-dimensional theory of reference, we can represent what Hans has said as a function from reference contexts to the proposition expressed by *I see you tomorrow* in this reference context: $\{\{w \in W|$ Hans will see a tomorrow in $w\} : a =$ Ede or $a =$ Paul$\}$. Of course, this function from reference contexts to propositions is formally a Kaplanian (1989) *character*.

If $\langle w_0, c_0\rangle$ is the actual world/reference context pair and A a sentence, the actual *horizontal* proposition expressed by A – that is, the set of *index worlds* where what is expressed by A with respect to the actual conversational situation is true – is determined as follows: $\{w \in W|\ [[A]]^{w,c_0} = 1\}$. Although it is normally the horizontal propo-

sition that the speaker intends to express, the hearer, as we have seen, doesn't always know which one this is because he does not know the relevant facts about the conversational situation. But even if the hearer doesn't know which horizontal proposition is expressed by the sentence, the information that he receives from the sentence can still be modeled by a set of possibilities.

A context that represents what is *presupposed* by the participants of a conversation should contain not only the information available for the interpretation of context-dependent utterances, but also the information accepted by speaker and hearer about the *subject matter* of the conversation. In a two-dimensional framework this means that a context, C, should be represented by a set of world/reference context pairs. Any element of C might, as far as the hearer can tell, be the actual possibility that makes true everything that is presupposed, both about the subject matter of conversation and about the conversational situation itself. If any element of C might, as far as the hearer can tell, be the actual possibility, he might update this information state C after accepting the utterance of A by eliminating any possibility $\langle w, c\rangle$ in C in which what is expressed by A in c is false in w. This new information, or presupposition, state, is $\{\langle w, c\rangle \in C|\ [[A]]^{w,c} = 1\}$, and is what Stalnaker (1978) has called the *diagonal* expressed by A with respect to C.

Stalnaker (1978) proposes that each time we can assume that the speaker assumes that it is unclear for the hearers which horizontal proposition is expressed by the use of a sentence, we should reinterpret what is said and assume that it is the diagonal proposition that the speaker has intended to communicate. Just as Stalnaker (1978) explains by means of diagonalization how the identity statement *Hesperus is Phosphorus* can be used successfully and informatively, although names are referential expressions, we can explain the successful use of anaphoric pronouns by means of diagonalization too.[43] According to this diagonalization solution, successful (enough) communication does not require there to be a unique individual that is the referent of a referential expression in all worlds/possibilities consistent with what is presupposed. For diagonalization to apply, it is only required that in each world/possibility consistent with what is presupposed the relevant expression must have a *unique* referent. In our case this means that it is required that there is a unique individual associated with each indefinite used specifically and pronoun used referentially in all possibilities

[43] In a very inspiring discussion, Sommers (1982) even suggests that we should reduce the referentiality and rigidity of proper names to that of (referentially used) pronouns. According to him, we should think of proper names as special purpose pronouns, pronouns that can be used in more than one conversation.

consistent with what is presupposed. Notice that this condition is not met in theories like DRT, FCS or DPL, but is in our analysis due to the use of reference functions. Thus, by means of diagonalization we can explain how successful communication can be achieved despite the uncertain reference of anaphoric pronouns.

In our case, a reference context is modeled by a pair consisting of a reference function and an assignment. It seems, thus, that after the acceptance of A in initial context C, the new information state should be modeled as $\{\langle w, \phi, g\rangle \in C |\ [[A]]^{w,\phi,g} = 1\}$. But this is obviously not what should be the case if we want to account for cross-sentential anaphora. The new information state should rather be $\{\langle w, \phi, Upd(A, \langle w, \phi, g\rangle)\rangle |\ \langle w, \phi, g\rangle \in C \ \&\ [[A]]^{w,\phi,g} = 1\}$. This, of course, is very close to what is normally assumed in dynamic semantics, though not the same.

