Preface

Choice Functions in Semantics

KLAUS VON HEUSINGER[1] AND RUTH KEMPSON[2]

[1]*Institute for General and German Linguistics, University of Stuttgart, Keplerstr. 17 D-70174, Stuttgart (E-mail: vonheusinger@ilg.uni-stuttgart.de);* [2]*King's College London, The Strand, London, WC2R 2LS (E-mail: kempson@dcs.kcl.ac.uk)*

Interest in the use of choice functions is increasing in formal semantics, demonstrated by recent discussions of the adequacy of analyses with choice functions. Their syntactic counterpart, the epsilon operator, was introduced into meta-mathematics in the epsilon calculus of Hilbert and Bernays (1939), which provides an explicit study of arbitrary names used as predicate logic proof terms, the epsilon operator being a generalized iota operator underpinning uses of both existential and universal quantification. Semantically, a choice function Φ is a function that assigns to a non-empty set s one of its elements, as defined in (1) or alternatively in (2). Intuitively, a choice function selects one element out of a set.

(1) $\Phi(s) \in s$ if $s \neq \phi$
(2) f is a choice function (i.e. CH(f) holds) iff P(f(P)), where P is non-empty.

This very general characterization makes choice functions an attractive and flexible semantic tool. For instance, choice functions, like Skolem functions, allow us to interpret the linguistic expressions associated with them *in situ*. The specific reading of the indefinite NP in (3) can be interpreted *in situ* if the indefinite article is associated with a choice function which takes wide scope, as illustrated in (4).

(3) Every student read a book.
(4) CH(f) & $\forall x$[student(x) $\rightarrow$ read(x, f(book)]

The range of application of choice functions in formal semantics has not yet been fully determined. The following contributions use choice functions as formal tools in as different domains as anaphoric definite descriptions, focus semantics, indefinite NPs, and conditionals, and there is also discussion of pragmatically related issues.

This special issue on "Choice Functions in Semantics" stems from papers delivered to the workshop "Choice Function and Natural Language Semantics" of the XIII European Summerschool in Logic, Language and Information in Helsinki (August 2001). We would specially like to thank the reviewers for their detailed and constructive comments on earlier versions of all papers.

Reference

Hilbert, D., Bernays P. [1939] (1970) *Grundlagen der Mathematik*. Vol. 2. 2nd edition Berlin; Heidelberg; New York: Springer.

Choice Functions and the Anaphoric Semantics of Definite NPs

KLAUS VON HEUSINGER
*Institute for General and German Linguistics, University of Stuttgart, Keplerstr. 17, D-70174
Stuttgart (E-mail: vonHeusinger@ilg.uni-stuttgart.de)*

Abstract. This article argues that definite NPs are interpreted depending on *contextual salience*, rather than on the uniqueness condition of their descriptive content. The salience structure is semantically reconstructed by a global choice function that assigns to each set one (most salient) element. It is dynamically modified by the context change potential of indefinite *and* definite NPs. The anaphoric potential of definite NPs can be accounted for by the interaction of the context change potential and contextual salience structure.

Key words: anaphora, choice functions, context change, contextual salience, definite NPs, semantics

1. Introduction

Since Russell's (1905) seminal paper on definite descriptions, definite NPs are interpreted as complex quantifiers that determine their referents by the uniqueness of their descriptive content. This semantic representation has not been changed, even though recent theories have developed new ways of analyzing the semantics of indefinite NPs, such as using choice functions or assigning context change potentials to indefinite NPs.

Choice functions are commonly used for representing indefinite NPs in LF for reasons of scope behavior, while definite NPs are analyzed according to Russell's classical theory. This view, however, is restricted to sentence semantics and, therefore, assumes static meanings of definite and indefinite NPs. Once we extend our analysis to (small fragments of) discourses, the picture changes dramatically – indefinite expressions receive a context change potential, while anaphoric definite expressions must be interpreted according to the updated context. This is the approach of dynamic semantics, such as File Changes Semantics, Discourse Representation Theory, or Dynamic Predicate Logic. Despite the "dynamic turn", definite descriptions aren't given interesting context change potentials. All dynamic theories give Russellian treatments of definite descriptions, which requires that the descriptive content determines uniqueness in a model. This won't work and

we need to get an account of definite descriptions which is consistent with non-uniqueness in a world but uniqueness within a (fine-grained) discourse context.

This paper extends a dynamic theory by a choice function approach for definite and indefinite NPs. Thus we will restore the parallelism between definite and indefnite NPs by analyzing both as having choice functions as their content and giving both similar (but distinct) context change potentials in terms of *contextual salience*. The motivation for such an extended semantic framework can be best illustrated by the discourse behaviour of definite descriptions as illustrated by Lewis.

1.1. LEWIS'S ARGUMENT AGAINST RUSSELLIAN UNIQUENESS

Lewis (1979, p.179) shows that the Russellian view of definite NPs cannot account for the different reference of the two occurrences of the definite NP *the cat* in (1) (indices are inserted by the author):

(1)   "the cat"
   i   Imagine yourself with me as I write these words. In the room is **a cat**$_1$, Bruce$_1$,
   ii who has been making **himself**$_1$ very salient by dashing madly about.
   iii **He**$_1$ is the only cat in the room, or in sight, or in earshot.
   iv I start to speak to you:
   v   The **cat**$_1$ is in the carton. **The cat**$_1$ will never meet **our other cat**$_2$,
   vi because **our other cat**$_2$ lives in New Zealand.
   vii **Our New Zealand cat**$_2$ lives with the Cresswells.
   viii And there **he**$_2$'ll stay, because Miriam would be sad if **the cat**$_2$ went away.

What is important to notice here is that this narrative contains two occurrences of the NP *the cat* which have different denotations. The situation being presented contains two individuals with the property of being a cat. While the first occurrence of the definite NP *the cat* in (1i) and in (1v) might be felicitously interpreted by a Russellian definite description, the last occurrence of the definite NP *the cat* in (1viii) shows that this cannot be the correct analysis. Still, the phrase occurs felicitously and refers to the second cat. Intuitively, this reference to the second cat is licensed by the fact that the second cat was introduced by the phrase *our other cat* and then made salient by talking about this second cat in (1vii–viii). The referent of the definite NP *the cat* is uniquely identifiable because it is the most salient cat, and not because it is the unique cat. Salience is understood as a property of the discourse that gives us for each set one (most salient) element of that set, rather than just a relation between discourse referents. Salience is not just created by talking about a certain individual, but by using definite NPs, and

the descriptive material of these definite NPs change the salience structure of the discourse.

We can account for the different denotations of the two occurrences of *the cat* in (1) by assuming the following context changes: The indefinite NP *a cat* in (1i) updates the context such that the subsequent term *he* in (1iii) and *the cat* in (1v) refer to that cat (let's call him Bruce). The second cat, let's call him Bobby, is introduced by the context change associated with the phrase *our other cat*. This new salient cat is available for further reference by the definite term *our New Zealand cat* in (1vii), which updates the context again: it makes the cat Bobby the most salient New Zealand cat, the most salient cat and the most salient (male) individual. Therefore, the definite pronoun *he* and the definite NP *the cat* in (1viii) refer to that cat Bobby in virtue of the assumption that they refer to the most salient object which meets their descriptive content.

1.2. STRUCTURE OF THE ARGUMENT

In the analysis presented below, both definite and indefinite descriptions will be interpreted via choice functions. Their analyses will differ in three important respects, however. First, indefinite NPs are represented by local or "minimalized" choice functions, while definite NPs are represented by global choice functions. Second, each indefinite NP introduces a new local choice function, while all definites are interpreted according to one global choice function. Third, the local choice functions for indefinite NPs are static, while the global choice function for definite is dynamic, i.e. it is updated in the discourse.

The values of all definite descriptions can be determined by a single choice function which is defined for many predicates. This single choice function serves simultaneously as a model of discourse salience and insures that the context-change potentials of all types of descriptions interact with each other. In this way, the global choice function can serve some of the functions that assignment functions do in other dynamic accounts. This global choice function as a contextual parameter is one principal innovation of this analysis. It is supplied from a global parameter of discourse. Thus definite NPs appear to take widest possible scope, rather than showing the interaction with other operators typical of the local choice functions which interpret indefinites. As a part of discourse context, the global choice function which fixes the interpretations of definite descriptions will be updatable. We discuss the context change potentials of all types of descriptions in terms of their update to the structure of salience in the discourse. A second principal innovation of this paper is to the context change potential of definite descriptions.

The paper is organized as follows: In section 2, I argue that we need global choice functions for the representation and interpretation of definite NPs. Definite NPs can only be interpreted according to the salience structure provided by the context. A global choice function stands for this salience structure. In section 3, I present some arguments for the use of local choice functions for indefinite NPs. These choice functions are only defined for the predicate by which they are locally introduced. This kind of choice function is discussed in the recent literature on choice function in natural language semantics. Furthermore, I motivate the assumption that both definite and indefinite NPs have the same context change potential. In section 4, I present a dynamic semantics based upon global choice functions and update functions of indefinite NPs, following Peregrin and von Heusinger (1995, 2004). In section 5, I extend this dynamic semantics by introducing update functions for definite NPs as well, which allows us to capture the semantics of anaphoric definite NPs.

2. Definite NPs and Global Choice Functions

2.1. THE CONCEPT OF SALIENCE

The concept of salience was introduced into the discussion of the semantics of definite NPs in the seventies (Lewis 1970, 1979). Lewis (1979, p. 178) uses it in order to replace Russell's problematic uniqueness condition for definite descriptions:

> The proper treatment of description must be more like this: 'the F' denotes x if and only if x is the most salient F in the domain of discourse, according to some contextually determined salience ranking.

The notion of salience itself seems to be influenced by the analysis of demonstrative expressions. A demonstrative like *this man* refers to the most prominent object in the physical environment of the speaker and hearer. Salience, however, does not depend only on the physical circumstances, or any other single cause. Rather it is a bundle of different linguistic and extra-linguistic factors, as noted by Lewis (1970, p. 63):

> An object may be prominent because it is nearby, or pointed at, or mentioned; but none of these is a necessary condition of contextual prominence. So perhaps we need a prominent-objects coordinate, a new contextual coordinate independent of the other.

In the following, salience assigned in a particular discourse context is assigned to one object relative to each set (or to each predicate).[1] The object so designated is the most salient or most prominent object of the extension of the predicate. We can therefore speak of 'the most salient F' in the context.

The present approach treats salience as a primitive which will not be analyzed further. The idea of salience was often criticized because of its pragmatic nature (cf. Heim 1982), however, an explicit formal account of salience and Lewis "prominent-object coordinate" was never seriously attempted, even though there are many different approaches towards the concept of salience [e.g. Sgall 1984 or Poesio and Stevenson (to appear)].

Here we propose to follow Lewis's salience solution to the problem posed above by the multiple cats by treating definite descriptions as terms, implementing a choice function analysis with epsilon operators.

2.2. Syntax and semantics of Hilbert's epsilon operator

The epsilon operator corresponds to a selection function that assigns to each non-empty set one element of this set.[2] Like the iota operator, the epsilon operator forms a term (constant) from a sentential form. Unlike the iota operator, it carries with it no existence or uniqueness presupposition as a condition of its reference. This is the key to solving Lewis's multiple cat problem. The main difference may be shown by the formalization and the paraphrase of the description *the island*, as given in (2) and (3):

(2) ιx [island(x)] the **unique** x, such that x is an island
(3) εx [island(x)] the **selected** x, such that x is an island

To introduce epsilon terms into a first-order predicate logic, we will adopt the axiom (4) by Hilbert and Bernays (1939, p.15), which they call the *epsilon formula*. From each formula of the form *Fa*, we can go directly to the corresponding formula F(εx [Fx]). The only new constant that has to be introduced is the symbol ε

(4) epsilon formula: Fa $\rightarrow$ FεxFx

Hilbert and Bernays did not give a semantic interpretation of their epsilon symbol, leaving this task for others. Schröter (1956) proposed interpreting the epsilon operator by a choice function. Asser (1957) then formulated this idea with the necessary detail. Following Asser we will interpret the epsilon operator by a (partial) choice function Φ, which assigns one of its elements to each set.[3]

We assume that models are pairs $\langle D, I \rangle = M$ with the domain of discourse D, an intepretation I of the constants. Denotations of expressions in the model are assigned relative to an assignment g of individuals to the variables as usual and in addition relative to a choice function Φ. The interpretation of an epsilon term $\varepsilon x\ \alpha$ is given by the following rule: $\|\varepsilon x\ \alpha\|^{M,g,\Phi} = \Phi(s)$, where s is the set of individuals $\{a : \|\alpha\|^{M,g^{x/a}} = 1\}$. An epsilon term is interpreted by applying the choice function Φ to the set of elements with the property $\lambda x\ \alpha$.

(5) $\|\varepsilon x\ \alpha\|^{M,g,\Phi} = \Phi(\{a\colon \|\alpha\|^{M,g^{x/a}} = 1\})$

To illustrate this, let us assume a domain of discourse, called "Lake Constance", and three objects in this domain, called "Mainau", "Reichenau", and "Lindau". All three individuals have the property of being an island. Let Φ be the choice function that assigns an arbitrary individual to the set of islands.

(6) $\Phi(\{\text{Mainau, Reichenau, Lindau}\}) \in \{\text{Mainau, Reichenau, Lindau}\}$

We only know that this element must be in the set of islands, but we do not know which island it is. This observation is often taken as argument for the "indefinite" character of epsilon terms, thus interpreting them as representing indefinite NPs. It is true that in the original epsilon theory, the choice of an element is arbitrary, but once the choice is made it is fixed for all subsequent expressions, which is the "definite" aspect of the classical epsilon calculus. This very general characterization makes epsilon terms and choice functions as their interpretations an attractive and flexible semantic tool that can reconstruct different linguistic categories. The operation of selecting one element out of a set (i.e. assigning one of its elements to a set) is common to all uses of the epsilon operator. This function very well captures the basic semantics of definite and indefinite NPs in their guise as terms. It corresponds to the assumptions of traditional grammar that the definite and indefinite article have an "individualizing" function. In order to distinguish between indefinite and definite NPs, we must modify the classical calculus.[4]

2.3. CONTEXT DEPENDENT EPSILON TERMS

Since Hilbert applied his classical epsilon terms only to the domain of numbers, a naturally ordered set, no determined choice function was necessary. However, in natural language the objects we refer to are not naturally ordered; rather, the order depends on a particular context. Egli (1991) approaches this problem by assuming a family of choice functions for representing definite NPs. Each context c has its own choice function Φ_C and the definite NP *the F* is represented as the epsilon term *$\varepsilon x\ Fx$ (in the context c)*, which can be paraphrased with *the selected x in the context c such that x is F* or *the most salient x in c such that x is F*. It is interpreted as the element that results from applying the choice function Φ_c to the set of all Fs. The "uniqueness commitment" of a definite NP is not understood as the uniqueness of the associated set, but as the "unique availability" of the referent (cf. Peregrin 2000). This is warranted by the definition of the choice function, which assigns one element to a set, independently of the size of this set.

Let us illustrate this point with our model "Lake Constance"; the property *island* is common to three objects: Mainau, Reichenau, and Lindau. The definite description *the island* is represented as the epsilon term *$\varepsilon x\ [island]$*. It

denotes different islands according to different situations. If we hear the expression from a Reichenau fisherman, he probably means the island Reichenau; if we encounter the same sentence during a guided tour through Lindau it will rather be the island Lindau that is meant; however, uttered by the Earl, owner and occasional inhabitant of Mainau, the sentence is sure to be about the island Mainau. We can assign one choice function Φ_c to each of these situations, representing the salience structure of that particular context c.

(7) *the island*
(8) $\|\varepsilon x \ [\text{island}]\|^{M,g,\Phi^{\text{fisherman}}} = \text{Reichenau}$
(9) $\|\varepsilon x \ [\text{island}]\|^{M,g,\Phi^{\text{tourist-guide}}} = \text{Lindau}$
(10) $\|\varepsilon x \ [\text{island}]\|^{M,g,\Phi^{\text{earl}}} = \text{Mainau}$

Thus we can conclude that the most appropriate representation for a definite NP *the F* is the epsilon term *εx Fx*, which denotes that individual with the property *F* that is selected in a situation *c*, as in (11):

(11) the F: $\|\varepsilon x \ Fx\|^{M,g,\Phi_c} = \Phi_c(\|F\|^{M,g,\Phi_c})$

To summarize: We represent definite NPs by epsilon terms which are interpreted by a global choice function representing the salience structure of the discourse. Thus we subsume the anaphoric use under the situational or salience use of definite NPs. Uniqueness is understood as "unique availability" of the referent rather than as a requirement that the corresponding descriptive material have a singleton set.

3. Indefinite NPs and Local Choice Functions

Recent work in linguistic semantics has explored the analysis of indefinites as terms rather than quantifier phrases as a response to concerns about constraining scope construal and systematic ambiguity of type. In this section I will add two further observations about the semantics of indefinites: First, the observation that indefinite NPs often behave like terms, and second the observation that indefinites and definites behave very similar with respect to their context change potential.

3.1. CHOICE FUNCTIONS AND MOVEMENT

One of the most celebrated arguments for using choice functions rather than existential quantifiers is based on a conflict between three principles of LF-representation: (i) scope ambiguities are reconstructed by movement, (ii) indefinite NPs are represented as existential quantifier phrases, and (iii) there are scope islands such as that-clauses. Fodor and Sag (1982) observe that (specific) indefinite NPs do not obey scope islands, as illustrated in (12b). Sentence (12) can receive a reading where the indefinite NP *a student* receives

wide scope over *the rumor* while the universal term *each student* in (13) cannot since the that-clause constitutes a scope island for quantifier phrases.

(12)   John overheard the rumor that a student of mine had been called before the dean.
(12a) the rumor > there is a student
(12b) a certain student > the rumor
(13)   John overheard the rumor that each student of mine had been called before the dean.
(13a) the rumor > each student
(13b) *each student > the rumor

One way to account for this "exceptional" scope behaviour is to assume that the indefinite is interpreted by a choice function. This can be illustrated on the example (14), where the indefinite NP *a girl* is represented by *f(girl)* with f being a choice function: *ch(f)*(cf. Reinhart, 1992; Winter, 1997; Kratzer, 1998; von Stechow, 2000; among others).

(14)   Every boy dates a girl.
(14a) $\forall x \ [boy(x) \rightarrow \exists \ f \ [ch(f) \ \& \ date \ (x, f(girl))]]$
(14b) $\exists f \ [ch(f) \ \& \ \forall x \ [boy(x) \rightarrow date \ (x, f(girl))]]$

The two readings of the example (14) are represented both with the indefinite *in situ,* while the existential binder of the choice function variable *f* appears at different locations resulting in the narrow scope reading (14a) and the wide scope reading (14b).[5]

3.2. INDEXED EPSILON TERMS

Egli (1991), Egli and von Heusinger (1995), Meyer-Viol (1995), von Heusinger (2000) argue that indefinite NPs often behave like terms. Therefore, they represent indefinite NPs as epsilon terms, rather than as quantifier phrases. Epsilon terms reconstruct the assumption that the indefinite article "picks up" one element of the set which is formed by all elements that fit the description in the NP. Thus an indefinite NP is of type *e*, rather than of the quantifier type ((et)t). This assumption about "selecting one element" is reconstructed by the interpretation of the epsilon operator by a choice function, which takes a set and yields an element of this set, as we have seen above. Other than definite NPs, each indefinite NP introduces a new choice function. This is represented in this framework by indexed epsilon terms which are interpreted by different choice functions, which are only defined for the descriptive content of the indefinite NP by which they are introduced. In contrast, all definite NPs are interpreted according to the one global choice function which stands for the salience structure of the context. This can be formalized as in (11), repeated as (15), and in (16):

(15) the F: $\|\varepsilon x \; Fx\|^{M,g,\Phi_c} = \Phi_c(\|F\|^{M,g,\Phi_c})$
(16) an F: $\|\varepsilon_i x \; Fx\|^{M,g,\Phi_c} =$ there is an Φ_i such that $\Phi_i(\|F\|^{M,g,\Phi_c})$

Here it also becomes obvious that the choice function for indefinites need not be defined for all sets, but only for the set that is associated with the descriptive material of the indefinite, thus we speak of "minimalized" or "local" choice functions.

To sum up: definite and indefinite NPs can be both represented by epsilon terms, which are interpreted by choice functions. They differ in that definite NPs are interpreted by the global choice function representing the salience structure, while indefinite NPs are interpreted by local and "minimalized" choice functions. Thus we use choice functions in two ways in representing definite and indefinite NPs.

3.3. CONTEXT CHANGE POTENTIAL OF DEFINITE AND INDEFINITE NPs

Definite and indefinite NPs exhibit another interesting common property: Their context change potential is identical, as it can be illustrated at the following examples. The anaphoric definite NP *the donkey* in the second sentence of (17) and (18) refers to its antecedent *a donkey* and *the donkey* in the same way.

(17) John owns *a donkey*. He beats *the donkey*.
(18) John owns *the donkey*. He beats *the donkey*.

The indefinite NP *a donkey* in (17) updates the given salience structure in such a way that the subsequent *the donkey* refers to the same object thus establishing the anaphoric reference by coreference. One could say that the indefinite updates the salience structure while the definite does not (since it already refers to the most salient donkey). Both definite NPs *the donkey* in (18) refer just to the same individual due to the uniqueness condition. However, in the presented theory, both definite NPs refer to the same individual due to the salience structure and establish coreference due to the same contextual parameters.

The structure of the contextual salience allows to account for coreference in examples (19) and (20) and illustrate an additional observation. Intuitively, in (19) the indefinite NP *a donkey* refers to a particular donkey that not only becomes the most salient donkey but also the most salient animal. This means that the indefinite not only updates the global choice function for the set that corresponds to the descriptive material by which it is introduced but also for some supersets of it. This also holds of definite NPs: *the donkey* in (20) not only (trivially) updates the global choice function for the set of donkey but also for some supersets (such as animals).

(19) John owns *a donkey*. He beats *the animal*.

(20) John owns *the donkey*. He beats *the animal*.

Therefore, I assume that both definite and indefinite NPs have the same context change potential (even though in simple cases it is invisible for definite NPs).

4. Dynamic Semantics with Choice Functions

Dynamic semantic theories like Discourse Representation Theory (DRT) (Kamp, 1981; Heim, 1982) or Dynamic Predicate Logic (DPL) (Groenendijk and Stokhof, 1991) take the following stand on the context dependent nature of interpretation. They assume that the meaning of a sentence is identified with its context change potential. Contexts are taken to be information states. Meanings are updates of such information states and interpretation of sentences creates context. Information states contain two kinds of information: information about the world, and discourse information. The information about the world is relevant for the truth conditions, while the information about the discourse restricts anaphoric relations. However, dynamic semantic theories only provide sets of accessible discourse referents, but no further ranking among them.

In a dynamic semantics with choice functions, the information states are sets of choice functions, rather than sets of assignment functions as in ordinary dynamic semantics. The discourse meaning of linguistic expressions (not only sentences) updates this information, which means that it potentially restricts the set of (possible) choice functions, which stand for the (possible) discourse structures. Here we model the discourse information of information states at an additional dynamic mechanism, which will developed below.[6]

4.1. THE PROBLEM OF COINDEXING

Dynamic approaches like Discourse Representation Theory (DRT) or Dynamic Predicate Logic (DPL) primarily investigate cross-sentential anaphoric pronouns. There is one problem of these approaches, which can be illustrated with our initial example (1): the pronoun *he* in line (viii) has two potential antecedents or already established discourse referents: the discourse referent for the cat Bruce and the discourse referent for the New Zealand cat Bobby. DRT cannot tell which is the more appropriate one, but must rely on additional knowledge, which is indicated by co-indexing the anaphoric term with its antecedent. However, it is the anaphoric relation that the theory should explain and not rely on.

(1) "the cat"
  i Imagine yourself with me as I write these words. In the room is **a cat**$_1$, Bruce$_1$,
  ii who has been making **himself**$_1$ very salient by dashing madly about.

iii **He**$_1$ is the only cat in the room, or in sight, or in earshot.
iv I start to speak to you:
v The **cat**$_1$ is in the carton. **The cat**$_1$ will never meet **our other cat**$_2$,
vi because **our other cat**$_2$ lives in New Zealand.
vii **Our New Zealand cat**$_2$ lives with the Cresswells.
viii And there **he**$_2$'ll stay, because Miriam would be sad if **the cat**$_2$ went away.

It seems very obvious from the discourse structure that the pronoun *he* can only refer to the New Zealand cat Bobby and therefore must be linked to that discourse referent. Therefore, I assume that the anaphoric link should follow from the theory and not be part of the input. This restriction of dynamic theories like DRT and DPL is described by Muskens and et al (1997, p. 606):

> Discourse Representation Theory models the way in which anaphoric elements can pick up accessible discourse referents, it tells us which referents are accessible at any given point of discourse, but it tells us little about the question which referent must be chosen if more than one of them is accessible. There are of course obvious linguistic clues that restrict the range of suitable antecedents for any given anaphoric element (...).

In the following we will concentrate on the information that is supplied by the descriptive material of definite and indefinite NPs, which updates the salience structure of the discourse. And the salience structure crucially contributes to the interpretation of definite anaphoric expressions. Peregrin and von Heusinger (1995, 2004) developed a dynamic semantics with choice functions in order to modell this linguistic information that is relevant for resolving the anaphoric reference. The context change potential of an expression is seen in its potential to update the global choice functions in the sense of section 2 consistent with discourse. The dynamic semantics with choice function is an extension of classical DPL: the dynamism of the salience structure is modeled in parallel to the information states that encodes the increasing information of the discourse.[7] In the following we concentrate on the context change potential. The context change potential of definite and indefinite expressions updates the context by changing the (global) choice function Φ_c. An indefinite NP introduces a discourse referent, which then becomes then the most salient of its kind. Hence, the global choice function is updated with respect to the set described by the indefinite. This set is assigned the referent of the indefinite. A definite NP refers to the most salient of its kind. The context change potential is generalized and therefore represented as a relation between two (potential) global choice functions.

4.2. CHOICE FUNCTIONS AND DYNAMIC INTERPRETATION

In the remainder of this section I present the dynamic choice functions approach of Peregrin and von Heusinger (1995, 2004): let us assume the non-empty universe U of individuals. A choice function (or "epsilon function" in Peregrin and von Heusinger) f is a partial function from the power-set of U into U such that $f(s) \in s$ for every $s \subseteq U$ for which f is defined. This means that the class CH_U of all choice functions based on U is defined as follows (where D(f) and R(f) denote the domain and the range of f, respectively):

$$DEF1 \ CH_U = \{f | D(f) \subseteq Pow(U) \ and \ R(f) \subseteq U \ and \ f(s) \in s$$
$$for \ every \ s \in D(f)\}$$

We further introduce update functions for choice functions, or choice function (cf-) updates in short. A cf-update is an operation that takes three arguments: a choice function, an element of the universe, and a subset of the universe; it yields a new choice function. The basic cf-update upd_1 applied to an choice function f, an individual d, and a set s, yields the choice function f' which is identical with f except for the assignment d for the set s.

DEF2 upd_1 is defined as follows

$$upd_1(f,d,s) = f' \ such \ that \quad f'(s') = d \ if \ s' = s \ and \ d \in s$$
$$and \quad f'(s') = f(s') \ otherwise$$

We use $f' \approx f^s$ as an abbreviation for $\exists d.f' = upd_1(f,d,s)$.[8] If $f_2 \approx f_1^s$ and $f_3 \approx f_2^{s'}$, then we also write $f_3 \approx f_1^{s,s'}$. Thus upd_1 can be seen as the first approximation to the salience change potential of an indefinite NP. The indefinite NP *a man* selects an arbitrary man and changes the actual choice function such that this arbitrarily chosen man becomes the current representative for the class of men. In the following, a formal fragment will be defined illustrating how choice functions act in a dynamic semantics. We do without quantifiers, since they play no role in the argument. However, for a detailed treatment of quantifiers in this framework see Peregrin and von Heusinger (1995, 2004).

DEF3a. (lexicon)
1. sentences
2. terms (**he, she, it**)
3. n-ary predicates for n > 0 (constants **man, walk, whistles, farmer, boring, woman, thing** for n=1; **own, beat** for n=2)
4. determiners (constants **a, the**)
5. n-ary logical operators for n=1, 2 (the constant ¬ for n=1; **&, v** for n=2)

In the syntax DEF3b we first define the operation of forming a term $D(P)$ with a determiner D (i.e. with the definite or indefinite article) and an open sentence P. Here we differ from other dynamic approaches, which interpret definite and indefinite NPs as quantifier phrases. The other clauses are as in other dynamic frameworks, they determine the construction of an atomic sentence in (2) and of complex sentences in (3) and (4):

DEF3b. (syntax)
1. If P is a unary predicate and D a determiner, then D(P) is a term.
2. if $T_1, \ldots, T_n$ are terms and R an n-ary predicate, then $R(T_1, \ldots, T_n)$ is a sentence.
3. If S is a sentence and o a unary logical operator, then oS is a sentence.
4. If S_1 and S_2 are sentences and o a binary logical operator, then $S_1 o S_2$ is a sentence.

DEF3c recalls the static interpretation of terms we have used in section 3. A term is interpreted in a model that consits of a domain U and an interpretation functions I and according to a choice function f. The interpretation of a constant term (such as proper names) does not depend on the choice function f, while the interpretation of a complex term of the type $D(P)$ crucially depends on the given choice function, as in 1b. Here *he, she, it, the* P and *a* P play the role of variables: they do not have a lexical meaning but only a meaning relative to the contextual choice function. The rule 2 for predicate constants is as usual.

DEF3c. (static semantics)
A model is a pair $\langle U, I| \rangle$, where U is a non-empty set and I is a function such that

1a. $\|T\|^{M,f} = I(T)$ if T is a constant term
1b. $\|T\|^{M,f} = f(\|P\|)$ if T is D(P) for a determiner D and a predicate P
2. $\|R\|^{M,f} = I(R) \subseteq U^n$ if R is an n-ary predicate

This static semantics determines the interpretation of terms, but does not show the update function of linguistic expressions. In order to model the context change potential of linguistic expression we have to assume that meaning is an update of information states. As noticed above, information states contain knowledge about the world and information about the discourse. In the following we are only concerned with information about the discourse, in particular with the salience structure of the discourse. We assume that the (discourse) meaning of linguistic expressions is their update function of choice functions (which stand for the salience structure). Information states can be modeled as sets of choice functions that are potentially changed by the salience change potential of the linguistic expression. We therefore define the dynamic semantics in the following way: The function $\| \ \|$

is extended to the categories of terms and sentences so that if E is a term or a sentence, then $\|E\| \subseteq CHxCH$:

DEF3d. (dynamic semantics)
1a. $\|\mathbf{a}(P)\| = \{\langle f, f' \rangle | f' = f^{\|P\|}\}$
1b. $\|\mathbf{the}(P)\| = \{\langle f, f' \rangle | f' = f \text{ and } f'(\|P\|) \text{ is defined}\}$
1c. $\|\mathbf{he}\| = \|\mathbf{the}(\mathbf{man})\|$
1d. $\|\mathbf{she}\| = \|\mathbf{the}(\mathbf{woman})\|$
1e. $\|\mathbf{it}\| = \|\mathbf{the}(\mathbf{thing})\|$
2. $\|P(T_1, \ldots, T_n)\| = \{\langle f, f' \rangle| \text{ there exist } f_0, \ldots, f_n \text{ so that } f = f_0 \text{ and } f' = f_n$
     and $\langle f_0, f_1 \rangle \in \|T_1\|$ and $\ldots$ and $\langle f_{n-1}, f_n \rangle \in \|T_n\|$ and
     $\langle \|T_1\|_{f_1}, \ldots, \|T_n\|_{f_n} \rangle \in \|P\|\}$
3. $\| \neg S \| = \{\langle f, f' \rangle| f = f' \text{ and there is no } f'' \text{ such that } \langle f, f'' \rangle \in \|S\|\}$
4a. $\|S_1 \& S_2\| = \{\langle f, f' \rangle| \text{ there is an } f'' \text{ such that } \langle f, f'' \rangle \in \|S_1\| \text{ and}$
     $\langle f'', f' \rangle \in \|S_2\|\} (= \|S_1; S_2\|)$
4b. $\|S_1 v S_2\| = \{\langle f, f' \rangle| f = f' \text{ and there is an } f'' \text{ such that } \langle f, f'' \rangle \in \|S_1\| \text{ or}$
     $\langle f, f'' \rangle \in \|S_2\|\}$

The (discourse) meaning of an indefinite NP *a P* is an choice function update, i.e. it updates the input choice function f to the output choice function f′ (here "the input choice function" means "whatever choice function from the set of choice function is taken"). f′ differs from *f* at most in the assignment for the set of P, which is the denotation of the indefinite NP. We write $f^{\|P\|}$ for an f′ resulting from the evaluation of *a P* with the input f. Clause (1b) describes the salience change potential of definite NPs: A definite NP *the P* denotes the representative of the set of P's according to a choice function f; it is taken to express the trivial cf-update. Further, it is required that there be at least one P – this expresses the existential presupposition of definite NPs. There is no uniqueness condition, since it is replaced by the condition that there exists the representative of the set of P's. The pronouns in (1c-1e) are taken to be semantically equivalent to the impoverished definite NP expressing merely the corresponding gender.

   The atomic sentence is semantically characterized in (2) via its potential to change the input choice function f to the updated function f′ by way of the subsequent application of the updates expressed by its terms. Thus, f and f′ must be connected by a sequence of choice functions such that the adjacent pairs of the sequence fall into the respective updates expressed by the terms; and the referents of the terms must fall into the extension of the predicate. Here we differ essentially from usual dynamic logic in that we consider atomic sentences internally and externally dynamic.[9] The logical operators ¬ and v are static (they act as tests) – they are in fact the classical operators only formally dynamized. & is the dynamic conjunction suitable for conjoining subsequent sentences.

Let us illustrate this mechanism by analyzing a simple atomic sentence with an indefinite NP. Sentence (21) is assigned the formula (21a) which is then interpreted as (21b) according to the definitions given above. As we have noted, a pair of choice functions $\langle f, f' \rangle$ falls into the update expressed by an atomic sentence iff f and f$'$ are connected by a sequence of choice functions such that the adjacent pairs of the sequence fall into the respective updates expressed by the terms and the referents of the terms fall into the extension of the predicate. Since we have only one term in (21), it is reduced to the condition that $\langle f, f' \rangle$ falls into the update expressed by *a(man)* and that the referent of *a(man)* falls into the extension of walk. This yields $f' = f^{\|man\|}$ and $f'(\|man\|) \in \|walk\|$. The resulting set of pairs is clearly non-empty just in case $\exists d. d \in \|man\|$ & $d \in \|walk\|$ (i.e. if the intersection of $\|man\|$ and $\|walk\|$ is non-empty) and our formula (21b) is thus in this sense equivalent to the classical formula $\exists x(man(x)\&walk(x))$.

(21)   A man walks
(21a) walk(a(man))
(21b) $\|walk(a(man))\|$
       $= \{\langle f, f' \rangle | \langle f, f' \rangle \in \|a(man)\|$ and $\|a(man)\|_{f'} \in \|walk\|\}$
       $= \{\langle f, f' \rangle | f' = f\|man\|$ and $f'(\|man\|) \in \|walk\|\}$

Sentence (22) with the definite NP *the man* is represented and interpreted similarly to (21). The only difference is the condition on the choice function – the interpretation of the definite NP is static (in the formalism developed so far). The only condition is that the referent of the NP, determined by the current choice function, falls within the extension of the predicate. The difference between the definite and the indefinite NP thus lies in their different behaviors with respect to the choice function — the indefinite NP updates it, whereas the definite NP acts merely as a test.[10] In both cases, the referent of the NP is yielded by the input choice function

(22)   The man whistles
(22a) whistle(the(man))
(22b) $\|whistle(the(man))\|$
       $= \{\langle f, f' \rangle | \langle f, f' \rangle \in \|\rangle(man)\|$ and $\|the (man)\|f' \in \|whistle\|\}$
       $= \{\langle f, f \rangle | f' = f'\|man\|$ & $f'(\|man\|) \in \|whistle\|\}$

The analysis of the conjunction (23) of (21) and (22) shows how the referent of the anaphoric NP the man gets identified with that of its antecedent a man.

(23)   A man walks. And the man whistles
(23a) walk(a(man))&whistle(the(man))
(23b) $\|walk(a(man))\&whistle(the(man))\|$
       $= \{\langle f, f' \rangle |$ there is an f'' such that $\langle f, f'' \rangle \in \|walk(a(man))\|$ and
       $\langle f'', f' \rangle \in \|whistle(the(man))\|\}$.

$$= \{\langle f,f'\rangle | \text{ there is an } f'' \text{ such that } \langle f,f''\rangle \in \{\langle f, f'\rangle | f' = f\|man\| \text{ and } f'(\|man\|) \in \|walk\|\} \text{ and } \langle f'',f'\rangle \in \{\langle f, f'\rangle | f = f' \text{ and } f'(\|man\|) \in \|whistle\|\}\}$$

$$= \{\langle f,f'\rangle | \text{ there is an } f'' \text{ such that } f'' = f\|man\| \text{ and } f'(\|man\|) \in \|walk\| \text{ and } f'' = f' \text{ and } f'(\|man\|) \in \|whistle\|\}$$

$$= \{\langle f,f'\rangle | \ f' = f\|man\| \text{ and } f'(\|man\|) \in \|walk\| \text{ and } f'(\|man\|) \in \|whistle\|\}$$

$\langle f,f'\rangle$ falls into the update expressed by (23b) if and only if there is an choice function f'' such that $\langle f,f''\rangle$ falls into the update expressed by (21b) and $\langle f'',f'\rangle$ falls into the update expressed by (22b). Using the results of the above analyses and eliminating redundancies, we reach the result that $\langle f,f'\rangle$ falls into the update expressed by (23b) iff f' differs from f at most in the representative of the class of men and this representative is a walker and a whistler.

5. The dynamics of definite descriptions

5.1. PROJECTING THE SALIENCE CHANGE

Using this formalism, we can give a first analysis of the variety of anaphoric relations in (1) above. The meaning of the first sentence in (1i) – (1v) consists of the pairs of choice functions f and f' such that f' is like f with the single possible exception that f' chooses a new representative for the class of cats, namely Bruce. Furthermore, the chosen representative must be in the extension of the predicate be in the room. The definite expression the cat in (1v) then refers to this chosen individual, namely Bruce. Thus, the anaphoric relation is not explained in terms of binding or by means of a Russellian description, but rather in the interaction of the context change potential of the antecedent together with the context dependent interpretation of the anaphoric term. However, this basic picture can only account for the anaphoric link between *a cat* and *the cat*, but not for the anaphoric link between *a cat* and *he* in (1iii) or the anaphoric link between *our New Zealand cat* in (vii) and *the cat* in (viii):

In order to account for an anaphoric relation between the indefinite NP *a cat* and the pronoun *he*, Peregrin and von Heusinger (1995, 2004) modify definition DEF 2 to DEF 2′. An indefinite NP *an F* does not only change the representative of the class of Fs, but also the representative of (certain) supersets. Hence, the anaphoric expressions *he* (as short for *the (male) object*) refers back to the mentioned representative.

DEF2 upd_1 is defined as follows

$$upd_1(f,d,s) = f' \text{ such that } \quad f'(s') = d \text{ if } s' \subseteq s \text{ and } d \in s$$
$$\text{and} \quad f'(s') = f(s') \text{ otherwise}$$

This modificaton can not only handle the coreference between a pronoun and its antecedent, but also between an anaphoric definite NP and its antecedent if the antecendent is more specific, such as in (24) and (25):

(24) John is looking at a collie. The dog barks.
(25) John is looking at a small white dog. The dog barks.

Still, this modification does not explain the anaphoric link between the definite NP *our New Zealand cat* and the definite NP *the cat*. An even more flexible account of salience change potential is necessary.

5.2. THE DYNAMICS OF DEFINITE DESCRIPTIONS

Lewis' example "the cat" illustrates that different occurrences of the definite NP *the cat* can refer to different referents, contrary to the classical assumption of Russell, which are also held in contemporary theories. The example rather shows that the context can change in a way that the second occurrence of the cat refers to a different object. The question I address in this subsection is what the contribution of other definite NPs to this context change is – we have already seen that an indefinite changes or updates an input choice function. Before we analyze the context change potential of a definite NP, I present another text in which we have more than one occurrence of one definite NP with different referents. The fragment is from the short novel "A clean, well-lighted place" from Ernest Hemingway ([1925] 1966, p. 379):

(26) **A clean, well-lighted place**
It was late and everyone had left the café except an old man who sat in the shadow the leaves of the tree made against the electric light. [...] **The two waiters** inside the café knew that the old man was a little drunk [...]. "Last week he tried to commit suicide," **one waiter** said. "Why? '[...] **The younger waiter** went over to him. [...] The old man looked at **him**. *The waiter* went away. [...]
The waiter who was in hurry came over. "Finished," **he** said [...]. "Another", said the old man. "No, finished." *The waiter* wiped the edge of the table with a towel and shook **his** head. The old man stood up [...]. "Why didn't you let him stay and drink?" **the unhurried waiter** asked.