4.3. The status of discourse referents and possibilities

In this section I will argue that our treatment of the antecedent-anaphor relation allows us to say that (the information associated with) a discourse referent in an information state is the representation of the (presumed) speaker's referent of the associated indefinite. In fact, I will argue that (i) a discourse referent is not only the representation in its information state of the *actual* speaker's referent, but also (ii) that it is *presupposed* that it represents this speaker's referent. The argument for the second claim will be somewhat more abstract than the argument for the first claim, so I will begin with the first one.

4.3.1. *Discourse referents represent speaker's referents*

According to the *functional* analysis of attitudes, an agent stands in a certain attitude relation to a proposition if by means of this relation, together with the assumption that the agent is rational, we can explain the agent's behavior. Attitudes are dispositional, or functional, states of a rational agent; and these states are individuated by the role that they play in determining the behavior of the agent who is in such a state. This picture (see Stalnaker, 1970) suggests that presupposition should also be thought of as a propositional attitude: we have to know what the speaker is presupposing in order to explain his behavior when he is engaged in a conversation.[44] The alternative possibilities helping to represent what the speaker is presupposing are the relevant

[44] With Stalnaker, I will assume that presupposition as a propositional attitude is a more basic notion than the semantic presupposition relation between sentences or propositions triggered by specific lexical items; and that the latter relation should be explained in terms of the former.

alternatives consistent with what the speaker assumes is commonly assumed, and are the alternatives with respect to which we must judge the informativity and acceptability of the speech acts made by speakers.

Notice that both the information states we used in section 4.2, and the information states of the kind used in standard dynamic semantics, are (supposed to be) representations of what is presupposed by the participants of a conversation. In both cases, what is presupposed is represented by a set of possibilities, and contains information about the values of discourse referents. It is sometimes argued that the possibilities in information states used in dynamic semantics should be finer-grained than possible worlds, because they represent something about the discourse that is going on in the actual world.[45] But what exactly do these possibilities represent about the discourse situation? Before we will answer this question, let us first have a closer look at contexts and the information associated with discourse referents.

Until now I have assumed that a possibility should be modeled by a triple consisting of a world, a reference function, and an assignment. If we limit ourselves to the singular case, this assignment assigns in each world a single individual to each discourse referent in its domain. This means that for each world, an assignment can also be thought of as a *sequence of individuals* that are the values of their respective discourse referents. If we write the variables in the domain of g in a particular order, we, thus, might rewrite possibility $\langle w, \phi, g\rangle$ as something like $\langle w, \phi, d_1, ..., d_m\rangle$, when the domain of g has m members. Each possibility in an information state that results during the interpretation of a discourse and represents what is presupposed might be the actual possibility, and, indeed, there will also be a possibility (a possibility that can, but need not, be an element of this information state) that figures as the *actual* possibility. Suppose that $\langle w_0, \phi_0, d^0_1, ..., d^0_m\rangle$ is, in fact, the actual possibility. In that case, each of the d^0_i will be the actual speaker's referent of the i'th used indefinite.[46] If we assume that C^0 is the information state with n possibilities that represents the discourse that is taking place in possibility $\langle w_0, \phi_0, d^0_1, ..., d^0_m\rangle$, it will also be the case that each element of C^0 will contain m individuals such that for each $j : 1 \leq j \leq n$, each d^j_i will be an individual that, as far as the hearer can tell, could have been the actual speaker's referent of the i'th indefinite. Suppose now that this indefinite introduced discourse referent r_i to the discourse. We might then think of the properties that all the i'th individuals of the possibilities of the information state C^0 share as the information associated with r_i in information state C^0.

[45] See Zimmermann (1998) for discussion.

[46] Some of the 'individuals' d^0_i, however, might be the impossible object, $*$.

For instance, if discourse referent r_i is introduced to the discourse due to the indefinite *a man* in the sentence *A man is walking in the park*, a property shared by all i'th individuals in the possibilities of C^0 will be that of being a man who is walking in the park.