In this fragment the two occurrences of the definite NP *the waiter* refer to different waiters – both refer to the last mentioned one. We can then extract the following anaphoric chains from this example:

(27) Anaphoric chains of definite NPs in (26)
the younger waiter ... him ... the waiter
the waiter who was in hurry.... he ... the waiter

It is obvious that the definite NP *the younger waiter* changes the context in a way that its referent is not only the most salient younger waiter (trivially), but also that its referent is the most salient waiter (at all). In order to implement this, we change the interpretation rule for the definite NP (1b) to (1b') by accommodating it to the one of indefinite NPs (1a) – the original rules are repeated below:

1a. $\|\mathbf{a}(P)\| = \{\langle f, f'\rangle \mid f' = f^{\|P\|}\}$
1b. $\|\mathbf{the}(P)\| = \{\langle f, f'\rangle \mid f' = f \text{ and } f'(\|P\|) \text{ is defined}\}$
1b'. $\|\mathbf{the}(P)\| = \{\langle f, f'\rangle \mid f' = f^{\|P\|}\}$

In the original rule (1b), the definite character of the NP was warranted by its static (i.e. non-updating) behavior, while in the new interpretation rule the definite NPs is also assigned an update function. However, this update function can only change the context if we allow for the more flexible updating function DEF2', otherwise it would trivially update the given global choice function only for the set associated with the descriptive material. This rule makes the definite and indefinite article synonymous with respect to their discourse meaning, i.e. to the salience change potential. This matches the intuition that both definite and indefinite NPs change the salience structure.

   The difference between an indefinite and a definite NP is not the dynamic vs. static behavior, but the way they find their referents. An indefinite NP refers to an arbitrarily selected element (by way of an existential quantifier or by a local choice functions), while a definite NP refers to its referents due to the global choice function (standing for the salience structure of a discourse). With these two modifications of the original dynamic semantics with choice functions, we can account for the example (1) of Lewis, the fragment from Hemingway (26) and many more natural language discourses with more than one occurrence of one and the same definite NP.

6. Summary

It was shown that definite descriptions exhibit two functions: (i) they are interpreted depending on the context and establish in this way anaphoric links by coreference, (ii) they change the context by raising new referents to the most salient ones for the set they describe as well as some supersets. Furthermore, it was argued that the most appropriate representation for definite descriptions are context dependent global choice function terms. These terms refer to the most salient object of the class of objects that fall under the descriptions and the referent of the term becomes the most salient element of the set, as well as some supersets. Indefinites are represented by local choice functions. Both NPs change the context by updating the global choice function, which represents the salience structure of the discourse.

Notes

[1] In the remainder of this paper I discuss choice functions that take sets as their arguments, rather then predicates. This approach has to be modified to choice functions that take predicates in order to get a more adequate picture.

[2] Hilbert's original proposal assumed that the selection function was total and assigned an arbitrary object to the empty set. In this paper, I assume that the choice function which represents the global salience parameter is a partial function, excluding the empty set from its domain. For the purposes of modeling salience, partial choice functions are more intuitive. I will continue to use the epsilon operator to highlight the parallel treatment of indefinite and definite descriptions in our analysis.

[3] Partial choice functions will do for the reconstruction of the semantic behavior of definite and indefinite NPs. If we would use total choice functions we have to take care of the "empty set problem" (see Winter, 1997). There are several possible solution for this problem (see Asser, 1957).

[4] Slater (1986) uses the classical epsilon calculus in order to describe E-type pronouns. He substitutes the iota-operator by the epsilon-operator and can therefore avoid the problematic uniqueness condition of the iota-term. Meyer-Viol (1995) develops such full epsilon calculus and applies it to semantic problems such as E-type and Bach-Peter pronouns. A variant of Meyer-Viol's calculus is used in Dynamic Syntax (e.g. Kempson et al. 2000; Kempson & Meyer-Viol, 2003).

[5] Kratzer argues that the choice function variables remain free at LF and that the context specifies their values. This makes the choice functions for indefinites similar to the global choice function we have employed for definite NPs. This reflects Kratzer's view that choice functions represent specific indefinites, i.e. indefinites that have some characteristics of definite expressions such as their wide scope. However, Kratzer's choice functions still differ considerably from the global choice functions for definite NP. They are only defined for the set by which they are introduced ("minimalized" choice functions), each indefinite introduces a new choice function variable, and their referent must not be mentioned in the discourse before (novelty condition).

[6] The original idea was developed in Egli and von Heusinger (1995) and formalized in Peregrin and von Heusinger (1995, 2003). Groenendijk et al. (1997) discuss this idea in detail and compare it with their own version of DPL.

[7] We need two levels of dynamic procedure: The one that keeps track of the (denotational) information in the discourse, and the other that models the salience structure. While the denotational component is montone increasing, the structural component is not. See also for a similar distinction von Heusinger (1997, chapt. 8).

[8] Here it is assumed that the referent for the indefinite is found by interpreting the indefinite as an existential operator (at least at the meta-language). This corresponds to the classical approach to indefinites. According to the choice function approach of section 3, we can modify this by saying that the referent is found by applying a local choice function to the set described by the descriptive material of the NP. For the argument in this section this does not make any difference, since we do not deal with scope interactions. Thus, in the remainder of the section

we use the classical view. However, for a more comprehensive account one would like to follow the indexed epsilon approach.

[9] Since each of the arguments of an atomic sentence potentially updates the given choice function, we have to account for different interpretations of one sentence depending on the order of its arguments.

[10] At this stage a context change potential of a definite NP would trivially update the given choice function – it would make the most salient referent most salient. See, however, section 5 for a revision of this position.

References

Asser G. (1957) Theorie der logischen Auswahlfunktionen. *Zeitschrift für mathematische Logik und Grundlagen der Mathematik*, 3, pp. 30–68.

Egli U. (1991). (In)definite Nominalphrase und Typentheorie. In: Egli U., K. von Heusinger (eds.), *Zwei Aufsätze zur definiten Kennzeichnung*. Arbeitspapier 27. Fachgruppe Sprachwissenschaft Universität Konstanz.

Egli U., von Heusinger K. (1995) The Epsilon Operator and E-Type Pronouns. In: U. Egli et al. (eds.), *Lexical Knowledge in the Organization of Language*. Benjamins, Amsterdam, pp. 121–141.

Fodor J., Sag I. (1982) Referential and quantificational Indefinites. *Linguistics and Philosophy*, 5, pp. 355–398.

Groenendijk J., Stokhof M. (1991) Dynamic Predicate Logic. *Linguistics and Philosophy*, 14, pp. 39–100.

Groenendijk J., Stokhof M., Veltman F. (1997) Coreference and Contextually Restricted Quantification. Is There Another Choice. In: Kamp H., B. Partee (eds.), Proceedings of the Workshop "Context Dependence in the Analysis of Linguistic Meaning". Prague and Stuttgart 1995, pp. 217–240.

Heim I. (1982) The Semantics of Definite and Indefinite Noun Phrases, Ph.D. Dissertation. University of Massachusetts, Amherst. University Microfilms. Ann Arbor.

Hemingway E. ([1925] 1966) A clean, well-ligthed place. *The short stories of Ernest Hemingway*. Scribner, New York, pp. 379–383.

von Heusinger K. (1997) Salienz und Referenz. Der Epsilonoperator in der Semantik der Nominalphrase und anaphorischer Pronomen. Studia grammatica 43. Akademie Verlag, Berlin.

von Heusinger K. (2000) The Reference of Indefinites. In: K. von Heusinger & U. Egli (eds.), *Reference and Anaphoric Relations*. Kluwer, Dordrecht, 247–265.

Hilbert D., Bernays P. [1939] (1970) *Grundlagen der Mathematik*. Vol. II. 2. ed. Berlin; Heidelberg; New York.

Kamp H. (1981) A Theory of Truth and Semantic Interpretation. In: Groenendijk J., Janssen T., Stokhof M. (eds.). Formal Methods in the Study of Language. Amsterdam Center, Amsterdam. pp. 277–322.

Kempson R., Gabbay D., Meyer-Viol W. (2000) *Dynamic Syntax. The Flow of Language Understanding*. Blackwell, London.

Kempson R., Meyer-Viol W. (2003) *Dynamic Syntax. The Flow of Language Understanding*. In: Peregrin J. (ed.), *Meaning in the Dynamic Turn*. Elsevier, Amsterdam.

Kratzer A. (1998) Scope or Pseudoscope? Are there Wide-Scope Indefinites? In: S. Rothstein (ed.), *Events and Grammar*. Dordrecht, Kluwer, pp. 163–196.

Lewis D. (1979) General Semantics. *Synthese*, 22, pp. 18–67.

Lewis D. (1979) Scorekeeping in a Language Game. In: Bäuerle R., Egli U., von Stechow A. (eds.), *Semantics from Different Points of View*. Springer, Berlin; Heidelberg, New York, pp. 172–187.

Meyer-Viol W. (1995) Instantial Logic. An Investigation into Reasoning with Instances. ILLC Dissertation Series 1995-11, Amsterdam.

Muskens R., van Benthem J., Visser A. (1997) Dynamics. In: van Benthem J., ter Meulen A. (eds.), Handbook of Logic and Language. Amsterdam, pp. 587–648.

Peregrin J. (2000) Reference and Inference: The Case of Anaphora. In: von Heusinger K., Egli U. (eds.), *Reference and Anaphoric Relations*. Kluwer, Dordrecht, pp. 269–286.

Peregrin J., von Heusinger K. (1995) Dynamic Semantics with Choice Functions. In: Egli U., von Heusinger K. (eds.). *Choice Functions in Natural Language Semantics*. Arbeitspapier 71. Fachgruppe Sprachwissenschaft Universität Konstanz, pp. 41–67.

Peregrin J., von Heusinger K. (2004). Dynamic Semantics with Choice Functions. In: Kamp H., Partee B. (eds.), Context Dependence in the Analysis of Linguistic Meaning. Elsevier Science, Amsterdam.

Poesio M., Stevenson R. (to appear). *Salience: Computational Models and Psychological Evidence*. Cambridge: University Press, Cambridge.

Reinhart T. (1992) Wh-in-situ: An Apparent Paradox. In: Dekker P., Stokhof M. (eds.), *Proceedings of the Eighth Amsterdam Colloquium*. ILLC. University of Amsterdam, pp. 483–491.

Russell B. (1905) On Denoting. *Mind*, 14, pp. 479–493.

Schröter K. (1956) Theorie des bestimmten Artikels. *Zeitschrift für Mathematische Logik und Grundlagen der Mathematik*, 2, pp. 37–56.

Sgall P. (1984) Remarks on Text, Language and Communication. In: Sgall P. (ed.), *Contributions to Functional Syntax, Semantics, and Language Comprehension*. Benjamins, Amsterdam, Philadelphia, 123–144 (Linguistic & Literary Studies in Eastern Europe 16).

Slater B. H. (1986) E-Type Pronouns and Epsilon-Terms. *Canadian Journal of Philosophy*, 16, 27–38.

von Stechow A. (2000) Some Remarks on Choice Functions and LF-Movements. In: von Heusinger K., Egli U. (eds.), *Reference and Anaphoric Relations*. Kluwer, Dordrecht, pp. 193–228.

Winter Y. (1997). Choice Functions and the Scopal Semantics of Indefinites. *Linguistics and Philosophy* 20, pp. 399–467.

Functional Quantification

YOAD WINTER

Computer Science, Technion – Israel Institute of Technology, Haifa 32000, Israel E-mail: winter@cs.technion.ac.il

Abstract. This paper develops a unified analysis of "functional" anaphora arid wide-scope indefinites. A new operator is added to Jacobson's variable-free semantics of functional readings, which leads to an analysis of theses readings using the general *Skolem function* interpretation of wide-scope indefinites. This accounts for the distriutional, technical and intuitive similarities between the two phenomena. Moreover, after formally characterizing the class of generalized quantifiers that are treated by the proposed mechanism, it is argued that this class is a good approximation of the quantifiers that empirically support functional readings.

Key words: functional readings, generalized quantifiers, indefinites, scope

1. Introduction

Traditional formal semantics assumes that quantification in natural language operates only on atomic entities. However, much recent work in natural language semantics has shown the advantages of more complex forms of quantification that involve functions over atomic domains. Two areas of functional quantification have received special attention. One area concerns the so-called *functional* and *pair-list* interpretations of questions and copular sentences. Another area deals with the *wide scope* interpretation of indefinite NPs. Quantification over *Skolem functions* is used to model both kinds of phenomena. Some theories restrict the usage of Skolem functions to the simple case of *choice functions:* Skolem functions that map any non-empty set to an entity in this set.

Despite the similarity between the mechanisms that are used for describing functional readings and wide scope indefinites, no attempt has so far been made to bring them into one framework. This paper proposes such a unified mechanism. It argues that functional readings and the interpretation of wide scope indefinites are restricted by the same principles – those that govern quantification over Skolem functions. Under this view, quantification over functions in natural language is existential only, and is furthermore restricted to a small subclass of noun phrases. This accounts for the similar distribution of functional quantification and wide scope phenomena, as well as for some

formal relations between functional quantification and Generalized Quantifier Theory. A novel hypothesis characterizes the class of quantifiers that license functional interpretations and explains their restricted distribution.

In a nutshell, the two general phenomena with which this article deals can be illustrated by the following two sentences.

(1)   a. The (only) woman that every man loves is his mother.
      b. Every man loves a (certain) woman.

Sentence (1a) illustrates *a functional reading*. In this reading, the pronoun *his* is "bound" by the noun phrase *every man* although it is not within its syntactic scope. Another familiar scope problem appears in sentence (lb). This sentence has a reading where the indefinite *a (certain) woman* is interpreted as taking *wide scope* over the subject, although the subject is not within its syntactic scope.

Popular analyses of both phenomena involve Skolem functions under some version or another. To illustrate these approaches, consider first the following intuitive analyses of sentences (1a) and (1b).

(2)   a. The (only) function in the set {f: f maps every man to a woman he loves} is the function that maps every man to his mother.
      b. There is a choice function f such that every man loves $f(\{x: x$ is a woman$\})$.

We refer to nominal expressions such as *woman* and *woman that every man loves* as the *restricting predicate*, or *restriction*, of the relevant noun phrase. The functions in (2a) are in the denotation of the surface restriction *woman that every man loves*. Consequently, this analysis derives a "bound" reading of the pronoun *his* without assuming that *every man* takes the pronoun within its scope. Similarly, since the function f in (2b) applies to the restriction *woman* in its surface position, there is no standard scope of the object over the subject in this analysis.

However, there is one important difference between the two analyses. In the analysis (2a) of the functional reading, it is the restricting predicate itself (i.e. *woman that every man loves*) that is assumed to range over functions. In the functional analysis (2b) of the scope of indefinites, the function is used as a variable that only applies to the restricting predicate *woman*, but it does not belong to the restricting predicate itself. Accordiingly, existential quantification over functions in (2b) is assumed as a (possibly contextual) default mechanism, and not as the denotation of the indefinite article *a (certain)* in (lb). By contrast, the quantifier *the (only)* over functions in (2a) is assumed to be the denotation of the definite article *the (only)* in (1a).

The first aim of this paper is to bridge this gap between the two kinds of analyses. It is argued that a unified analysis is not only justified by the technical and conceptual similarity between the two theories, but also

because functional readings with copular sentences, questions and wide scope indefinites appear with the same class of NPs. The first step in unifying the two mechanisms, which is quite uncontroversial, is to generalize the choice function analysis of wide scope indefinites to Skolem functions of arbitrary arity. This is independently necessary for treating indefinites such as *a certain woman he knows,* where the pronoun *he* is bound from outside the indefinite. The second step is to modify the analysis in (Jacobson, 1994) of functional readings in variable-free semantics. The revised analysis allows the Skolem function mechanism of wide scope indefinites to apply to the restricting predicate also with functional copular sentences, while keeping the quantification over functions existential only.

The main ingredient in unifying Jacobson's account with the general treatment of indefinites is a novel type-shifting operator from objects of type $(ee)t$ (sets of functions from entities to entities) to binary relations of type $e(et)$ (functions from entities to sets of entities). Over finite domains this mapping can be one-to-one only under certain restrictions on its domain of $(ee)t$ objects. It is shown that this is naturally guaranteed when the quantifier within the functional NP (e.g. *every man* in (1a)) is a so-called *bounded* quantifier – a quantifier that can be expressed as an intersection of a positive universal quantifier *every*(A) with a negative universal quantifier *no*(B), for some sets A and B. This characterization leads to the linguistically plausible hypothesis that only bounded quantifiers can give rise to functional readings.

The structure of this paper is as follows. Section 2 briefly overviews the problems of functional readings and wide scope indefinites, gives necessary technical details about previous approaches and discusses the motivation for a unified analysis. Section 3 introduces the mapping that allows an extended theory to treat both phenomena, and exemplifies its applications. Section 4 motivates the restrictions on the proposed mapping, and proves the relations between this restriction and the class of bounded quantifiers. Section 5 points out the strong relationships between the present work and the data and mechanisms that are studied in a recent work (Jacobson, 2002).

2. Functional Readings and Wide Scope Indefinites

2.1. FUNCTIONAL READINGS

The so-called *functional* reading of questions can be illustrated by the following familiar question-answer pairs.

(3) a. Which woman does every man love? His mother.
     b. Which woman does no man love? His mother-in-law.

The problem of interpreting questions that exhibit this kind of readings was discussed extensively in the literature.[1] The puzzle is often related (cf.

Chierchia, 1993) to the problem of *pair-list* readings of questions, as illustrated by the following short discourse.

(4) Which woman does every man love? John loves Mary, Bill loves Sue, etc.

In order to analyze both kinds of reading that questions exhibit, it has been proposed that quantification over function plays a role in the interpretation. Similar mechanisms have been proposed for the functional reading of *copular sentences* like the following.

(5)   a. The woman that every man loves is his mother.
       b. The woman that no man loves is his mother-in-law.

(Sharvit, 1999) convincingly argues that such copular sentences have the same distribution and syntatic/sematic properties of functional questions as in (3). In order to illustrate the mechanisms that will be considered, we will therefore concentrate on such indicative copular sentences, without getting into the more intricate sematics of questions.

    A fully worked-out account of functional readings in copular sentences as in (5) is given in Jacobson (1994, 1999). Jacobson claims that sentences like (5a) and (5b) cannot be treated by giving the noun phrase *every man* sentential scope, as informally illustrated below.

(6)   For every (no) man x, the woman that x loves is x's mother (mother-in-law).

One reason is syntactic: to obtain such an analysis of the sentences in (5), the noun phrases *every/no man* would have to cross a complex NP (*the woman that...loves*), which is generally acknowledged to be a scope island. For instance, sentence (7a) does not have the reading that is paraphrased in (7b).

(7)   a. The woman that no man loves came to the party.
       b. For no man x, the woman that x loves came to the party.

Moreover, noun phrases normally do not bind pronouns that are not within their syntactic scope (their *c-command domain*) as witnessed by the fact that the informal analysis in (8b) is not an available reading of sentence (8a).[2]

(8)   a. The woman that no man loves pinched him.
       b. For no man x, the woman that x loves pinched x.

Another argument against a "scopal" analysis, which Jacobson attributes to Dahl (1981), is semantic. Consider sentence (9a) below. This sentence is clearly not equivalent with (9b), which is obtained by giving the noun phrase *no man* sentential scope.

(9)   a. The only woman that no man loves is his mother-in-law.
       b. For no man x, the only woman that x loves is his mother-in-law.

To see that, consider a situation in which John is a man who loves both his wife and his mother-in-law. In this situation (9a) is false. However, (9b) may still be true, as long as also other men do not love *only* their mother-in-law.

Jacobson's account of functional readings is based on her general theory of variable-free semantics. The theory itself is introduced in much detail in Jacobson (1999, 2000), and I will not try to review all of its parts here. Jacobson's assumptions that are important for our present purposes are the following, which for convenience are given the names (J1)–(J4).

(J1) An expression E that contains a "free pronoun" P denotes a function from entities of the standard type of P to entities of the standard type of E.

Consider for instance the noun phrase *the woman who gave him birth* or, equivalently *his mother* (as in (5)). Assume that the standard type of NPs is e and that this is also the standard type of the pronoun *him*. Jacobson therefore assumes that the whole NP denotes a function of type *ee:* a function from entities to entities. In the example, this is the function that maps every (adult male human) entity to its mother.

(J2) Transitive predicates like *love,* of the standard type e (et), have a secondary meaning of type (ee) (et) that ranges over ee functions in the object argument.

This reading enables the subject NP to "bind" a pronoun within the objects. The operator that derives this additional meaning of transitive predicates is denoted 'Z' and is defined as follows.

(10) $Z_{(e(et))((ee)(et))} \overset{\text{def}}{=} \lambda R_{e(et)}.\lambda f_{ee}.\lambda x_e.R(f(x))(x).$

In words: the Z function maps a binary relation R to the relation $Z(R)$ that holds exactly between those ee-type functions f and entities x that satisfy $R(f(x))(x)$. For instance, the following example (11a) is analyzed as in (11b).[3]

(11) a. Every man loves his mother.

   b. $\textbf{every}'_{(et)((et)t)}(\textbf{man}'_{et})(Z(\textbf{love}'_{e(et)})(\textbf{his\_mother}'_{ee}))$

   $\Leftrightarrow \textbf{every}'(\textbf{man}')(\lambda x_e.\textbf{love}'(\textbf{his\_mother}'(x))(x))$

   $\Leftrightarrow \forall x[\textbf{man}'(x) \rightarrow \textbf{love}'(\textbf{his\_mother}'(x))(x)].$

The determiner *every* here standardly denotes the subset relation between sets, or in lamda format: $\lambda A_{et}.\lambda B_{et}.\forall x_e[A(x) \rightarrow B(x)]$.

(J3) Items like the definite article *the*, relative pronoun *that* and the copula *be* can range over ee functions as well as "ordinary" e-type entities. In essence, we can assume that such items denote the (polymorphic) *iota*, intersection and identity functions respectively.

(J4) Intransitive restricting predicates like *woman*, of the standard type et, have a second meaning of type (ee)t. This meaning ranges over ee

functions and allows the restriction to combine with functional relative clauses. I use 'N' to denote the operator that derives this additional meaning of transitive predicates. This operator is defined as follows.

(12) $N_{(et)((ee)t)} \stackrel{\text{def}}{=} \lambda P_{et}.\lambda f_{ee}.\forall x_e[P(f(x))]$.

For instance, the $(ee)t$ denotation of the noun *woman* that the N operator derives is the set of *ee* functions that map each entity to a woman. Note that this set is empty when there are no women in the model.

For sake of exposition, I will use here a slightly modified version of the Z operator that Jacobson uses for binding. This revised operator, which is denoted 'Z^0', allows generalized quantifiers of type $(et)t$, rather than e type entities, to combine directly with the binary relation that is modified by the operator. Its definition follows.

(13) $Z^0_{(e(et))(((et)t)((ee)t))} \stackrel{\text{def}}{=} \lambda R_{e(et)}.\lambda Q_{(et)t}.\lambda f_{ee}.Q(\lambda x_e.R(f(x))(x))$.

This operator has essentially the same consequences of Jacobson's Z operator, but its arguments are now a quantifier and an *ee* function (in this order), instead of an *ee* function and an entity as in Jacobson's analysis. This revised formulation of Z only comes to allow a generalized quantifier such as *every man* in the relative clause *that every man loves* to "saturate" the subject argument of the transitive predicate *loves,* without getting into complex questions concerning the derivation of this option within a general categorial theory.[4]

Now we can get back to the sentences in (5) and illustrate their analysis in Jacobson's approach. The meaning derivation of sentence (5a) is summarized in Figure 1. The type variable τ stands for any monomorphic type. It is not hard to verify that the result of this derivation is tantamount to the following statement:

(14) There is only one *ee* function f s.t. f maps every man to a woman he loves, and this function is the function that maps every man to his mother.

Jacobson argues that this paraphrase captures the intuitive meaning of sentence (5a).[5] Furthermore, this meaning is derived without any scope shifting
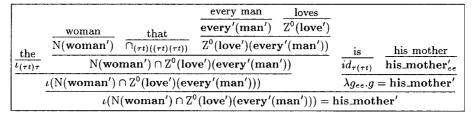

Figure 1. Jacobson's derivation of meaning for sentence (5a).

of the noun phrase *every* man, which, as mentioned above, would have been problematic.

Alexopoulou and Heycock (2002) observe that copular sentences exhibit functional readings also in cases where their subject is not necessarily a singular definite as in (5). Some examples for this observation follow.

(15)   a. One/a (certain) woman that every man loves is his mother.
       b. One/a (certain) woman that no man loves is his mother-in-law.

(16)   A woman that every/no man would be happy to see again is his childhood sweetheart.

(17)   a. (The) two women that every Frenchman admires are his mother and Brigitte Bardot (after Engdahl (1986, ch. 4)).
       b. (The) two women that no Frenchman admires are his mother-in-law and Margaret Thatcher.

Jacobson does not mention such sentences, and their functional readings are not immediately captured by her mechanism. However, along the lines of her proposal, it is natural to assume that pre-nominal items such as *one*, *a* and, (*the*) *two* should also be analyzed as polymorphic operators, similar to the polymorphic *iota* operator that Jacobson assigns to the definite article. For instance, the items *one* and *a* in (15) and (16) could be analyzed as polymorphic existential determiners, of type $(\tau t)((\tau t)t)$. When τ is the *ee* type, of functions from entities to entities, this would allow the determiner to compose with a restricting predicate of type $(ee)t$ as in Jacobson's polymorphic analysis of the definite article. Consider the existential determiner of the relevant type.

(18) $\mathbf{a}'_{((ee)t)(((ee)t)t)} \overset{\text{def}}{=} \lambda F_{(ee)t}.\lambda G_{(ee)t}.\exists f_{ee}[F(f) \wedge G(f)].$

Substituting this existential determiner for ι in Jacobson analysis in figure 1, we get the following meaning.

(19) $\mathbf{a}'(\mathrm{N}(\mathbf{woman}') \cap \mathrm{Z}^0(\mathbf{love}')(\mathbf{every}'(\mathbf{man}')))(\lambda g_{ee}.g = \mathbf{his\_mother}')$

       $\Leftrightarrow \exists f[(\mathrm{N}(\mathbf{woman}') \cap \mathrm{Z}^0(\mathbf{love}')(\mathbf{every}'(\mathbf{man}')))(f) \wedge f = \mathbf{his\_mother}']$

       $\Leftrightarrow (\mathrm{N}(\mathbf{woman}') \cap \mathrm{Z}^0(\mathbf{love}')(\mathbf{every}'(\mathbf{man}')))(\mathbf{his\_mother}').$

This analysis makes the requirement of the mother function that is intuitively expressed by sentence (15a), and similar analyses of the sentences in (15)–(17) can be derived in the same method. However, a general analysis of determiners as ranging over *ee* functions would be unnecessarily strong. This is because other NPs do not lend themselves so easily to functional readings. Consider for instance the following unacceptable sentences.

(20)   ??At most/at least one woman that every man loves is his mother.

(21)   ??No woman that every man would be happy to see again is his
        childhood sweetheart.

(22)   ??Between two and three women that every Frenchman admires are his
        mother, Brigitte Bardot and possibly Isabelle Adjani.

In (22), for instance, it could be expected that the sentence should entail that
there are between two and three functions that send each Frenchman to a
woman he admires, in a similar way to the entailment from (17a) to the
existence of two such functions. However, sentence (22) is quite incoherent.
Thus, it would not be too helpful to assign a polymorphic meaning to all the
determiners in natural language, since we only need a proper subset of them
(if any) to range over functions.[6] In this paper I propose that in fact, *no*
determiner should range over functions. According to the proposed mecha-
nism, functional quantification is only existential and it is derived by the same
general functional process that is responsible for the interpretation of wide
scope indefinites. The unified process that will emerge will also cast some
doubts on the usefulness of the other polymorphic entries that are assumed in
(J3) and of the N operator that is assumed in (J4).[7]

2.2. WIDE SCOPE INDEFINITES

The main reason for introducing functions in the semantic analysis of inde-
finites is their ability to take scope beyond syntactic islands, which has re-
ceived much attention in the literature since (Fodor and Sag, 1982). Consider
for instance the following example.

(23)   If some/a (certain) girl I know arrives to the party then John will be
        happy.

This sentence has the reading that is paraphrased in (24) below, which makes
a statement about a ''specific girl''.

(24)   There is a girl x I know such that if x arrives to the party then John will
        be happy.

We say that under this reading, the indefinite takes *wide scope* (WS) over the
conditional. This reading is sometimes sloppily referred to as the *WS reading*
of the indefinite. This behavior of indefinites is in sharp contrast with the
behavior of other NPs. For instance, the following sentence in (25a) does not
have the analysis that is paraphrased in (25b).

(25)   a. If every girl arrives to the party then John will be happy.
        b. For every girl x, if x arrives to the party then John will be happy.

This claim can be attested by considering that sentence (25a) can be intui-
tively false in situations where the statement in (25b) is true. For instance,

when John would be happy if *any* girl comes to the party, while not wanting *all* the girls to come to the party together.

The special scopal behavior of indefinites calls for an explanation. In recent years, many works have followed (Reinhart, 1992, 1997; Kratzer, 1998), and assumed that the problem should be solved by allowing indefinites to be interpreted using *choice functions* (CFs), whose definition is given below.

DEFINITION 1 (**Choice functions**). Let E be a non-empty set. *A* function f from. $\wp(E)$ to E is a choice function over E iff for every $A \subseteq E$: if A is not empty then $f(A) \in A$.

In extensional type logical frameworks like the one assumed throughout this paper, the set CH^τ of *choice functions over type* τ is defined as follows:

(26) $CH^\tau \overset{\text{def}}{=} \lambda f_{(\tau t)\tau}.\forall P_{\tau t} \neq \emptyset[P(f(P))].$

Convention: we often write 'CH' instead of 'CH^e'.

The WS behavior of indefinites is treated using CFs by letting a free CF variable apply to the restriction of the indefinite. We assume that an existential quantifier applies to this variable at the matrix level, and derives the following interpretation of sentence (23).

(27) $\exists f[CH(f) \wedge [\textbf{arrive}'(f(\textbf{girl}')) \rightarrow \textbf{glad}'(\textbf{j}')]]$.

The introduction of an existential quantifier over CFs – especially at levels lower than the matrix level – is controversial. Some authors, notably (Kratzer, 1998), favor a usage of CFs as "deictic" entities, without existential quantification over them. This controversy is not central for the purposes of this paper, and readers may easily modify the analyses below to the "deictic" version of CFs. Another problematic point which is not directly relevant to our purposes here is the treatment of indefinites with an empty restricting predicate. For example, consider sentence (23) when there happens to be no girl in the given situation. In Winter (1997, 2001, ch. 3) this case is given a solution using CFs of a higher type. For the sake of presentation I will not employ here this more complex analysis.

We have seen that indefinites with *some* and *a certain* show WS behavior beyond the "adjunct island" of the conditional. The same holds for such indefinites in other syntactic environments that behave like scope islands.[8] However, many other NPs (e.g. universal NPs as in (25a)) do not show this island-free behavior. It is therefore important to characterize those NPs to which the CF mechanism applies. This question is taken up in Winter (2001, ch. 4), where it is argued that in addition to simple indefinites as in (23), the CF mechanism also applies to simple numeral indefinites (e.g. *three students*), *WH* phrases and singular and plural definite NPs. The WS potential of

simple numerals can be easily illustrated by the interpretation (28b) of sentence (28a).

(28)   a. If three girls I know arrive to the party then John will be happy.
        b. There is a set A of three girls that I know such that if the girls in A
           arrive to the party then John will be happy.

The WS potential of *WH* elements is exemplified by Reinhart (1997) using question-answer pairs like the following:

(29)   Who will be offended if we invite which philosopher? John will be offended if we invite Putnam.

To interpret the noun phrase *which philosopher,* early semantic theories of questions would have to assign it sentential scope over the conditional. Reinhart however shows that the CF mechanism can treat such effects of *WH in situ* in a similar fashion to the treatment of indicative indefinite that was reviewed above. As for definite NPs, because of their uniqueness requirement it is not easy to test their scopal behavior. I refer the reader to Winter (2001, ch. 4) for other interpretative effects with definites that are accounted for by their CF interpretation.

Crucially, we have seen above that all these kinds of NPs – simple singular/plural (in) definites and *WH* elements – also lead to functional readings (cf. (3), (5) and (15)–(17)). Conversely, NPs as in (20)–(22) do not give rise to wide scope readings beyond islands, and therefore they do not require the CF analysis. Consider for instance the contrast between *at least one* in the following sentence and *some* in (23) above.

(30)   If at least one girl (I know) arrives to the party then John will be happy.

This sentence does not have the reading that is paraphrased below.

(31)   There is at least one girl I know x such that if x arrives to the party then John will be happy.

Similar cases where WS interpretations are unavailable can also be illustrated for the other NPs in (20)–(22). This distinction between NPs where the CF analysis is available and other NPs where it is unavailable, is addressed by the syntactic-semantic mechanism that is proposed in Winter (2000, 2001, ch. 4). In these works I propose that the two kinds of nominals correspond to two different syntactic classes of DP structures. It is proposed that so-called *flexible DPs,* like simple (in)definites and WH phrases, are basically predicative, and therefore in argument position they are interpreted using CFs. Other DPs, which are not amenable to application of flexibility principles, are accordingly called *rigid.* These NPs include complex numerals (e.g. *at least one* or *exactly one*) and universal quantifiers, and involve a full DP structure.

Consequently, these DPs are purely quantificational and have no CF interpretation.

Of course, the parallelism we observed between WS interpretation and the availability of functional readings is not accounted for in the absence of a unified theory of both phenomena. However, before moving on to the details of this theory, we need to discuss a crucial ingredient of the proposal: general Skolem functions of arbitrary arity.

2.3. GENERAL SKOLEM FUNCTIONS

It is interesting to note that, quite independently of the scope of indefinites that has occupied semantic theories in recent years, some earlier works had proposed to use general *Skolem functions* (SFs) for capturing other semantic effects with indefinites and interrogatives. Intuitively, we can think of an n-ary Skolem function as a choice function with n parameters. While a choice function f maps any non-empty set X to a member of X, a general Skolem function f_n of arity n maps such a set X, together with tuple of n parameters $\langle x_1, \ldots, x_n \rangle$, to a member of X. Since we want to treat parameters in such tuples as "free variables" in Jacobson's variable-free framework, we assume that the input to a Skolem function is an n-place *function* that has the power-set $\wp(E)$ as its range. A variable-free Skolem function modifies such a function to a function that has E as its range. This is formally stated in the following definition.

DEFINITION 2 (**Skolem functions of arity** ≥ 1). Let E and A be non-empty sets. A function f from $(\wp(E))^A$ to E^A is a Skolem function iff for every function $g \in (\wp(E))^A$, for every $x \in A$ s.t. $g(x) \neq \emptyset : (f(g))(x) \in g(x)$. If A is a Cartesian product of arity $n \geq 1$ then we say that f is a Skolem function of arity n.

In a purely functional type-theoretical format, the set A cannot be a Cartesian product. However, standard 'Currying' allows us to simulate products using functions and generalize the notion of Skolem functions. SFs of arity $n \geq 0$ are defined as objects of the following type scheme:

$$(\tau_1(\ldots(\tau_n(\tau t))\ldots))(\tau_1(\ldots(\tau_n\tau)\ldots)).$$

In this type scheme, the argument of an SF is a function of type $\tau_1(\ldots(\tau_n(\tau t))\ldots)$, which is isomorphic to a function from the Cartesian product $D_{\tau 1} \times \cdots \times D_{\tau n}$ to the sets of τ-type elements. An SF maps such a function to a function of type $\tau_1(\ldots(\tau_n\tau)\ldots)$, which is isomorphic to a function from the Cartesian product $D_{\tau_1} \times \cdots \times D_{\tau_n}$ to τ-type elements.

The set SK^n of general Skolem functions of this type is defined as follows.

(32) $\quad SK^n \stackrel{\text{def}}{=} \lambda f. \forall g_{\tau_1(\dots(\tau_n(\tau t))\dots)} \forall x_1 \dots \forall x_n$

$$[g(x_1)\dots(x_n) \neq \emptyset \rightarrow (g(x_1)\dots(x_n))((f(g))(x_1)\dots(x_n))]$$

Convention: We usually assume that $\tau_1 = \dots = \tau_n = \tau$, and say that a function in $SK^{\tau,n}$ is a *Skolem function* (of arity n) *over type* τ. We often write 'SK^n' instead of '$SK^{n,e}$'. With this notation it is clear that the SF of arity 0 over type τ are simply the CFs over τ.

Consider the following example for a Skolem function f of arity 1. Assume that $E = \{a, b, c\}$. Let g be a function of type $e(et)$ that maps a to the singleton set $\{c\}$ and maps both b and c to the set $\{a, b\}$. Then the function $f(g)$ must be one of the four *ee* functions $h_1, \dots, h_4$ that are described in Table I. Note that h_2 and h_3 illustrate the liberty of $f(g)$ to "choose" a different element of the set $\{a, b\}$ for each of the two "parameter values" b and c. It is in this sense that SFs of arbitrary arity are more powerful than CFs.

The close relationships between WS indefinites and general SFs on the one hand, and Jacobson's treatment of functional readings on the other hand, are most easily demonstrated by the following examples due to Groenendijk and Stokhof (1984, ch.3) and Hintikka (1986), respectively.

(33)   Every man loves a (certain) woman – his mother.

(34)   According to Freud, every man unconsciously wants to marry a certain woman – his mother.

Both Groenendijk & Stokhof and Hintikka propose to treat such examples using an existential quantifier over SFs that takes matrix scope. In the current setting, this means that the restricting predicate that the noun *woman* denotes must be a two-place relation, of type $e(et)$. This is surprisingly reminiscent of Jacobson's assumption (J4), according to which restricting predicates are systematically of a higher type. Jacobson's N operator maps the set of women to the set of *ee* functions that map any entity to some woman or other. Instead, now we need to map the set of women to a "parameterized set of women": the constant function of type $e(et)$ that maps any entity to the set of women. In view of this similarity, let us use the notation 'N^0' to denote the general operator that derives such functions from ordinary denotations of nouns. Formally, the N^0 operator is defined as follows.

Table I. An $e(et)$ function and some *ee* functions

	g	h_1	h_2	h_3	h_4
$a \mapsto$	$\{c\}$	c	c	c	c
$b \mapsto$	$\{a,b\}$	a	a	b	b
$c \mapsto$	$\{a,b\}$	a	b	a	b

(35) $N^0_{(et)(e(et))} \stackrel{\text{def}}{=} \lambda P_{et}.\lambda x_e.\lambda y_e.P(y)$

Using SFs, we can now model discourses such as (33) and (34) by existential quantification over an SF at the matrix level. The result is that each of these sentences is analyzed as a statement about the existence of a certain function, where the second part of each sentence specifies its identity. Even without getting into the technical derivation of this "discourse specification", it is clear that existential quantification over SFs accounts for the interpretation of such sentences.[9] To give an illustration of this treatment, consider the following analysis of (33).

(36) $\exists f_{(e(et))(ee)}[SK^1(f) \wedge \textbf{every}'(\textbf{man}')(Z(\textbf{love}')(f(N^0(\textbf{woman}'))))]$.

In words: there is an Skolem function f of arity 1, which maps the function $N^0(\textbf{woman}')$ – the constant function that maps every entity to the set of women – to an ee function h. This h function furthermore has the property that every man is in $Z(\textbf{love}')(h)$. More simply: every man loves the woman that h assigns to him. Assuming that the set of women is not empty, this is equivalent to the standard analysis of sentences like (33), with narrow scope existential quantification over e-type entities:

(37) $\forall x \in \textbf{man}' \exists y \in \textbf{woman}'[\textbf{love}'(y)(x)]$

We see that in these examples, general SFs are employed to account for *narrow scope* readings of indefinites that due to the anaphora have a wide scope "functional flavor". It is curious to note that Reinhart's usage of the simpler 0-arity SFs (i.e. CFs) for deriving "ordinary" *wide scope* readings of indefinites was discovered rather late in the development of the theory.

In addition to the use of general Skolem functions in the treatment of indefinites as in (33) and (34), there are (at least) four other types of motivations that were given in the literature for their introduction:

1. Kratzer (1998) objects to the existential closure of CFs in Reinhart's analysis. Instead, Kratzer proposes to use general SFs as a means for deriving readings that under Reinhart's account require existential quantification over CFs with scope that is narrower than matrix scope.
2. Schlenker (1998) and Winter (1998, 2001, ch. 3) argue (independently) that certain sentences with indefinites show readings that are inexpressible using CFs alone and require SFs of arbitrary arity.[10]
3. In Winter (2001, ch. 3) it is argued that SFs are required in order to eliminate problematic analyses that are derived using CFs for indefinites with free pronouns (e.g. *a woman he knows*).
4. Chierchia (2001) argues that certain effects with indefinites, including weak crossover and *de re* interpretations, require introduction of SFs of arbitrary arity.

I will not discuss these issues in this paper, but simply assume, as in most works on this subject, that general SFs are required for the treatment of indefinites. The exact restrictions on their introduction is a topic that is still under extensive investigation.[11] Section 3 shows how this assumption allows us to achieve a unified treatment of functional readings and WS indefinites.

3. A Unified Mechanism for Functional Readings and Wide Scope Indefinites

In Section 2, we have seen technical, intuitive and distributional reasons to assume that "functional readings" and "wide scope indefinites" are two phenomena that should be derived by the same mechanism. Technically, the SF treatment of the scope of indefinites involves functions from entities to entities, like the ones that were independently employed by Jacobson and others for treating sentences with functional readings. Intuitively, the "wide scope functional" interpretation of discourses as in (33) and (34) seems to be essentially the same kind of thing as the functional interpretation of questions such as (3) or copular sentences such as (5). From a distributional point of view, we saw that the same NPs that are treated using SFs and can show WS effects beyond islands also show "functional" effects. It is therefore natural to expect the two phenomena to be amenable to the same treatment.

However, as mentioned in the introduction, one link is missing between the two kinds of theories that were reviewed above. Jacobson assumes that the restricting predicate in NPs with functional readings *ranges* over functions. Consequently, in her account any NP should in principle allow quantification over functions. By contrast, the CF treatment assumes that choice functions *apply to* the restricting predicate, which denotes a set of "ordinary" entities of type *e*. Quantification over CFs, if needed at all, is only existential, and applies syncategorematically: independently of the syntax/ semantics of the NP. The same holds of the way SFs of higher arity are used for treating scopal effects with indefinites.