In fact, it turns out that the information associated with a discourse referent can be modeled as a function from the possibilities in C^0 to individuals, a kind of individual concept. The trick, extensively used in Dekker (1996), is to think of discourse referents as functions from possibilities to individuals. Dekker calls these functions the *subjects* of C^0 associated with the discourse referents. The information that context C^0 associates with discourse referent r_i, i.e. the subject of C^0 associated with r_i denoted by $[r_i]_{C^0}$, can now be defined as follows:

$$[r_i]_{C^0} \stackrel{def}{=} \text{the function } f \in [C^0 \rightarrow (D \cup \{*\})] \text{ s.t. } \forall \alpha \in C^0 : f(\alpha) = [[r_i]]^\alpha$$

Thus, we might now say that each function $[r_i]_{C^0}$ represents the information associated with discourse referent r_i in information state C^0.

But what does such a function represent about the actual possibility? It is not clear how to answer this question for proponents of standard dynamic semantics. But if we assume that the actual possibility contains enough information to determine the actual speaker's referent (if there is one) of a specifically used indefinite, as we do, our question regarding what a subject represents about the actual possibility has a straightforward answer. We might say that each function $[r_i]_{C^0}$ is the representation of the actual speaker's referent (if there is one) d_i^0 in the state C^0 that represents what is presupposed about the discourse.[47] So, just like C^0 represents what is presupposed about the actual possibility $\langle w_0, \phi_0, d_1^0, ..., d_m^0 \rangle$, so does subject $[r_i]_{C^0}$ represent what is presupposed about the actual speaker's referent d_i^0.

4.3.2. *Presupposed representation of speaker's referent*

In this subsection I will argue that a subject not only actually represents something about a speaker's referent, but that this is also *presupposed* to be the case.

To explain the actions of rational agents, we must normally assume that believers know their own minds, i.e. have *introspective* access to their own minds; if an agent believes or doesn't believe something, he also believes that he does or doesn't believe it. I have argued above that

[47] In the terminology of Kamp (1990), we might think of the subject $[r_i]_{C^0}$ as being *anchored* to object d_i^0. See also Dekker (1997) for a somewhat similar notion. In terms of counterpart theory (e.g. van Rooy, 1997) we might say that each value of r_i in the possibilities of C^0 is the *counterparts* of d_i^0 in this possibility.

presuppositions, just like beliefs, should be thought of as propositional attitudes, needed to explain the communicative actions of agents. But if speech is action, and if the appropriateness of the speech acts of agents is to be explained partly in terms of what they presuppose, then we also have to assume that the attitude of presupposition is liable to introspection, so that if an agent presupposes something, he also presupposes that he presupposes this something, and if he doesn't presuppose something, he presupposes that he doesn't presuppose it.

Just as each world of a belief state might, as far as the agent believes, be the actual world, each element of the context (the possibilities consistent with what is presupposed) might, as far as the participants in a conversation assume, be the actual possibility where the discourse is taking place. In each of those possibilities something is presupposed.[48] So, not only for the actual possibility $\langle w_0, \phi_0, d^0_1, ..., d^0_m \rangle$ there is a context C^0 that represents what is presupposed in this possibility, but we should associate such a context with each possibility in C^0. As a result, each possibility α in C^0 should be represented by something like $\langle w_\alpha, \phi_\alpha, d^\alpha_1, ..., d^\alpha_m, C^\alpha \rangle$. The assumption that the presupposition state, C^0, is introspective means that it will be the case for each such α it holds that $C^\alpha = C^0$.

In the previous section we have seen that from an information state like C^0, we can extract a set of *subjects*, a set of functions representing the information associated with discourse referents. If each possibility α in C^0 also contains a context C^α, also from each of those C^α's we can extract a set of subjects. In fact, our introspection condition demands that for each α in C^0 it holds that the set of subjects of C^α is the same as the set of subjects of C^0. In the previous section I have argued that subjects in C^0 represent something about specific objects in the actual possibility $\langle w_0, \phi_0, d^0_1, ..., d^0_m, C^0 \rangle$. Now I want to argue that each subject in C^α also represents something about a specific object in possibility α.