The main argument of this section is that this discrepancy can be resolved by renouncing Jacobson's general polymorphic scheme of quantification over functions. In the proposed modification of her mechanism, functions are involved in the internal semantics of the NP as part of the variable-free mechanism in assumptions (Jl) and (J2). However, the restricting predicate and the determiner range over ordinary entities and are standardly treated using the general SF mechanism. This allows us to renounce Jacobson's assumption (J4) about the N operator. Further polymorphism as in (J3) will be needed only with the copula construction and not with determiners or relative pronouns.

To achieve this unified analysis, the main new part of the proposal is an operator RG ('range' operator) that is defined below.

DEFINITION 3 (**RG operator**). Given two non-empty sets, A and B, the RG operator is a function from $\wp(B^A)$ (the sets of functions from A to B) to $(\wp(B))^A$ (the functions from A to subsets of B) s.t. for all $F \subseteq B^A$, for all $x \in A$:

$$(\mathrm{RG}(F))(x) \overset{\mathrm{def}}{=} \{f(x) : f \in F\}.$$

In other words, if F is a set of functions from A to B, then RG(F) maps each entity x in A to the subset of B that consists of the images of x by the functions in F. As an example, reconsider the functions in table I above, and observe that RG($\{h_1, h_2, h_3, h_4\}$)=g.

For our purposes, the most useful instance of RG is when A and B are both equal to the domain of entities. The RG operator of the respective type is given below in lambda format.

(38) $\mathrm{RG}_{((ee)t)(e(et))} = \lambda F_{(ee)t}.\lambda x_e.\lambda y_e.\exists f \in F[f(x) = y]$.

To illustrate how this operator allows us to unify the two mechanisms, consider again Jacobson's analysis of sentence (5a) that was illustrated in Figure 1. In the revised analysis, the denotation of the "gapped" constituent *every man loves* still denotes the same $(ee)t$ set of functions as in Jacobson's analysis. This is the set $F = Z^0(\mathbf{love'})(\mathbf{every'}(\mathbf{man'}))$ – the set of functions that map every man to something he loves. The RG operator maps F to a binary predicate – the function that maps every man to the *set of* things he loves, provided that every man loves something.[12] Let us henceforth denote this binary relation RG(F) by 'S'. The binary relation S should be intersected with the denotation of common noun *woman*. This can be achieved by lifting the set **woman'** using the N^0 operator that was defined in (35) above, and was used in the treatment of sentence (33) in (36).[13] Consequently, the restricting predicate *woman that every man loves* denotes the function that maps every man to the set of *women* he loves. Let us refer to this binary relation ($N^0(\mathbf{woman'}) \cap S$) by the letter '$R$'. To the binary relation R we can apply the general SF mechanism, as in (36). It is easier to see how this SF mechanism works with indefinites. Therefore, let us consider how the meaning of sentence (39) below is derived. This sentence (=(15a)) is a slight variation on Jacobson's example (5a).

(39)   A (certain) woman that every man loves is his mother.

The analysis of this example using SFs is given in Figure 2, where *EC* stands for "existential closure" – here of an SF of arity 1. In this analysis, the Skolem function f applies to the denotation of the restricting predicate *woman that every man loves*: the binary relation R that sends each man to the set of women he loves. Assuming that $R(x)$ is non-empty for every man x, any SF sends R to an *ee* function that maps every man to one of the women he loves. The

statement that is derived in Figure 2 claims that one of these *ee* functions is the mother function. This is the desired interpretation of sentence (39), and it is equivalent to the Jacobsonian analysis of the same example in (19) above.

The proposed treatment of sentence (39) as described above is presented in a way that highlights the main differences from Jacobson's treatment. These differences (cf. Figure 1 vs. Figure 2) are the introduction of the RG operator and the application of SFs to the restricting predicate. However, a more general treatment is obtained if we adopt the mechanism of Winter, (2001, ch. 3) for handling "free variables". This mechanism uses a special functional type constructor "$\rightarrow$" in order to distinguish the type of denotations that contain "free variables" from the type of other functions.[14] An expression of type $\tau \rightarrow \sigma$ denotes a function from D_τ to D_σ, but it behaves compositionally like an expression of type σ that contains a "gap" (a free variable) of type τ. For instance, a transitive predicate such as *love* gets the standard type $e(et)$. By contrast, a binary relation like the relation R that is denoted by the expression *woman that every man loves* gets the type $e \rightarrow (et)$: it behaves compositionally like an ordinary noun of type *et,* but involves a "free variable" of type e. In order to capture this double nature of expressions with gapped meanings, we use the following categorial rule.

$$\frac{\Gamma, \sigma_1 \vdash \sigma_2}{\Gamma, \tau \rightarrow \sigma_1 \vdash \tau \rightarrow \sigma_2}\text{ACOND}, \qquad \frac{X, y_1(x_\tau) \Rightarrow y_2(x_\tau)}{X, y_1 \Rightarrow y_2}.$$

This rule, presented here in sequent format with the appropriate semantics, is called ACOND (for "Argument Conditionalization"), and it is a restricted version of the general Conditionalization rule of the Lambek Calculus (see Van Benthem, 1991). An example for its application is the following derivation of meaning for the constituent *that every man loves in* (39). Unlike Jacobson, we assume now that the relative pronoun *that* standardly denotes the *monomorphic* intersection function of type $(et)((et)(et))$. The derivation of meaning for the constituent *every man loves* proceeds as in Figure 2. Recall that we denote this meaning by 'S'. Accordingly, we assume the following types and meanings:
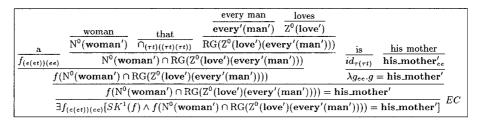

Figure 2. Derivation of meaning for sentence (39).

(40) $[\![\text{that}]\!] = \cap_{(et)((et)(et))} \stackrel{\text{def}}{=} \lambda A_{et}.\lambda B_{et}.\lambda x_e.A(x) \wedge B(x)$

   $[\![\text{every man loves}]\!] = S_{e \to (et)} \stackrel{\text{def}}{=} \text{RG}(\text{Z}^0(\textbf{love}')(\textbf{every}'(\textbf{man}')))$.

Composition of these two types and meanings is achieved by the following application of ACOND:

$$\frac{(et)((et)(et)),\ et \vdash (et)(et)}{(et)((et)(et)),\ e \to (et) \vdash e \to ((et)(et))} \text{ACOND}.$$

To see the use of the semantics of this rule, note the following valid derivation (using application):

$$\cap,\ S(x) \Rightarrow \cap(S(x)),$$

which is tantamount to:

$$\cap,\ (\lambda y.S(y))(x) \Rightarrow (\lambda y.\cap(S(y)))(x).$$

Using ACOND, we can use this derivation as follows:

$$\frac{\cap,\ (\lambda y.S(y))(x) \Rightarrow (\lambda y.\cap(S(y)))(x)}{\cap,\ S \Rightarrow \lambda y.\cap(S(y))} \text{ACOND}.$$

Intuitively, when composing $\cap$ and S, this rule amounts to: introduction of a "fresh free variable" y, application of $\cap$ to $S(y)$ and abstraction over y. Motivations for this particular formulation and for the "$\to$" constructor are given in Winter (2001, ch. 3), and I will not repeat them here.

   In a similar way, the ACOND rule allows us to derive the meaning of the constituent *woman that every man loves,* without applying the N^0 operator to *woman* as in Figure 3. This derivation is given below.

(41)
$$\frac{\dfrac{\dfrac{et,(et)(et) \vdash et}{et, e \to ((et)(et)) \vdash e \to (et)} \text{ACOND},}{\textbf{woman}',\ \cap(S(x)) \Rightarrow \textbf{woman}' \cap (S(x))}}{\dfrac{\textbf{woman}',\ (\lambda y.\cap(S(y)))(x) \Rightarrow (\lambda y.\textbf{woman}' \cap (S(y)))(x)}{\textbf{woman}',\ \lambda y.\cap(S(y)) \Rightarrow \lambda y.\textbf{woman}' \cap (S(y))}} \text{ACOND}.$$

Note that the result is the same as the relation *R,* which was obtained above by intersecting S with $\text{N}^0(\textbf{woman}')$ (cf. Figure 3). Formally, let $\cap_1$ and $\cap_2$ denote the intersection operators of one-place and two-place predicates respectively. Then we have:
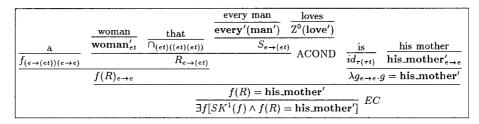
$$\cfrac{\cfrac{f_{(e\to(et))(e\to e)}}{\quad}\quad \cfrac{\cfrac{\cfrac{\text{woman}}{\text{woman}'_{et}}\quad \cfrac{\text{that}}{\cap_{(et)((et)(et))}}}{R_{e\to(et)}}}{f(R)_{e\to e}}\quad \cfrac{\cfrac{\cfrac{\text{every man}}{\text{every}'(\text{man}')}\quad \cfrac{\text{loves}}{Z^0(\text{love}')}}{S_{e\to(et)}}}{}\;\text{ACOND}\quad \cfrac{\cfrac{\text{is}}{id_{\tau(\tau t)}}\quad \cfrac{\text{his mother}}{\text{his\_mother}'_{e\to e}}}{\lambda g_{e\to e}.g=\text{his\_mother}'}}{\cfrac{f(R)=\text{his\_mother}'}{\exists f[SK^1(f)\wedge f(R)=\text{his\_mother}']}\;EC}$$

Figure 3. Revised derivation of meaning for sentence (39).

(42) $[\![\text{woman that every man loves}]\!] = R$

$= (\mathrm{N}^0(\textbf{woman}')) \cap_2 S$

$= (\lambda y.\lambda z.\textbf{woman}'(z)) \cap_2 S$

$= \lambda y.((\lambda z.\textbf{woman}'(z)) \cap_1 S(y))$

$= \lambda y.\textbf{woman}' \cap_1 S(y).$

In words, for any binary relation S: the intersection of S with the relation that sends each entity to the set of women, is the relation that sends each entity y to the set of women in $S(y)$. This means that when the second conjunct in the construction *woman that...* ranges over functions as in Jacobson's account, the outcomes of the N^0 operator are now derived rather than stipulated.[15]

The revised analysis of sentence (39) using the ACOND rule is given in Figure 3, with the abbreviations S and R for the denotations of the expressions *every man loves* and *woman that every man loves* respectively, as derived above. The meaning that is derived in this way is equivalent to the meaning that is derived using the more *ad hoc* mechanism in Figure 3. It is important to mention that the introduction of the free SF variable in this analysis (and the following one) is only for the sake of presentation. In fact, the ACOND rule as introduced above is designed to get rid of such free variables, as explained in Winter (2001, ch. 3).

With the ACOND rule, it is also no longer necessary to adopt Jacobson's assumption (J3) that items such as relative pronouns (e.g. *that)* or the definite article (i.e. *the)* are polymorphic. We adopt the treatment in Winter (2001, ch. 3) of the definite article as a predicate modifier that imposes singularity on the restricting predicate. The Strawsonian (presuppositional) version of this operator is defined as follows.

(43) $\textbf{the}'_{(et)(et)} \overset{\text{def}}{=} \lambda A_{et}.\lambda x_e.(x = \iota(A)).$

Thus, the $\textbf{the}'$ operator sends a set to itself if it contains a unique element, and is undefined otherwise.[16] Using these assumptions, the analysis of sentence

(5a) in Figure 4 becomes analogous to the analysis of sentence (39) in Figure 3.[17]

A similar analysis becomes now possible for copular sentences with bare numerals as illustrated in (17). We adopt the common assumption that the numeral *two* is a predicate modifier, similarly to the definite article. Concretely, given a cardinality function *card* on "plural" *e*-type entities, we can assume the following denotation for the numeral *two*:

(44) $[\![two]\!] = \mathbf{two}'_{(et)(et)} \stackrel{\text{def}}{=} \lambda A_{et}.\lambda x_e.card(x) = 2 \wedge A(x)$

This makes the analysis of (17) similar to the analysis of (5a) in Figure 4.[18]

Note that in the proposed modification of Jacobson's system, the only polymorphic item in the analysis of functional readings is the copula *be*. This is not coincidental. The polymorphism of *be* is what allows predicates such as *be his mother* to range over functions of type $e \rightarrow e$. With a monomorphic transitive predicate such as *love* or *pinch* this is not possible. In the proposed system, there are two ways to analyze a verb phrase such as *love his mother*. The object *his mother* is of type $e \rightarrow e$. The transitive predicate *love* is standardly of type $e(et)$. One way to compose their denotations is using ACOND; another is using the Z (or Z^0) function and direct application. These two options are illustrated below.

(45)   a. $\mathbf{love}'_{e(et)}, \mathbf{his\_mother}'_{e \rightarrow e} \stackrel{\text{ACOND}}{\Rightarrow} (\lambda x.\mathbf{love}'(\mathbf{his\_mother}'(x)))_{e \rightarrow (et)}$

       b. $Z(\mathbf{love}')_{(e \rightarrow e)(et)}, \mathbf{his\_mother}'_{e \rightarrow e} \Rightarrow (\lambda x.\mathbf{love}'(\mathbf{his\_mother}'(x))(x))_{et}$

The first option accounts for the deictic interpretation of the pronoun (as in *Mary loves his mother*). The second option accounts for the bound interpretation of the pronoun (as in *every man loves his mother*). However, the VP denotation cannot combine with $e \rightarrow e$ functions that are denoted by NPs. This accounts for the impossibility to get a functional reading in non-copular sentences such as the following (cf. Sharvit, 1999).

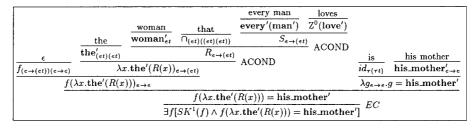(46)   a. The woman that no man admires hates his mother.

Figure 4. Derivation of meaning for sentence (5a).

b. The woman that no man loves pinched him.(= (8a))

Assuming that no man admires his mother-in-law, sentence (46a) cannot mean something like "a mother-in-law always hates her son-in-law's mother". Thus, the sentences in (46), as opposed to (5b), do not have a functional interpretation. This is accounted for by our assumption that ordinary transitive predicates, as opposed to *be*, do not have a polymorphic meaning.[19]

As pointed out by an anonymous reviewer, relying on polymorphism for the derivation of functional readings has some further implications. One implication is that the functional interpretation of questions as in (3) should also involve a polymorphic item. It seems natural to assume that *wh* elements can indeed range over functions, which is the source of this interpretation. Another phenomenon that (Jacobson, 1999) proposed to treat using her functional mechanism is the case of "unexpected" (or "sloppy") inferences. Consider for instance the following inference, of the type that (Jacobson, 1999) attributes to Reinhart (1990).

(47)   Mary will buy what(ever) Bill buys; Bill buys his favorite car ⇒ Mary will buy her favorite car.

Jacobson shows that her binding mechanism using *ee* functions accounts for this kind of inferences and shows some advantages of her analysis over previous one. Again, in order to achieve this in the present setting of Jacobson's mechanism we need a polymorphic item to blame. I think that polymorphism is quite natural to assume for the relative pronoun *what*(*ever*), but I will not try to substantiate this suggestion here, and leave for further study the details of the adaptation of the present account for the sake of analyzing functional questions and "unexpected" inferences.

Let us summarize the main proposals of this section:
– Sets of *ee* functions are mapped to binary relations using the RG operator.
– This allows the general SF mechanism to derive functional readings.
– The analysis of functional readings does not require Jacobson's N operator or polymorphic meanings of items such as *that* or *the*. These are treated using a more general variable-free compositional mechanism.
– Polymorphism of the copula *be* is required in order to account for its licensing of functional readings as opposed to other transitive predicates.

Note that SFs of arity greater than one are needed in order to analyze sentences with more than one pronoun in the post-copular NP. For example:

(48)   The present that every man will send to every woman is the present that she asked him to send her.

For sake of exposition, I will not consider such complex examples in the sequel.

4. Generalized Quantifiers and Functional Readings

Mapping sets of functions to binary relations using the RG operator may potentially result in loss of information. To see that, note that in a model with n entities, the cardinality of the $(ee)t$ domain is $2^{(n^n)}$, whereas the cardinality of the $e(et)$ domain is only $(2^n)^n = 2^{(n^2)}$. Consequently, the RG operator is not one-to-one. This means that the present proposal is, at least in principle, less expressive than Jacobson's original quantification over functions. In this section it is argued that this loss of expressivity is not only innocuous, but actually desired. In the mechanism that was introduced in the previous section, a restricting predicate such as *woman that Q loves* denotes a set of functions F for any quantifier Q. This section introduces a notion of *closed sets of functions,* for which the RG operator is one-to-one, and shows that the set F is closed if and only if Q is a *bounded quantifier:* an intersection of a principal filter and a principal ideal. We will then hypothesize that it is exactly the bounded quantifiers that are licensed in functional interpretations of questions and copular sentences.

To illustrate the intuitive reasoning that underlies the formal discussion in this section, consider the following contrast.

(49)   A woman that $\left\{ \begin{array}{c} \text{no} \\ {}_*\text{at most one} \end{array} \right\}$ man loves is his mother-in-law.

As in Jacobson's analysis, the denotation of the constituent *woman that no man loves* is the set of functions that send no man to a woman he loves. Let us denote this set of functions by F_{no}. Assuming that F_{no} is not empty, the relation $RG(F_{no})$ sends every man to the set of women he does not love. Applying any Skolem function of arity 1 to $RG(F_{no})$ gives us again a function in F_{no}. This is not the case when we consider the determiner *at most one*, which does not support a functional reading in (49). Let us denote the set of functions that send at most one man to a woman he loves by $F_{\leq 1}$. Assume that the men are John and Bill and that the women are Mary and Sue. Assume further that John loves Mary and hates Sue, and that Bill loves Sue and hates Mary. Thus, the set $F_{\leq 1}$ includes the following functions f_1, f_2 and f_3:

$$f_1: \quad \text{John} \mapsto \text{Mary}, \quad \text{Bill} \mapsto \text{Mary},$$
$$f_2: \quad \text{John} \mapsto \text{Sue}, \quad \text{Bill} \mapsto \text{Sue},$$
$$f_3: \quad \text{John} \mapsto \text{Sue}, \quad \text{Bill} \mapsto \text{Mary}.$$

This means that $RG(F_{\leq 1})$ sends both John and Bill to the set {Mary, Sue}. Consequently, a Skolem function of arity 1 may map $RG(F_{\leq 1})$ to the function that sends each of the two men to the woman he loves. But this function is not in $F_{\leq 1}$. This formal distinction between the quantifiers that the noun phrases *no woman* and *at most one woman* denote will be given a general

characterization in this section in terms of bounded and unbounded quantifiers. It will be hypothesized that (un)boundedness of quantifiers is responsible for contrasts as in (49).

4.1. Closed sets of functions and bounded quantifiers

Consider first the following definition.

DEFINITION 4 (**Closed sets of functions**). Let $F \subseteq B^A$ be a set of functions from A to B. The closure of F is the set of functions $\bar{F} \subseteq B^A$ that is defined by:

$$\bar{F} \stackrel{def}{=} \{f \in B^A : \text{ for every } x \in A \text{ there is } g \in F \text{ s.t. } f(x) = g(x)\}.$$

We call F a *closed set of functions* if $F = \bar{F}$.

EXAMPLE: Consider the functions in Table I above. The set $\{h_1, h_2\}$ is closed. By contrast, the set $\{h_1, h_4\}$ is not closed: its closure is $\{h_1, h_2, h_3, h_4\}$.

Note that $F \subseteq \bar{F}$ trivially holds for any set of functions F. It is also easy to observe the following fact.

FACT 1. If F and G are sets of functions in B^A then $RG(F) = RG(G)$ iff $\bar{F} = \bar{G}$.

As a result, the RG operator is one-to-one when it is restricted to apply to closed sets of functions:

COROLLARY 2. If F and G are closed sets of functions in B^A then $RG(F) = RG(G)$ iff $F = G$.

A related fact about closed sets of functions, which was informally mentioned and illustrated above, concerns their relations with Skolem functions and the RG operator.

FACT 3. A non-empty set F of functions in B^A is closed iff for every Skolem function f from $(\wp(B))^A$ to $B^A : f(RG(F)) \in F$.

Proof ($\Rightarrow$): Assume that F is a non-empty closed set of functions in B^A and that f is an SF of the given signature. For every $x \in A$: from $F \neq \emptyset$ it follows that $(RG(F))(x) \neq \emptyset$, and by definition of SFs: $(f(RG(F)))(x) \in (RG(F))(x)$. By definition of RG, this means that for every $x \in A$ there is $g \in F$ such that $(f(RG(F)))(x) = g(x)$. Hence $f(RG(F)) \in \bar{F}$. But $F = \bar{F}$ by closure of F, hence $f(RG(F)) \in F$.

($\Leftarrow$): Assume that for every Skolem function f of the given signature: $f(RG(F)) \in F$. By definition of SFs, this means that for every function $g \in B^A$: if g sends every $x \in A$ to a member of $(RG(F))(x)$ then $g \in F$. But this means that $\bar{F} \subseteq F$, hence $\bar{F} = F$, so F is closed. $\square$

Thus, we know that for a non-empty closed set of functions F, sequential application of RG and a Skolem function maps F to one of its members. Recall that in the compositional mechanism of Section 3, the set F is the denotation of a restricting predicate such as (*woman that*) *every man loves,* which is derived by applying the Z^0 operator to a quantifier Q (e.g. the denotation of *every man*) and a binary relation R (e.g. the denotation of *loves*). Let us use the following abbreviation:

(50) $F_{QR} \overset{\text{def}}{=} Z^0(Q_{(et)t})(R_{e(et)}) = \lambda f_{ee}.Q(\lambda x_e.R(f(x))(x)).$

Further modification of the set F_{QR} within the relative clause is immaterial for our present purposes, hence ignored.[20] In set-theoretical format, for any quantifier $Q \subseteq \wp(E)$ and binary relation $R \subseteq E^2$, the set of functions F_{QR} is $\{f \in E^E : \{x : R(f(x), x)\} \in Q\}$.

The main theorem of this section characterizes the class of quantifiers Q that guarantee that F_{QR} is a closed set of functions for every R. It will be shown that this class is precisely the class of *bounded* quantifiers: those quantifiers that are an intersection of a principal filter (e.g. the denotation of *every student*) and a principal ideal (e.g. the denotation of *no teacher*). This class of quantifiers is officially defined below.

DEFINITION 5 (**Bounded quantifiers**). A quantifier $Q \subseteq \wp(E)$ is called *bounded* iff there are two sets X and $Y \subseteq E$ s.t. $Q = \{A \subseteq E : X \subseteq A \subseteq Y\}$. In this case we say that Q is bounded by X and Y.

Table II gives some NPs that denote bounded quantifiers with the sets that bound them.

THEOREM 4. *A quantifier $Q \subseteq \wp(E)$ is bounded iff the set of functions F_{QR} is closed for any binary relation $R \subseteq E^2$.*

The proof of this theorem makes use of the *convexity*[21] property of quantifiers, which is defined below.

DEFINITION 6 (**Convex quantifiers**). A *quantifier $Q \subseteq \wp(E)$ is* convex iff for all $A \subseteq B \subseteq C \subseteq E$: if $A \in Q$ and $C \in Q$ then $B \in Q$.

For instance, the noun phrase *between three and five students* denotes a convex quantifier in any model, but this is not true of the disjunctive noun phrase *exactly three or exactly five students*.

The following simple fact will be useful in proving Theorem 4.

FACT 5. A quantifier $Q \subseteq \wp(E)$ is bounded iff Q is convex and closed under arbitrary intersections and unions. In this case $Q = \{A \subseteq E : \cap Q \subseteq A \subseteq \cup Q\}$.

Table II. Bounded quantifiers

Noun phrase	Quantifier bounded by	
Every student	**student**$'$	E
No student	$\emptyset$	$E\backslash$**student**$'$
Every student but no teacher	**student**$'$	$E\backslash$**teacher**$'$
Every student but Mary	**student**$'\backslash\{$**m**$'\}$	$E\backslash\{$**m**$'\}$
No student but Mary	$\{$**m**$'\}$	$E\backslash($**student**$'\backslash\{$**m**$'\})$

Proof of Theorem 4

$\Rightarrow$: We assume that $Q \subseteq \wp(E)$ is bounded by $X, Y \subseteq E$ and show that F_{QR} is a closed set of functions for any $R \subseteq E^2$. It is enough to show that any $f_0 \in \overline{F_{QR}}$ is also in F_{QR}. Thus, we need to show that for any $f_0 \in \overline{F_{QR}}$, for any $R \subseteq E^2$:

the set $A_0 \overset{\text{def}}{=} \{x \in E : R(f_0(x), x)\}$ is in Q, which holds iff $X \subseteq A_0 \subseteq Y$.

To show that $A_0 \subseteq Y$, we assume $x_0 \in A_0$ and show $x_0 \in Y$. From $f_0 \in \overline{F_{QR}}$ it follows that some $g_0 \in F_{QR}$ satisfies $g_0(x_0) = f_0(x_0)$. Because $g_0 \in F_{QR}$, we conclude from the definitions of F_{QR} and Q that $\{x \in E : R(g_0(x), x)\} \subseteq Y$. (i)

From the assumption $x_0 \in A_0$, it follows that the relation $R(f_0(x_0), x_0)$ holds. Because $g_0(x_0) = f_0(x_0)$, we conclude that $R(g_0(x_0), x_0)$ holds too. (ii) Facts (i) and (ii) entail that x_0 is in Y. The proof that $X \subseteq A_0$ is similar.

($\Leftarrow$): Assume that for a quantifier $Q \subseteq \wp(E)$, the set of functions F_{QR} is closed for any binary relation $R \subseteq E^2$. By Fact 5, we need to show that Q is convex and closed under arbitrary intersections and unions. The proofs of the convexity and closure properties of Q are all similar.

Convexity of Q: Suppose, for contradiction, that there are $A \subsetneq B \subsetneq C \subseteq E$ s.t $A, C \in Q$ but $B \notin Q$. We conclude that $|E| \geq 2$ and denote $a, b \in E$ for two arbitrary elements $a \neq b$. Consider the relation $R \overset{\text{def}}{=} \{a\} \times E$ and the characteristic functions χ_A, χ_B and χ_C in $\{a, b\}^E$ of the sets $A, B, C \subseteq E$. By saying that $\chi_Y \in \{a, b\}^E$ is the characteristic function of a set $Y \subseteq E$ we mean that χ_Y is the function that satisfies for any $x \in E$:

$$\chi_Y(x) = \begin{cases} a & x \in Y, \\ b & x \notin Y. \end{cases}$$

Observe now the following equalities:

$$\{x \in E : R(\chi_A(x), x)\} = A,$$
$$\{x \in E : R(\chi_B(x), x)\} = B,$$
$$\{x \in E : R(\chi_C(x), x)\} = C.$$

By definition of F_{QR} and the non-convexity assumption about Q it therefore follows that $\chi_A, \chi_C \in F_{QR}$, but $\chi_B \notin F_{QR}$.

However, by the assumption $A \subsetneq B \subsetneq C$, we have for every $x \in E$:

$$x \in B \Rightarrow \chi_B(x) = a = \chi_C(x),$$
$$x \notin B \Rightarrow \chi_B(x) = b = \chi_A(x).$$

Thus, for every $x \in E$ there is $f \in \{\chi_A, \chi_C\}$ s.t. $\chi_B(x) = f(x)$.

We conclude that $\chi_B \in \overline{\{\chi_A, \chi_C\}} \subseteq \overline{F_{QR}}$, which means that $\overline{F_{QR}} \neq F_{QR}$.

By assuming non-convexity of Q, we contradicted the assumption that F_{QR} is a closed set of functions for any relation R. Hence: Q is convex.

Closure of Q under intersections: Suppose, for contradiction, that there is a set $\mathcal{A} \subseteq Q$ s.t. $\cap \mathcal{A} \notin Q$ ($|\mathcal{A}| \geq 2$). Again, we conclude that $|E| \geq 2$, denote $a, b \in E$ for two arbitrary elements $a \neq b$, and consider the binary relation $R = \{a\} \times E$. Now consider the characteristic functions in $\{a, b\}^E$ of the sets in $\mathcal{A}$ and of their intersection $\cap \mathcal{A}$. Observe now the following equalities:

$$\{x \in E : R(\chi_A(x), x)\} = A, \quad \text{for every} A \in \mathcal{A},$$
$$\{x \in E : R(\chi_{\cap \mathcal{A}}(x), x)\} = \cap \mathcal{A}$$

By definition of F_{QR} and the assumption about Q it therefore follows that $\chi_A \in F_{QR}$ for every $A \in \mathcal{A}$, but $\chi_{\cap \mathcal{A}} \notin F_{QR}$.

However, for every $x \in E$:

$$x \in \cap \mathcal{A} \Rightarrow \chi_{\cap \mathcal{A}}(x) = a = \chi_A(x) \text{ for any } A \in \mathcal{A},$$
$$x \notin \cap \mathcal{A} \Rightarrow \chi_{\cap \mathcal{A}}(x) = b = \chi_B(x) \text{ for at least one } B \in \mathcal{A}.$$

Thus, for every $x \in E$ there is $f \in \{\chi_A \in \{a, b\}^E : A \in \mathcal{A}\}$ s.t. $\chi_{\cap \mathcal{A}}(x) = f(x)$.

We conclude that $\chi_{\cap \mathcal{A}} \in \overline{\{\chi_A \in \{a, b\}^E : A \in \mathcal{A}\}} \subseteq \overline{F_{QR}}$, which means that $\overline{F_{QR}} \neq F_{QR}$.

This contradiction to our assumption about the closure of F_{QR} entails that Q is closed under arbitrary intersections.

The proof of the closure of Q under unions is similar. $\qquad \square$

4.2. THE DISTRIBUTION OF QUANTIFIERS IN FUNCTIONAL READINGS

Consider the following examples.

(51) The/A woman that $\left\{ \begin{array}{l} \text{every/no man (but John)} \\ \text{*at most one man} \\ \text{?exactly/at least one man} \end{array} \right\}$ loves is his mother

(52).   Which woman does $\left\{ \begin{array}{l} \text{every/no man (but John)} \\ \text{*at most one man} \\ \text{?exactly/at least one man} \end{array} \right\}$ love? His mother

These examples show a contrast in acceptability of functional readings between bounded quantifiers such as *every, no,* etc. and unbounded quantifiers such as *exactly one, at most/least one,* etc. The following hypothesis makes a natural connection between this linguistic contrast and the formal properties that were proven in the previous subsection.

(53)   **Hypothesis:** *the* RG *operator applies only to closed sets of functions.*

Let us examine how this hypothesis makes the desired connection. As shown by Theorem 4, it is exactly the bounded quantifiers that generate closed sets of functions. Thus, if hypothesis (53) is correct, then the functional mechanism is analyzed as sensitive to the (un)boundedness property of the quantifier. This is due to Theorem 4. Such sensitivity is not at all trivial to explain in compositional frameworks, because the functional mechanism applies at a much higher syntactic level than the level where the quantifier is present. However, Theorem 4 makes the necessary connection between the restrictions on the set of functions that is generated, and the generalized quantifier that participates in its generation.

But *why* should the RG operator be sensitive to whether the set of functions it applies to is closed or not? Fact 3 gives at least a partial answer to this question. According to this fact, these are exactly the closed sets F of functions that guarantee that Skolem functions do not return functions outside F. For instance, an SF that applies to the denotation RG ($[\![$ *woman that every man loves* $]\!]$), does not return functions that were not already in the basic denotation of *woman that every man loves.* Thus, if the hypothesis in (53) is correct, then in a sense it makes the grammar of functional quantification semantically "optimal": functional readings derive "natural functions" exactly in the cases when they are grammatically licensed.[22]

Despite this attractiveness of hypothesis (53), there are (at least) two potential empirical problems that it has to face. One problem concerns the felicitous functional interpretation of sentences such as the following, which is based on examples due to Alexopoulou and Heycock (2002).

(54)   The woman that almost every/no man loves is his mother.

The natural interpretation of the noun phrase *almost every man* is the following generalized quantifier.

(55) $\{A \subseteq E : 1 \leq |\overline{A} \cap \mathbf{man'}| \leq n\}$,

where n is a (small) number that is determined by the context.

This quantifier in not bounded. However, sentence (54) can also be interpreted as in the rough paraphrase below.

(56)   There is a (small) set of men B s.t. the woman that every/no man except for the men in B loves is his mother.

In other words: the "exception set" B that *almost* quantifiers require can take here sentential scope, which allows the noun phrase *almost every/no man* to be interpreted as a bounded quantifier.[23]

Another problem for the hypothesis in (53) may come from certain facts that are pointed out by Sharvit (1999, (18)–(21)). Sharvit considers plural sentences in Hebrew that are parallel to the following English sentence.[24]

(57)   The woman that most of the students invited was their mother.

The determiner *most* is unbounded, hence functional anaphora is *a priori* not expected here by hypothesis (53).

However, with noun phrases such as *most of the students,* also discourse anaphora as in the following example may appear and complicate the picture.

(58)   Most of the students admire their mother. They invited her.

By contrast, discourse anaphora is unacceptable with the other quantifiers that illustrated (un) availability of functional readings in the above examples. For instance:

(59)   No/at most one student hates his mother. *He invited her.

This means that the contrast we observed above between *no* and *at most one* cannot originate from different discourse anaphora potentials, and therefore it supports hypothesis (53). The functional reading in Sharvit's example (57) may result from discourse anaphora and therefore it does not seriously challenge this hypothesis. For recent works that deal with discourse anaphora using a mechanism of functional quantification see Steedman (1999) and Peregrin and Von Heusinger (2003).

5. Further Motivation: A Note on Jacobson (2002)

In her recent article, Jacobson shows motivation for an operator that is very similar to the RG operator, although her motivation comes from different phenomena than the ones that were addressed throughout this paper. The examples that motivate Jacobson's proposal are all similar to the following ones.

(60)   The woman *he loves* that every/no man invited (to his wedding) is his mother.

(61)   A (certain) woman *he loves* that every/no man invited (to his wedding)
       is his mother.

Jacobson (1994) lifts the denotation of the noun *woman* to a set of *ee* func-
tions. However, in her variable-free treatment of anaphora, the denotation of
the relative *he loves*, which should be intersected with the denotation of
woman, is a binary relation. Hence, Jacobson (2002) proposes to shift this
binary relation into a set of functions using an operator that she calls '**m**'. The
effect of this operator is to map any binary relation $R \subseteq A \times B$ to the set of
functions $\{f \in B^A :$ for all $x \in A : R(x,f(x))\}$. In type-theoretical format:

(62) $\mathbf{m}_{(e(et))((ee)t)} \stackrel{\text{def}}{=} \lambda R_{e(et)}.\lambda f_{ee}.\forall x_e R(f(x))(x).$

In Jacobson's categorial treatment of extraction, the relative *he loves* denotes
the following relation:

(63) $\mathbf{he\_loves}' = \lambda x.\lambda y.\mathbf{love}'(y)(x)$

Intersection with the functional denotation of the noun *woman* leads to the
following analysis:

(64) $N(\mathbf{woman}') \cap \mathbf{m}(\mathbf{he\_loves}') =$
       $(\lambda f_{ee}.\forall z[\mathbf{woman}'(f(z))]) \cap (\lambda f_{ee}.\forall u[\mathbf{love}'(f(u))(u)]) =$
       $\lambda f_{ee}.\forall z[\mathbf{woman}'(f(z))] \wedge \forall u[\mathbf{love}'(f(u))(u)]$

In words, this is the set of *ee* functions that send every x to a woman that x
loves. Once this set of functions is derived, the rest of the analysis of sen-
tences (60) and (61) is analogous to the analysis of sentences (5a) and (39) in
Figure 1. (see also (19)). For instance, the analysis of sentence (61) using
Jacobson's **m** operator is given below.

(65) $\mathbf{a}'(N(\mathbf{woman}') \cap \mathbf{m}(\mathbf{he\_loves}') \cap Z^0(\mathbf{invite}'(\mathbf{every}')(\mathbf{man}')))(\lambda g_{ee}.g = \mathbf{his\_mother}')$
       $\Leftrightarrow \exists f[N(\mathbf{woman}') \cap \mathbf{m}(\mathbf{he\_loves}') \cap Z^0(\mathbf{invite}'(\mathbf{every}'(\mathbf{man}')))(f) \wedge f = \mathbf{his\_mother}']$
       $\Leftrightarrow (N(\mathbf{woman}') \cap \mathbf{m}(\mathbf{he\_loves}') \cap Z^0(\mathbf{invite}')(\mathbf{every}'(\mathbf{man}')))(\mathbf{his\_mother}')$

   In the present proposal, since the noun *woman* is anyway lifted to a binary
relation using the N^0 operator, there is no need to use further shiftings, and
the analysis of the relative nominal *woman he loves* is as follows.

(66) $N^0(\mathbf{woman}') \cap (\mathbf{he\_loves}') =$
       $(\lambda x.\lambda y.\mathbf{woman}'(y)) \cap (\lambda x.\lambda y.\mathbf{love}'(y)(x)) =$
       $\lambda x.\lambda y.\mathbf{woman}'(y) \wedge \mathbf{love}'(y)(x)$

From this point the analysis is as in figure 3:

(67) $\exists f_{(e(et))(ee)}[SK^1(f) \wedge f(N^0(\mathbf{woman}) \cap$

   $(\mathbf{he\_loves}) \cap RG(Z^0(\mathbf{invite}')(\mathbf{every}'(\mathbf{man}')))) = \mathbf{his\_mother}']$

This statement is equivalent to the one that is derived above in the Jacobsonian analysis (65).

We conclude, in agreement with Jacobson, that her strategy is in a sense the inverse of the present proposal. Jacobson shifts relatives from binary relations into sets of functions, while in this paper we used the opposite direction – from sets of functions to binary relations – using the RG operator. It is notable that the RG mechanism, which was developed for entirely different reasons than Jacobson's **m** operator, also handles the problem that motivated this latter operator. I also believe that it is likely that Jacobson's mechanism can be adapted to handle at least some of the problems that were treated in the present paper.

6. Conclusions

In this paper we have seen intuitive, distributional and technical support for the claim that "functional" readings of questions and copular sentences and "wide scope" readings of indefinites are two names for the same phenomenon. One simple operator that maps sets of functions to binary relations was used as the key for unifying separate mechanisms that had been previously introduced for treating these phenomena using Skolem functions. This unified mechanism establishes some new relations between functional quantification and generalized quantifier theory. In particular, it was shown that there are both empirical and mathematical reasons to expect functional quantification to be restricted to the newly introduced class of bounded quantifiers. Of course, many hard linguistic-logical questions about functional quantification in natural language are still open. However, I believe that the results in this paper exemplify the benefits that can be gained by the on-going study into this challenging area.

Acknowledgements

Carolyn Heycock, Pauline Jacobson, Manfred Krifka, Ya'acov Peterzil, Tanya Reinhart and Eddy Ruys for discussion. Special thanks to Ruth Kempson, Klaus von Heusinger and an anonymous reviewer of *Research on Language and Computation* for their remarks. Earlier versions of this work have appeared in *Proceedings of the Workshop on Choice Functions and Natural Language Semantics* and in *Proceedings of SALT12*.

Notes

[1] See Engdahl (1986, ch. 4) and Groenendijk and Stokhof (1984, ch. 3) for two classical works on this topic.

[2] Sharvit (1999) claims, following Doron, (1982), that if *no* in (8a) is replaced by *every,* then binding may sometimes become possible, but argues that this option should be derived by the same mechanism that derives pair-list readings with *every* as in (4). A similar point may be related to the fact that some speakers find the statement in (7b) more acceptable as a reading of (7a) when *no* is replaced by *every* or *each.* Since this paper does not deal with pair-list readings, these points are not directly relevant to its main purposes.

[3] Throughout this paper, we employ a typed lambda calculus with equality over the extensional types e, t and their functional compounds. In this formalism, the propositional connectives are easily defined, as well as the logical quantifiers $\exists$ and $\forall$ over any type (see Van Benthem, 1991, p. 7). Lambda terms are freely mixed with the set-theoretical correlates of their denotations, where the domain of entities D_e is also denoted 'E'. For instance: a binary relation is viewed either as a lambda term of type e (et) or as a subset of E^2; a generalized quantifier is viewed either as a lambda term of type (et)t or as a subset of $\wp(E)$; etc.

[4] Jacobson's mechanism achieves this using (a modified version of) the "Geach Rule" and Function Composition.

[5] Strictly speaking, this is not true, since requiring uniqueness of .an *ee* function is much too strong a requirement. In fact, Jacobson argues that domain restriction of the definite article plausibly leads here only to a requirement about the uniqueness of a function from men to women with the properties required above and which is furthermore "natural" in the given discourse.

[6] An anonymous reviewer points out that quantified NPs like those in (20)–(22) are also infelicitous in copular sentences without functional interpretations:

(i) ??At most/at least one woman that every man loves is Sue.

(ii) ??No woman that every man would be happy to see again is Sue.

(iii) ??Between two and three women that every Frenchman admires are Catherine Deneuve, Brigitte Bardot and possibly Isabelle Adjani.

Thus, while I am not claiming here to have an account of unacceptabilities as in (20)–(22) or (i)–(iii), my claim is more modest than that: that a 'functional' meaning of determiners such as *at most one, at least one, no, between two and three,* etc. is unnecessary for analyzing functional readings.

[7] See the derivation in (41) below.

[8] See Ruys (1992, ch. 3) for a concise description of the facts.

[9] Manfred Krifka (p.c.) points out to me that the discourse processing of similar examples may be more complicated than a simple substitution of a Skolem function for a variable. He mentions examples like the following:

(i) Every man loves at most two women – his mother, and if he has one – his sister.