Suppose that possibility α is consistent with what is presupposed, and represents what a possibility represents in standard dynamic semantics. Suppose in addition that in the world of the possibility there are two men walking in the park, and we are looking at an information state resulting from the update of an earlier information state with the assertion *A man$_r$ is walking in the park*. The question that arises now is what is represented about possibility α by the subject associated with

[48] That the possibilities used in dynamic semantics should contain the information that is presupposed in the conversation is something that I learned from Fernando (1995) and Stalnaker (1998). But Fernando, at least, does not draw from this the conclusion that I will argue we should draw: that the pronouns analyzed in dynamic semantics should be treated as referential expressions.

r in C^α? Proponents of standard dynamic semantics must deny the usefulness of this question, but as we have seen in the previous section, at least in the actual possibility the question does seem to make sense. It seems natural that to answer the questions for the actual possibility, we would look at the specific individual the speaker had in mind by his use of the antecedent indefinite. But if this is so for the actual possibility, why not, then, also for possibilities compatible with what is presupposed to be the actual possibility, as our α?

So far, we have not said much about the fact that each possibility α in C^0 also contains a reference function. This function will be crucial now, because it assigns a unique individual in that possibility to each specifically used indefinite, namely its speaker's referent (in that possibility). Due to our assumption that possibilities also contain reference functions, we can now answer our above question what each subject $[r_i]_{C^\alpha}$ represents about possibility α: Each subject associated with discourse referent r_i is not only the representation in C^0 of d_i^0, but (due to the introspection condition on C^0) also the representation in each C^α of d_i^α, for any α in C^0. This accounts for the intuition that the subject is not only the representation of the *actual* speaker's referent (if there is one) of the indefinite that introduced r_i to the discourse, but also that it is *presupposed* that it is the representation of the actual speaker's referent.[49]

The question that arises now is how we should account for the introspectiveness of presupposition states, such that also after the acceptance of a new assertion the presupposition state remains introspective. This is quite a complicated task, but when we assume that after acceptance of assertions in a discourse only the presupposition state changes, things can be worked out rather straightforwardly by making use of recent work of Gerbrandy and Groeneveld (1997).

According to our analysis so far, a possibility should be represented by a quadruple of the form $\langle w, \phi, g, C\rangle$, where C represents what is presupposed in the possibility. The same information could also be represented by the triple $\langle w, \phi, g\rangle$, if we introduce a distinguished variable, p, to our language, and assume that assignments assign to this variable a set of world/reference function/assignment triples. Let us now assume that $\langle w_0, \phi_0, g_0\rangle$ is the actual possibility, where $g_0(p)$ denotes the set of possibilities consistent with what in this possibility is presupposed, i.e. our former C^0. Notice that not only the assignment in the actual

[49] Just like I did in the previous section, also Dekker (1997) seeks to account for the intuition that subjects of presupposition states represent what is presupposed about the individuals the speaker's had in mind for their use of the indefinites. But he does this in a different way than I do, however, and cannot treat what I account for in this section.

possibility, but also each assignment of the possibilities in $g_0(p)$ assigns to p a set of possibilities. By our introspectiveness condition it follows that for each $\langle v, \psi, h\rangle$ in $g_0(p)$ it holds that $h(p) = g_0(p)$. Observe that this, in turn, has the result that for each $\langle v, \psi, h\rangle$ in $g_0(p)$ it holds that $\langle v, \psi, h\rangle \in h(p)$. This latter fact would be problematic when we used standard set theory, but, as shown by Gerbrandy and Groeneveld, is okay when we use Aczel's *non-wellfounded* set theory.

This is not the place to go into non-wellfounded set theory, nor to discuss how Gerbrandy and Groeneveld (1997) use it in update semantics to account for introspection (for the propositional case). Still, I would like to sketch how we could change our definitions to account for introspection of our presupposition states. In fact, the only thing we really need to do if we assume that only the presupposition state changes after the acceptance of an assertion is to change the definition of '$Upd(A, \langle w, \phi, g\rangle)$' so that a possibility is updated with the terms introduced by A; and that it is *presupposed* that A is true and accepted and that the referents of the terms of A are also introduced. For atomic sentence A, $Upd(A, \langle w, \phi, g\rangle) = g[p \rightarrow \{\langle v, \psi, Upd(A, \langle v, \psi, h\rangle)\rangle : \langle v, \psi, h\rangle \in g(p) \ \& \ [[A]]^{v,\psi,h} = 1\}]$. Thus, after the interpretation of, for instance, a sentence of the form $P(t)$, the possibility is enriched by the object introduced by term t, and $P(t)$ is presupposed after the update to be true *and presupposed*. Due to this richer representation of possibilities, the discourse referents used in the information states not only represent the speaker's referents associated with the relevant indefinites, but it is also presupposed that they represent their speaker's referents.