Here, no Skolem function is likely to be operational. However, Skolem functions are required if we like to account for contrasts like the following, where the indefinite NP is within a (complex NP) island.

(ii) Every child loves every man who admires a certain woman – his mother.

(iii) ?Every child loves every man who admires at most two women – his mother, and if he has one – his sister.

We see that the same discourse process that applies in (i) does not apply in (iii), although simple indefinites, as expected, lead to wide scope interpretations beyond islands, as in (ii), as expected by the Skolem function mechanism.

[10] Following Hintikka, Schlenker also uses general SFs for deriving "branching readings" with indefinites. I will not discuss this phenomenon here.

[11] In addition to Chierchia (2001), see also Schlenker (1998) and Schwartz (2001) for some recent work in progress that deal with this topic.

[12] If there is a man who does not love anything, then F is empty and consequently $\mathrm{RG}(F)$ sends everything to the empty set.

[13] Later we will see that this special operator is actually not needed here, and its postulation reflects a more general property of the "binding" mechanism.

[14] Jacobson (1999) identifies expressions that contain "free variables" by their syntactic category rather than by their semantic type, and changes the categorial syntax accordingly.

[15] Note that for cases such as (33) or (34) above, where there is no functional conjunct, an operation like N^0 is still required in order to introduce the "implicit variable" within the noun *woman*. I do not address here the linguistic status of such implicit variables.

[16] A Russellian (truth-conditional) version of the definite operator of this type is: $\mathbf{the}'_{(et)(et)} \overset{\text{def}}{=} \lambda A_{et}.\lambda x_e.|A| = 1 \wedge A(x)$. I use here the Strawsonian operator for sake of compatability with Jacobson's analysis.

[17] The only material difference between the two analyses is that the SF variable corresponds to the indefinite article in Figure 3, and to a null element in Figure 4. In fact, in Winter (2001) it is argued that SF variables correspond to null elements in both cases, and the deviation from this treatment is here only for the sake of simplicity.

[18] Post-copular nominals such as *his mother and Brigitte Bardot* are treated using the ACOND rule. Suppose (for the sake of presentation alone) that *and* denotes here the i-sum operator of Link (1983). This operator, of type $e(ee)$, composes using ACOND with the e-type denotation of *Brigitte Bardot* and the $(e \to e)$-type denotation of *his mother,* to derive another $(e \to e)$-type function that maps each man to the i-sum of his mother and Brigitte Bardot. A similar analysis can be derived within the more complex treatment of plurals and predication that is proposed in Winter (2001, ch. 4).

[19] In fact, as argued in Winter (2001, ch. 4), it may be advantageous to treat copulas as meaningless rather than polymorphic, but whether this alternative is adopted or not is irrelevant for our present purposes.

[20] As an illustration for why this is justified, consider a modified restricting predicate of the form *A that Q R,* where A is the denotation of a noun (e.g. *woman*). Our analysis of this construction in Section 3 is $\mathrm{N}^0(A) \cap \mathrm{RG}(F_{QR})$, which is equal to $\mathrm{RG}(\mathrm{N}(A) \cap F_{QR}) = \mathrm{RG}(F_{QR'})$, of the form of the analysis, where $R' = \lambda x.\lambda y.R(x)(y) \wedge A(x)$.

[21] Thijsse (1983), who introduced this notion into generalized quantifier theory, gave it the misnomer *continuity,* which was used in some subsequent works.

[22] I believe that a similar reasoning underlies the semantic account in Barwise and Cooper (1981) of the grammaticality of *th*ere sentences with various NPs, using a non-triviality assumption.

[23] This is not always the case with *almost* quantifiers. For instance, a sentence such as *every Frenchman admires almost every actress,* is compatible with a situation where one Frenchman admires every actress but Brigitte Bardot, while another Frenchman admires every actress but Isabelle Adjani.

[24] Sharvit also judges as felicitous sentences like (57) with *more than two* and *at most two* instead of *most of the*. However, my Hebrew informants disagree, and consider these cases significantly worse than the Hebrew parallel of (57).

References

Alexopoulou, T. and Heycock, C. (2002) 'Relative clauses with quantifiers and definiteness'. In: *Proceedings of the Workshop 'Choice Functions and Natural Language Semantics'*. Arbeitspapier 110. FB Sprachwissenschaft. Universität Konstanz. Can be downloaded from http://ling.uni-konstanz.de/pages/home/heusinger/publications/ftp/ap110/AP110.Alexopoulou.pdf.

Barwise, J. and Cooper, R. (1981) 'Generalized quantifiers and natural language'. *Linguistics and Philosophy,* 4, pp. 159–219.

Benthem, J. Van (1991) *Language in Action: categories, lambdas and dynamic logic.* Amsterdam: North-Holland.

Chierchia, G. (1993) 'Questions with quantifiers'. *Natural Language Semantics,* 1, pp. 181–234.

Chierchia, G. (2001) 'A puzzle about indefinites'. In Cecchetto G. C. C. and Guasti M. T. (eds.), *Semantic Interfaces: Reference, Anaphora and Aspect.* CSLI Publications, Stanford.

Dahl, O. (1981) 'Binding relations in dislocated constituents'. Talk presented at the Sloan Workshop on Alternatives to Transformational Grammar, Stanford University.

Doron, E. (1982) 'The syntax and semantics of resumptive pronouns'. *Texas Linguistic Forum* 19, University of Texas, Austin.

Engdahl, E. (1986) *Constituent Questions: the syntax and semantics of questions with special reference to Swedish.* Dordrecht: D. Reidel.

Fodor, J. D. and Sag I. (1982), 'Referential and quantificational indefinites'. *Linguistics and Philosophy* 5, 355–398.

Groenendijk, J. and Stokhof M. (1984), 'Studies on the Semantics of Questions and the Pragmatics of Answers'. Ph.D. thesis, University of Amsterdam.

Hintikka, J. (1986) 'The semantics of *a certain'*. *Linguistic Inquiry,* 17, pp. 331–336.

Jacobson, P. (1994) 'Binding connectivity in copular sentences'. In *Proceedings of Semantics and Linguistic Theory, SALT4.*

Jacobson, P. (1999) 'Towards a variable-free semantics'. *Linguistics and Philosophy* 22, pp. 117–185.

Jacobson, P. (2000) 'Paycheck pronouns, Bach-Peters sentences, and variable-free semantics'. *Natural Language Semantics,* 8, pp. 77–155.

Jacobson, P. (2002) 'Direct compositionality and Variable-Free Semantics: the case of binding into heads'. In *Proceedings of Semantics and Linguistic Theory, SALT12.*

Kratzer, A. (1998) 'Scope or pseudoscope? Are there wide scope indefinites?'. In Rothstein S. (ed.), *Events and Grammar.* Kluwer, Dordrecht.

Link, G. (1983) 'The logical analysis of plurals and mass terms: a lattice theoretical approach'. In Bauerle R., Schwarze C and von Stechow A. (eds.), *Meaning, Use and Interpretation of Language.* De Gruyter, Berlin.

Peregrin, J. and von Heusinger, K. (2003) 'Dynamic semantics with choice functions'. In Kamp H. and Partee B. (eds.), *Context Dependence in the Analysis of Linguistic Meaning.* Elsevier, Amsterdam.

Reinhart, T. (1990) 'Self representation'. Paper presented at the Princeton Conference on Anaphora, Princeton.

Reinhart, T. (1992) 'Wh-in-situ: an apparent paradox'. In *Proceedings of the 8th Amsterdam Colloquium.* University of Amsterdam, Institute for Logic, Language and Computation.

Reinhart, T. (1997) 'Quantifier scope: how labor is divided between QR and choice functions'. *Linguistics and Philosophy*, 20, 335–397.

Ruys, E. G. (1992) 'The Scope of Indefinites'. Ph.D. thesis, Utrecht University.

Schlenker, P. (1998) 'A Note on Skolem functions and the scope of indefinites'. Unpublished ms., MIT.

Schwartz, B. (2001) 'Two kinds of long-distance indefinites'. In *Proceedings of the 13th Amsterdam Colloquium*. University of Amsterdam, Institute for Logic, Language and Computation.

Sharvit, Y. (1999) 'Functional relative clauses'. *Linguistics and Philosophy*, 22, pp. 447–478.

Steedman, M. (1999) 'Alternating quantifier scope in CCG'. In *Proceedings of the 37th annual meeting of the Association for Computational Linguistics (ACL)*.

Thijsse, E. (1983) 'On some proposed universals of natural language'. In ter Meulen A. (ed.), *Studies in Modeltheoretic Semantics*. Foris, Dordrecht.

Winter, Y. (1997) 'Choice functions and the scopal semantics of indefinites'. *Linguistics and Philosophy,* 20, pp. 399–467.

Winter, Y. (1998) 'Flexible Boolean Semantics: coordination, plurality and scope in natural language', Ph.D. thesis, Utrecht University.

Winter, Y. (2000) 'DP structure and flexible semantics'. In *Proceedings of the 30th conference of the Northeast Linguistic Society, NELS30*.

Winter, Y. (2001) *Flexibility Principles in Boolean Semantics: coordination, plurality and scope in natural language*. MIT Press, Cambridge, Massachusetts.

The Pragmatic Dimension of Indefinites

PAUL DEKKER[1]
*ILLC/Department of Philosophy, University of Amsterdam, Nieuwe Doelenstraat 15, NL 1012
CP, Amsterdam (E-mail: dekker@hum.uva.nl)*

Abstract. This paper sets out to give a natural pragmatic explanation of several aspects of
the interpretation of singular indefinite noun phrases. We develop a uniform account of
characteristic features of their use which have been dealt with only partly in other semantic
paradigms (in particular the dynamic, the E-type and the choice function one). We give an
intuitive motivation for the familiar discourse dynamic features of the use of these expres-
sions, and, taking due account of the structuring of information in more involved contexts,
account for their behaviour in negated, conditional, quantified, and intensional construc-
tions.

Key words: anaphora, information structure, pragmatics, scope islands, semantics, Skolem
functions

1. Introduction

It is generally agreed upon that pronouns and other definite terms may be
coreferential with antecedent indefinite noun phrases. There is no general
consensus, however, about how this is possible. If, as we have been taught by
Frege, Russell and Montague, among many others, indefinite noun phrases
behave like existentially quantified terms, then these phrases may be
'denoting' expressions, but they are not referential. So how can a pronoun
then be coreferential? In the modern tradition this issue has first been dis-
cussed by Peter Geach in the early sixties.[2] Geach in a sense did away with
the problem, assuming that, whereas an indefinite noun phrase indeed is like
an existential quantifier, the associated quantifier is taken to bind the pro-
noun. According to (Geach, 1968), in a sentence like:

(1) Socrates owned a dog, and it bit Socrates.

the conjunctive "and" is not the main operator, but the restrictive term "a
dog" and the sentence, thus analyzed, states that there is some dog x such
that Socrates owned x and x bit Socrates.

Geach's analysis has been criticized on two points, which are closely re-
lated but which have given rise to two different families of alternatives.

Gareth Evans has forcefully argued against the idea that pronouns, like the one we find in example (1), are like bound variables. Rather, they are referential expressions and they should be assigned a reference of their own (Evans, 1977). Thus:

(2) John owns some sheep and Harry vaccinates them.

should not be taken to state that there are some sheep which John owns and Harry vaccinates, rather, it states that, first, Johns owns some sheep, and, second, that Harry vaccinates (all) the sheep that John owns. According to Evans, the pronouns in (1) and in quite a few other examples (which he called 'E-type pronouns') really are disguised definite descriptions, which must be interpreted by first reconstructing the description with the help of the context of use, and next interpreting the description in a more or less standard fashion. Evans's work, together with that of Cooper (1979), has inspired a lot of work, notably in the early nineties, which tries to pin down what descriptions pronouns ought to be associated with, and how these in their turn ought to be interpreted.

Another research tradition has agreed with the E-type tradition that example (1) is a plain conjunction, but it rejects the idea that pronouns as in (1) are referential expressions. In various versions of discourse representation theory (Kamp and Reyle, 1993), file change semantics (Heim, 1982) and dynamic predicate logic (Groenendijk and Stokhof, 1991), the pronoun in (1) is taken to be like a variable, as assumed by Geach, but it is not directly bound. Syntactically it is a free variable, which eventually is bound, semantically, using either an intermediary process of discourse representation construction or of dynamic interpretation. These dynamic and discourse representational systems build upon the observation by Lauri Karttunen that indefinite noun phrases as in (1) in some sense introduce or set up discourse referents, which stand in for possible witnesses, and which can be subsequently picked up by pronouns (Karttunen, 1968).

One family of approaches shares features with both the E-type and the dynamic analysis of sentences like (1), the choice function approaches. [Early references are (Ballmer, 1978 p. 122ff and p. 307ff) and (Slater, 1986); see also papers from the Konstanz school (Egli and von Heusinger, 1995; Peregrin and von Heusinger, 2003) and (Meyer-Viol, 1995)]. The basic intuition of these approaches is that an indefinite term like "a dog" is associated with an ε-term, and that it denotes an individual chosen from the set of dogs. The pronoun "it" in (1) then can be taken to refer to the same dog. Example (1) is held true iff there is a choice function by means of which both conjuncts in (1) can be seen to be true, that is, a choice function Φ which selects an element from any non-empty set of individuals and such that $\Phi(dog)$ is a dog d such that Socrates owns d and d bit Socrates.[3]

Choice functions have also been used to deal with a local or in situ interpretation of indefinites, an issue which we will come back to in Section 6.

In this paper I argue for a specific way of understanding the dynamic semantic notion of interpretation, which, like the E-type and choice function approaches, is more consistent with classical, referentially based theories of meaning. This understanding of a dynamic semantics is based on ideas exposed earlier in (Kamp, 1990; van Rooy, 1997; Stalnaker, 1998; Zimmermann, 1999), and worked out in (Dekker, 2002a, more formally), (Dekker, 2002b more linguistically), (Dekker, 2003 more philosophically). With the E-type and choice function theorists, I agree that examples like (1) are ordinary conjunctions with a classical meaning. Like the choice function theorists, I assume that indefinite noun phrases and pronouns behave like terms, with an indefinite or variable interpretation which depends on the context of use. A pronoun's interpretation is functional on its occasions of use as it can be taken to refer to what is its referent on these occasions. It is assumed that also an indefinite noun phrase, like we find in the first conjunct of (1), is generally used with referential intentions. Thus, the pronoun which we find in the second conjunct can be taken to refer to whatever individual can be the intended referent associated with the indefinite noun phrase.

Without any need of shifting our notion of meaning, this can be accounted for by extending the semantic information expressed by the first conjunct of (1) with pragmatic information about what are the possibly intended referents associated with the uses of the indefinite. A dynamic semantics thus can be seen to be the result of a simple and systematic extension of a classical notion of meaning with quite general pragmatic features of the use of indefinites.[4]

I will proceed as follows. In the next section I will show in more detail how the essential features of anaphoric relationships with indefinite noun phrases can be accounted for on the basis of this pragmatic understanding of the use of these terms. In Section 3 I will sketch an equally pragmatic explanation of the 'fact' that indefinite noun phrases in certain configurations (for instance, in the scope of a negation) tend to lose their anaphoric potential, and I will briefly discuss the interpretation of donkey-conditionals. Then I will discuss two natural generalizations dealing, first, with the anaphoric potential of indefinites in 'derived' or dependent contexts (Section 4) and, second, with wide and so-called 'intermediate' interpretations of indefinites on scope islands (Section 5). Section 6 compares our approach to indefinites with the choice function approach and Section 7 summarizes the results. The paper concludes with an appendix specifying some of the main technical machinery underlying the main claims made in this paper.

2. Surface Indefinites

This section gives a concise review of the understanding of indefinites and pronouns as it is worked out in (Dekker, 2002a; Dekker 2003), and sketch the outlines of a formal pragmatic interpretation of them. (Technical details are provided in the appendix.) The main assumptions are the following:

- first, I want to stick to a classical (say, possible worlds) notion of meaning which can be specified as a satisfaction semantics
- second, conjunction amounts to intersection, basically, but when utterances are actually conjoined, we can incorporate the pragmatic fact that the first conjunct literally precedes the second
- third, I want to take to heart Stalnaker's observation that (surface) indefinite noun phrases can and generally are used with referential intentions
- fourth, I assume that anaphoric pronouns may pick up the individuals which are the intended referents of terms used earlier

I believe this to be a coherent set of ideas which can be used to motivate a formal toy system of interpretation *PLA* (Predicate Logic with Anaphora) which is a conservative extension of a classical system. The formal rendering of these assumptions may require some additional comments though.

The semantics of *PLA* is spelled out as a satisfaction relation $\models$ between, on the one hand, a model M, a variable assignment g, a world or possibility w and a sequence of individuals $\vec{e}$, and, on the other hand, a formula ϕ. The first three parameters are the usual ones, and the sequence of individuals parameter constitutes the extension over a classical system. These sequences provide the possible referents or witnesses of terms used in a formula, and they thus model what can be the targets of the referential intentions associated with these terms.

Satisfaction of conjoined utterances is defined, in the classical way, as the joint satisfaction of the conjoined conjuncts, but which accommodates the pragmatic fact that terms in the first conjunct are used before those in the second, so that we can account for the fact that pronouns in the second may refer back to the witnesses of terms used in the first.[5]

That an indefinite can be used with referential intentions is formally modeled by specifying what is a possible witness of that term. Formally, an individual d is said to be a witness satisfying $\exists x\phi$ iff ϕ gets satisfied if we map x to d. This account of the use of indefinites may need some comments and qualifications. First, notice that I do not model what are *actually* intended referents on certain occasions of use, but what are possible intended referents on ideal occasions of use. However, second, when this extensional system gets lifted into a calculus defining the support speakers can be required to have for their utterances, indefinites have to be linked up with individuals or rather

individual concepts they 'have in mind'. It would go too far here to fully explain what this 'having in mind' precisely means, but intuitively it means that the speaker must have some idea who it is about, even if it is as vague as for instance 'the individual whoever it is which somebody else must have intended to refer to when he told me this.' Thus, referential intentions do not need to allow the speaker to identify the referent in any contextually relevant sense, she is only required to have the idea that, eventually, it concerns some definite person, possibly via a causal intentional chain.

Third, the basic *PLA* system, which is basically a first order logic, does not actually model the referential intentions associated with full terms, but only that of the bare existential quantifier (something). However, later in the paper, and more specifically in the appendix, a more fine-grained interpretation is presented in which the contribution of indefinites is separated from what is said about their possible referents. Fourth, such a more fine-grained analysis will also allow us to clarify the pragmatic differences between the use of definite and indefinite noun phrases. Both types of terms are taken to be used with referential intentions. However, the first (and not the second) are of a presuppositional nature, and come with the assumption that the hearer should be able to identify the witness or referent; indefinites, on the other hand, typically do not allow the hearer to do so, and do not give him any further clue than that their reference should be the individual which the speaker should have been intended to refer to when he used the indefinite.

As for the fourth assumption, it proves very useful to formally distinguish anaphoric pronouns from bound variables, so they are introduced as a separate category in the language of *PLA*. Furthermore, in order to make them unambiguous they carry indices. So, apart from constants and variables, our language contains a set of pronouns $p_1, p_2, \ldots$ as terms, and basically such a pronoun p_i will be interpreted as the intended referent of a specific indefinite, viz., the i-th indefinite found when looking back in the discourse.[6]

Building on the stated assumptions, it is now fairly easy to show how indefinites and pronouns are dealt with in the basic system of *PLA*. Apart from the satisfaction of atomic formulas, which is straightforward, the two main clauses deal with the existential quantifier, modeling indefinites in a rudimentary way, and conjunction:

- $M, g, w, d\vec{e} \models \exists x \phi$ iff $M, g[x/d], w, \vec{e} \models \phi$

  $M, g, w, \vec{c}\vec{e} \models \phi \wedge \psi$ iff $M, g, w, \vec{e} \models \phi$ and $M, g, w, \vec{c}\vec{e} \models \psi$

A sequence $d\vec{e}$, with d its first element, satisfies $\exists x \phi$ relative to g iff the sequence $\vec{e}$ satisfies ϕ relative to the assignment $g[x/d]$ which assigns the witness d to x. A sequence $\vec{c}\vec{e}$ satisfies a conjunction $\phi \wedge \psi$ iff $\vec{c}$ are witnesses for indefinites in ψ, and $\vec{e}$ witnesses for pronouns in ψ, which have been introduced earlier, in ϕ, or even earlier.

It is easily established that the following two formulas turn out equivalent:

(3) $\exists x(D(x) \wedge O(s,x)) \wedge B(\mathrm{p}_1, s)$
(4) $\exists x(D(x) \wedge O(s,x) \wedge B(x,s))$

Actually, we can take the first formula to be the natural first order rendering of our example (1) above, and from the equivalence with the second we can see it gets interpreted correctly. An individual satisfies formula (3) if it is a dog which Socrates owns, and if it, that same dog, bit Socrates. Example (4) has the same satisfaction conditions.