5. Conclusion

In this paper I have defended the view that pronouns should always be interpreted exhaustively by (i) suggesting that some empirical phenomena are problematic for the non-exhaustivity treatment of pronouns in standard dynamic semantics, (ii) showing that the referential analysis of pronouns can be pushed much further than is usually assumed, and implementing it in a dynamic semantics, (iii) combining this referential analysis with an account that treats pronouns as abbreviations for their antecedent clauses, and finally (iv) arguing that when exhaustivity/uniqueness is assumed, we can give a natural answer to the question what the discourse referents used in information states represent.

Acknowledgements

I am endebted to Stalnaker's writings on this subject, as is obvious. Discussions with and comments from Maria Aloni, Paul Dekker, Jelle Gerbrandy, Jeroen Groenendijk, Hans Kamp, Peter Krause, Henk Zeevat, some anonymous reviewers from L & P, and especially Ede Zimmermann were also important for writing this paper. I would also like to thank Ben Shaer and Jason Mattausch for correcting my English. The paper was written as part of the 'Sources and Streams of Information' project, sponsored by the Dutch Organization for Scientific Research (NWO), which is gratefully acknowledged.

References

Chastain, C.: 1975, 'Reference and context', In: K. Gunderson (ed.), *Menesota Studies in the Philosophy of Science, vol. VII- Language, Mind and Knowledge*, Minneapolis: University of Menneasota Press.

Chierchia, G.: 1992, 'Anaphora and dynamic binding', *Linguistics and Philosophy*, **15**: 111-183.

Chierchia, G.: 1996, *Dynamics of Meaning, Anaphora, Presuppositions and the Theory of Grammar*, University of Chicago Press, Chicago.

Cooper, R.: 1979, 'The interpretation of pronouns', In: F. Heny and H.S. Schnelle (eds.), *Selections from the third Groningen round table, Syntax and Semantics*, 10, Academic Press, New York.

Dekker, P.: 1993, *Transsentential Meditations*, Ph.D. thesis, University of Amsterdam.

Dekker, P.: 1996, 'The values of variables in dynamic semantics', *Linguistics and Philosophy* **19** 211-257.

Dekker, P.: 1997, 'On first order information exchange', In: A. Benz and G. Jaeger (eds.), *Proceedings of Mundial '97. Munich Workshop on the Formal Semantics and Pragmatics of Dialogue*, CIS, München.

Dekker, P. & R. van Rooy: 1998, 'Intentional identity and information exchange', In: R. Cooper and T. Gamkrelidze (eds.), *Proceedings of the Second Tbilisi Symposium on Language, Logic and Computation*, Tbilisi State university.

Does, J. van der: 1994, 'Formalising E-type logic', In: P. Dekker and M. Stokhof (eds.), *Proceedings of the Ninth Amsterdam Colloquium*, Amsterdam, 229-248.

Donnellan, K.: 1978, 'Speaker reference, descriptions, and anaphora', In: P. Cole (ed.), *Syntax and Semantics, vol. 9: Pragmatics*, New York: Accadamic Press, 47-68.

Evans, G.: 1979, 'Pronouns, quantifiers and relative clauses (1)', *The Canadian Journal of Philosophy* **7**: 467-536.

Fernando, T.: 1995, 'Are context change potentials functions?' In: H. Kamp and B. Partee (eds.), *Context in the Analysis of Linguistic Meaning*, Stuttgart/Prague.

Fodor, J. and I. Sag: 1982, 'Referential and quantificational indefinites', *Linguistics and Philosophy* **5**: 355-398.

Fraassen, B. van: 1967, 'Meaning relations among predicates', Nous **1**: 161-179.

Forbes, G.: 1985, *The Metaphysics of Modality*, Clarendon Press, Oxford.

Gerbrandy, J. and W. Groeneveld: 1997, 'Reasoning about information change', *Journal for Language, Logic and Information* **6**: 147-169.

Groenendijk, J. and M. Stokhof: 1990, 'Dynamic Montague Grammar', In: L. Kalman and L. Polos (eds.), *Papers from the second symposium on logic and language*, Budapest, 1990. Adademia Kiado.

Groenendijk, J. and M. Stokhof: 1991, 'Dynamic predicate logic', *Linguistics and Philosophy* **14**: 39-100.

Groenendijk, J, M. Stokhof, and F. Veltman: 1996, 'Coreference and modality', In: Lappin (ed.), *Handbook of Contemporary Semantic Theory*, Oxford, Blackwell.

Heim, I.: 1982, *The Semantics of Definite and Indefinite Noun Phrases*, Ph.D. dissertation, University of Massachusetts, Amherst.

Heim, I.: 1990, 'E-type pronouns and donkey anaphora', *Linguistics and Philosophy* **13** 137-178.

Heusinger, K. von: 1997, 'The reference on indefinites', In: K. von Heusinger and U. Egli (eds.), *Reference and Anaphoric Relations*, Kluwer, Dordrecht.

Kamp, H.: 1981, 'A theory of truth and semantic representation', In: Groenendijk et al. (eds.), *Formal Methods in the Study of Language*, Amsterdam, 277-322.

Kamp, H.: 1988, 'Comments on Robert Stalnaker: 'Belief attribution and context'¿ In: R. Grimm and D. Merrill (eds.), *Contents of Thought*, Tuscon, University of Arizona Press.

Kamp, H.: 1990, 'Prolegomena to a structural account of belief and other attitudes', In: R. Grimm and D. Merrill (eds.), *Contents of Thought*, Tuscon, University of Arizona Press.

Kamp, H. and U. Reyle: 1993, *¿From Discourse to Logic*, Kluwer, Dordrecht.

Kaplan, D.: 1989, 'Demonstratives', In: I. Almog et al. (eds.), *Themes from Kaplan*, Oxford University Press.

Krahmer, E. and R. Muskens: 1995, 'Negation and disjunction in discourse representation theory', *Journal of Semantics* **12**: 357-376.

Kripke, S.: 1977, 'Speakers reference and semantic reference', *Midwest Studies in Philosophy, II*, 255-276.

Lewis, D.: 1979, 'Scorekeeping in a language game', *Journal of Philosophical Logic* **8**: 339-359.

Montague, R.: 1974, 'Pragmatics', In: R. Thomason (ed.), *Formal Philosophy*, Yale University Press, New Haven, 95-118.

Neale, S.: 1990, *Descriptions*, MIT Press, Cambridge.

Partee, B.: 1989, 'Binding implicit variables in quantified contexts', In: *Papers from the 25th regional meeting: Parasession on Language in Context*, Chicago Linguistic Society, Chicago.

Quine, W.V.: 1960, *Word and Object*, Cambridge, Mass.

Rooy, R. van: 1997, *Attitudes and Changing Contexts*, Ph.D. thesis, University of Stuttgart.

Rooy, R. van: 2000, "Anaphoric relations across attitude contexts", In: K. von Heusinger and U. Egli (eds.), *Reference and Anaphoric Relations*, Kluwer, Dordrecht.

Sommers, F.: 1982, *The Logic of Natural Language*, Oxford: Clarendon Press.

Stalnaker, R.: 1970, 'Pragmatics', *Synthese* **22**: 272-289.

Stalnaker, R.: 1978, 'Assertion', In: P. Cole (ed.), *Syntax and Semantics, vol. 9: Pragmatics*, New York: Academic Press, 315-332.

Stalnaker, R.: 1998, 'On the representation of context', *Journal of Logic, Language, and Information* **7**: 3-19.

Strawson, P.: 1952, *Introduction to Logical Theory*, London, Methuen.

Zimmermann, T. E.: 1998, 'Remarks on the epistemic role of discourse referents', In: L. Moss (ed.), *Logic, Language and Computation, Vol 2*, CSLI publications, Stanford.