Example (2) can be dealt with in basically the same way, if we make sure that a witness for the plural noun phrase is not an individual sheep, but the set of sheep John owns. (Similar witness set constructions can and have to be made for structures with generalized quantifiers.) It should be pointed out, however, that plural indefinites can also be used with specific referential intentions, cf., e.g., (Kamp and Reyle, 1993). We will not go into the interpretation of plural noun phrases here though.

Before we inspect more fancy constructions in the subsequent sections it is useful to agree upon some further terminology. Conforming to quite general practice, I will refer to satisfying individuals as 'witnesses', to satisfying sequences of individuals as 'cases', and to sets of pairs consisting of a world and a case as 'information states'. Actually, such states can be taken to present the combined semantic and pragmatic 'content' of our formulas. Relative to a model and a variable assignment the content of a formula is simply the set of world-sequence pairs which satisfy that formula.[7]

The contents of two formulas can be merged in a sophisticated fashion whereby anaphoric pronouns may get resolved by a previous indefinite noun phrase. I will not go into the details of that here [but cf., e.g., (Dekker, 2002a)], but a similar notion of merging can be used to define two other pragmatically crucial notions: the 'update' which a hearer may get from accepting an utterance, and the 'support' which a speaker can be required to have for it. We can think of the information states of a hearer and a speaker also as sets of world-sequence pairs. For a hearer such a state may serve to embody the information about 'discourse referents' which he has obtained from previous discourse. For a speaker such a state is required to embody the information about discourse referents which she herself introduces, and which she is supposed to associate with some defining characterization. A more detailed exposition of these notions can also be found in the appendix.

The ensuing notions of update and support have some nice formal features. As is shown in detail in (Dekker, 2002a), satisfaction, content, update and support are interdefinable, so this gives us the pleasant theoretical freedom to take any one of these notions as basic. Moreover, properly

resolved and supported updates never generate information which the involved agents together did not already have before the update. This is a highly desirable soundness result, which means, basically, that information does not get corrupted in an exchange. If speaker and hearer have true information, and they exchange some of it, their information is still true.[8]

There is one final point about the support relation which will become relevant later in the paper. For a speaker's state to support a formula it must be true in all the possible worlds in that state, in all possibilities which the speaker has not excluded as not being actual. But there is more. Like I said, discourse referents introduced by a speaker must themselves be associated with specific 'subjects' of her state, representatives of individuals which she believes to be uniquely specifiable. To model this I use a linking relation, which associates discourse referents with the speaker's subjects. In this way we can require a speaker to have support for an assertion in the sense that the things she attributes to discourse referents are really properties of the associated subjects of her information state.[9]

3. Background Indefinites

So far I only discussed what I called 'surface' indefinites, indefinite noun phrases which do not occur in the scope of other operators, like a negation, an implication, or, if we extend the language, quantifiers and mood indicators other than the indicative. It seems to belong to the received wisdom, however, that non-surface indefinites have a limited anaphoric potential, or even none. Consider:

(5) Onno doesn't run a sushibar.
(6) Is there a doctor in the audience?
(7) Give me a screwdriver, please.

In the first example [adapted from (Kamp and Reyle, 1993)] the indefinite noun phrase "a sushibar" is in the scope of a negation, and somehow it does not seem to license any anaphoric pick up. That is, if one were to continue with "It's in Soho." our responder would most probably ask something like: "What? What's in Soho? What are you talking about?" From my point of view this means that in a regular utterance of (5), or upon a regular understanding of it, no referential intentions are associated with the indefinite noun phrase. Similarly, asking (6) does not seem to make much sense if one has a particular doctor in mind and if one wants to know whether he is in the audience. (Think how odd it would be if somebody stood up and responded: "Do you mean me? You want to know if I am in the audience?") Furthermore, as a request for a very particular screwdriver uttering (7) would be quite beside the point. One can not, indeed, comply with a request if it is not (fully) specified.

So it seems, quite generally, that certain non-surface indefinites come without referential intentions, and in the case of (6) and (7) indeed some partial pragmatic explanation can be given for this fact. For a question about or a request for a particular thing to make sense the thing itself must be specified. Still, the question remains, if indefinite noun phrases are generally associated with referential intentions then why should these at all vanish in certain constructions? And do they, really? Let us consider some more examples:

(8) If Merl throws a party tonight, I'll be there!
(9) Many boy scouts who keep a pet develop into animal liberators later.
(10) If a client comes in, I'll give her a folder.

There is clearly something odd about continuing (8) with "It starts at 21.00." What could be supposed to start at 21.00? Not Merl's party, since (8) at least implicates that maybe there is not going to be such a thing. But a speaker may have something special in mind with asserting such a sentence. For it is really natural to continue (8) with "It will be fun!" If we then ask what is going to be fun, the straightforward answer, of course, is "The party, if any, which Merl is going to give." It is a hypothetical entity, but it is very clearly circumscribed.

Example (9) pretty much resists an interpretation which relates to a particular pet. Singular pets are generally not kept by any great number of boy scouts. Intuitively, such a sentence may serve to assert something about boy scouts, about pet-owning boy scouts, or about the relation between boy scouts and pets owned. The last construal is interesting, since it licenses a continuation with for instance "They then feel sorry about the way they treated it." The pronoun "it" then refers, not necessarily to a particular pet, but, for each boy scout, to the pet he is related to, i.e., the one he kept.

Examples like (10) are of the notorious 'donkey sentence' kind. In such a conditional assertion the indefinite "a client" in the antecedent clause is picked up by a pronoun in the consequent clause. But is this, therefore, about a particular client? Obviously not, since it would be quite odd again to try and pick up the indefinite later with: "She is from Oklahoma." Nevertheless, the sentence can again be read as being about clients in general. "What do you do when a client comes in?" "I'll give her a folder."

The above observations, which are not at all new, suggest the following generalization. Whereas non-surface indefinites can (if needed) relate to specific individuals (some more examples which have popped up in the literature are discussed in Section 4), also when they do not relate to specific individuals, they can relate to classes of individuals, which the assertions can be conceived of as being about. That is to say, these assertions can be assigned a so-called information structure, part of which is a ground, which in a sense can be assumed to be given, and another part which can be called its

focus. Indefinites in the ground are not associated with referential intentions then, because they are not part of the speaker's own contribution (which is laid down in the focus part) but part of the assumed given ground.

Consider again a statement made by means of (5). Typically, but not inflexibly, a negation "Not S" may serve to answer the issue – raised explicitly or implicitly – whether "S" is true or not. An utterance of (5) may serve to state – possibly in answer to the question whether Onno runs a sushibar – that he doesn't, that is, that there is no such bar which Onno runs.[10] A speaker need, in other words, not have a particular sushibar in mind when uttering (5), because the existence of such a sushibar is not part of what she claims to have evidence for. Rather, the existence of such a bar is part of the issue which the speaker addresses–negatively, with (5) – or even part of what the hearer might have claimed just before. So actually, when somebody utters (5), she is normally not coming up with a sushibar herself, but she is claiming to have evidence against the existence of such a bar, were anybody else thinking of the possibility of there being one. And although the indefinite is clearly part of the string of words which the speaker utters, it is not part of what she asserts, or of what she can be required to (be able to) support.

But indeed an utterance of example (5) may have other interpretations than the mere negation of a proposition, much dependent on the way it is uttered, and the context in which it is used. Like I said, on a most regular interpretation such an assertion is about Onno (being presupposed) and it rejects the existence of a sushibar he runs. A truth-conditionally equivalent, but pragmatically different, construal takes the assertion to be about sushibars in general, and to reject that Onno owns one. Such an analysis seems to be appropriate if an utterance of (5) is followed by the subsequent assertion of "They are all run by non-residents." Truth-conditionally different is a construal which takes the assertion of (5) to be about a specific sushibar after all. An interpretation like this seems to be appropriate if the sushibar is picked up again, as in the following extension of (5):

(5) Onno doesn't run a sushibar.
(11) He only does its financial administration.

Finally, an odd, but certainly not impossible construal has it that there is no Onno owning a sushibar, one that allows for the possibility that Onno does not even exist.

As is shown in some detail in the appendix, this variety of readings is elegantly accounted for in a multi-dimensional interpretation architecture. The nice thing about such a multi-dimensional set up is that the variety of interpretations is obtained without postulating a semantic ambiguity in the negative element. The different readings emerge from the different pragmatic ways in which that element can be taken to act upon the information structure of the embedded clause. (See the appendix for more details.)

A flexible, multi-dimensional architecture is also very well equipped to deal with the previously mentioned donkey sentences in a principled way. Consider the following examples:

(12) If a farmer leases a donkey he beats it.
(13) A farmer beats a donkey only if he leases it.
(14) Only if a farmer leases a donkey does he beat it.
(15) A farmer beats a donkey if he leases it.

Example (12) is modeled after Geach's donkey sentences. It has constituted a major problem for standard theories of interpretation, because, in order for the indefinites in the antecedent of this conditional to 'bind' the pronouns, the indefinites need to gain wide scope. However, they will not, thus, gain the universal force which the indefinites in (12) arguably have. E-type pronoun theories, discourse representation theory and systems of dynamic semantics have given a neat and well-known account of this example, and most probably this account extends to example (13). However, these accounts terribly fail in the face of seemingly similar examples like (14) and (15). The reason is that the *only if*-clause in (14) is semantically (conditionally) dependent on the main clause, whereas the pronouns in the main clause are structurally (anaphorically) dependent on the *only if*-clause, and on the mentioned accounts this creates a paradox of interpretation. The same goes for example (15). [See, e.g., (von Fintel, 1994; Dekker, 2001b) for more discussion].

Much of the mystery around (12–15) is cleared, however, if we acknowledge that indefinites may serve what can be labeled a 'topical' role. Their possible generic interpretation has first been discussed in, e.g., (Schubert and Pelletier, 1989), and more recently a 'presuppositional' or 'non-novel' use of them has been discussed in (Gawron, 1996; Aloni et al., 1999; Krifka, 2001). We can account for this if indefinites are assigned a special *use* (not: *meaning*) in such conditional (and quantified) contexts. For, intuitively, the examples (12–15) are about farmers and donkeys and each of them can be used to state a conditional dependency between a leasing and a beating relation between them. More precisely, the indefinites "a farmer" and "a donkey" may contribute to establish a domain of quantification consisting of pairs of farmers x and donkeys y, and on that domain of pairs example (12) can be used to state that if any such x leases any such y, then x beats y, example (13) that any such x beats any such y only if x leases y, example (14) that only if any such x leases any such y does x beat y, and example (15) that any such x beats any such y if x leases y.

I hope that this suffices to make clear how I think these sentences should be understood. Formally, such an interpretation can be obtained if we assume a partition of the contents of assertions into a ground and a focal part, a distinction familiar from many different types of theories of information

structure, like that of, e.g., (Jackendoff, 1972; van der Sandt, 1989; von Stechow, 1991; Vallduví, 1992). As can be seen in more detail in the appendix, the contents of all expression can be distributed over various dimensions of interpretation and indefinites can be taken to contribute to a dimension of their own. Possible witnesses then can be used to 'communicate' between the various dimensions. We thus can say that, for instance, a pair of individuals *bd satisfies the ground* of "a farmer leases a donkey" iff *b* is a farmer and *d* a donkey, and that such a pair *satisfies its focus* iff *b* leases *d*. The examples (12–15) then are adequately dealt with if they are taken to quantify (universally, or generically) over the (tuples of) individuals which satisfy the ground of the embedded clauses, and assert that all of them satisfy the asserted conditional dependency.

Let us take stock at this point. Typically – that is, if context or intonation do not imply otherwise – one can say that e.g., the contents of negated sentences, of questions and commands, but also the antecedents of conditional sentences, and the restrictions on quantifiers, all constitute or relate to a background or topic. Topics are assumed to be given in some sense and they do not belong or contribute to the conversational commitments which a participant takes upon herself when making a certain statement. A speaker then can be taken to have support for what the focal part of her utterance contributes to the ground, but not to the ground itself. Hence, she is not required to have any referential commitments associated with indefinites used in the ground, and indeed these indefinites may be topical (or 'non-novel') themselves.[11]

4. Dependent Indefinites

So far I have given an idea of how surface indefinites get associated with referential intentions and why some non-surface indefinite noun phrases do not. But of course this is still only part of the story. I already mentioned indefinites which figure in the focal part of an assertion. If a speaker can be held conversationally responsible for what the focal part of her utterance contributes to a ground, then we would indeed expect indefinites there to be associated with referential intentions again. Interestingly, this seems to be precisely what we find in a couple of examples familiar from the literature. And although these examples are arguably somewhat marginal, something which eventually has to be explained, I think they provide strong support for the pragmatic kind of analysis I am pursuing in this paper. Before I turn to the relevant examples it is useful to inspect the notion of implication which naturally suggests itself in the system of *PLA*.

If an implication $\phi \rightarrow \psi$ is defined, in a fairly usual way, as $\neg(\phi \wedge \neg\psi)$, it turns out that support for stating such an implication boils down to having support for ψ after one has updated one's information with ϕ. One could

reformulate this as follows. I have epistemic support for *If A then B* if, and only if, if you were to tell me that *A*, or if I find out otherwise that *A*, then I have sufficient evidence for *B*. This sounds pretty fair, and close to the interpretation of conditional sentences in systems of game theoretical semantics. But notice that an utterance of *If A then B*, thus, does not commit a speaker to having support for *B*, but only for *B* in functional dependence upon (learning that) *A*. So if there are referential intentions to be associated with indefinites in *B*, we can assume them to be functionally dependent upon whatever is contributed by the ground in *A*.

Conditional statements thus can be interpreted as supported comments upon the types of situation provided by their antecedent clauses, which they are dependent upon. Likewise, many quantified constructions (in particular upward monotonic ones) can be interpreted as qualified comments on a domain set up (or presupposed) by the clauses restricting the main determiner. And then it can be expected, again, that a speaker has qualified (i.e., functional) support for terms used in the determiners' nuclear scope. Actually, such functional readings of noun phrases (*Wh*-phrases, definite and indefinite noun phrases) are familiar from the literature from the eighties and the nineties. [See (Jacobson, 1999) for a recent overview, or (Alexopoulou and Heycock, 2003)]. Typical examples include:

(16) Whom does every Englishman admire? His mother.
(17) Every Englishman loves, but no man wants to marry his mother.

The question in (16) is about, and relative to, a domain consisting of Englishmen. Although it is most likely to be interpreted as asking for the individual x which is such that every Englishman likes x, the question also allows for a functional reading: what is a or the function f such that every Englishman e likes $f(e)$? As appears from the continuation in (16), a felicitous answer can be the mother-function, assigning to every Englishman e e's mother. Similarly, a functional interpretation of the definite noun phrase in (17) seems to be most appropriate.

Also indefinites and pronouns may license a functional interpretation in these configurations:

(18) Every Englishman loves some woman, but no one wants to marry her.

The indefinite in this example may relate to a particular woman which every Englishman loves, but which no one wants to marry. However, it can also be interpreted functionally, in case it yields a reading like that of (17). The difference is, of course, that an utterance of (18) does not specify which woman or woman-function it is about, and a speaker might continue the example with "But I forgot whether it is his mother, his grandmother, or his oldest sister."
The following examples display an essentially similar pattern:

(19) If a book is printed with Kluwer it has an index. It can always be found at the end. (after Heim)

(20) Harvey courts a girl at every convention. She always comes to the banquet with him. (Karttunen)

(21) Most men had a gun, but only a few used it. (Sandu)

(22) Mary believes there is a burglar in the house. She thinks he came in through the chimney. (Landman)

In Heim's example we find an indefinite noun phrase "an index" in the consequent clause of a conditional sentence. If it is associated with referential intentions, the speaker can be assumed to have a specific index in mind (like 'the index of a book printed with Kluwer') which, however, is not one particular index, but functional upon Kluwer books. Assuming the speaker to have such a function in mind, we can also assume that it is that function which is picked up by the pronoun. That is, the second sentence of Heim's example can be taken to state that always (that is, in every Kluwer book) the index of that book can be found at the end of the book.

Karttunen's 'girl' can be thought of as a function associating every convention which Harvey visits with a girl he courts there. Sandu's 'gun' is appropriately associated with a function from gun owning men to their guns and Landman's burglar with a function from the alternatives which Mary believes to be possible to burglars which are there in the house.[12]

The above observations are precisely those expected on an analysis of indefinites like the one defended here. Focal indefinites are associated with referential intentions, and since they are functionally dependent on some ground, so are the entities which the speaker can be said or required to have in mind. That is to say, my notion of support and the way in which it is supposed to function, naturally suggests an analysis of these examples. Basically, they can be dealt with by generalizing our notion of a case, which are not only sequences of satisfying individuals, but also of witness functions from world and cases to individuals.[13] Satisfaction and support for the various operators (conditional, quantified, epistemic) then makes the witnesses for their embedded clauses parametric upon the indices they quantify over. (Basically, this is nothing but a suitable generalization of Geach's so-called rule of 'division'.) The net effect is that they generate Skolem equivalences of the following form, and they do this in an entirely compositional manner:

(23) $\exists x \phi(x) \rightarrow \exists y \psi(y) \Leftrightarrow \exists f (\exists x \phi(x) \rightarrow \psi(f(\mathbf{p}_1)))$

(24) $\forall x \exists y \phi(x, y) \Leftrightarrow \exists f \forall x \phi(x, f(x))$

(25) $B_x \exists y \phi(y) \Leftrightarrow \exists z B_x \phi(^{\vee}z)$

Satisfaction of an implication of the form $\exists x \phi(x) \rightarrow \exists y \psi(y)$ requires a witness function f, which applies to the type of entities which the antecedent

requires to be ϕ, and which can be picked up by the pronoun p_1. By the set up of the system of interpretation the implication can be followed by another one in which an anaphoric pronoun picks up this function. Notice that for this to work, the second pronoun must be functionally dependent upon the same type of things which the original indefinite is functionally dependent upon. Thus, in order to effectively deal with Heim's example, it must be made sure that the adverbial quantifier "always" relates to books printed with Kluwer.

Something similar holds of the seemingly regular Skolem equivalence (24). The difference with an ordinary Skolem equivalence is that the use of f is associated with referential intentions, so that it is available for anaphoric take up. This gives us a handle on Karttunen's example, if, again, the quantifier "always" is made to range over the right types of things. In the equivalence (25) I have suggestively used Montague's extension operator $^\vee$ (which is not actually part of my own formal language).

On the basis of the above equivalences we can account for the examples (19–22). For instance, an appropriate interpretation of (1000) satisfies the following sequence of equations:

$$(26)\ B_x \exists y \phi(y) \wedge B_x \psi(p_1) \Leftrightarrow \exists z B_x \phi(^\vee z) \wedge B_x \psi(p_1) \Leftrightarrow$$
$$\exists z (B_x \phi(^{vee}z) \wedge B_x \psi(^\vee z)) \Leftrightarrow \exists z B_x (\phi(^\vee z) \wedge \psi(^\vee z)) \Leftrightarrow B_x \exists y (\phi(y) \wedge \psi(y))$$

The first and the last equivalence in this sequence corresponds to the Skolem one in (25). The second serves to display the effect of anaphoric take up. The third is part of the logic of belief. The other examples can be dealt with analogously. (See the appendix for some more details).

The preceding discussion may also serve to answer the question which I raised earlier in this section. Although the general idea behind the support of indefinites indeed serves to predict the type of functional dependencies discussed here, such dependencies are hardly felt to be there in the great majority of conditionals and quantified statements. Why should that be? Part of the answer is this. For a functional indefinite to be picked up by a subsequent pronoun, it is absolutely necessary that the pronoun is evaluated relative to precisely the same ground as the indefinite is. For, in short, functions require arguments, and these must be of the right type. In a lot of discourse and dialogue, however, contexts seem to change so quickly and subtly that in many cases anaphoric pick up of dependent indefinites is impossible, and functional readings are therefore invisible. Besides, of course, individual concepts and Skolem functions do not really belong to the most familiar things which linguistics agents deal with, so this will certainly be a further reason why the functional interpretation of indefinite noun phrases and pronouns does not belong to our most basic linguistic skills.

5. 'Intermediate' Indefinites

We have seen, on the one hand, that surface indefinites are generally used with referential intentions, which can be functional if they figure in the scope of quantified constructions, and that they may acquire a topical nature when they figure in a ground. However, indefinites can also be used with referential intentions when they are not in, say, focal position. In this section we discuss two types of examples which have puzzled logicians and linguists alike, which have given rise to non-standard systems of interpretation, but which naturally fit in the pragmatic outlook on indefinites argued for in this paper.

The first type of example is due to Charles Sanders Peirce (1906), and is discussed in detail in (Dekker, 2001a):

(27) There is some married woman who will commit suicide in case her husband fails in business.

Peirce notes that on what we understand as a relatively straightforward predicate logical analysis the sentence would be equivalent with:

(28) Some married woman will commit suicide if all married men fail in business.[14]

Most people judge that an utterance of (27) conveys something stronger and more specific than an utterance of (28), however. Peirce puts the blame for this "absurd result" on "admitting no reality but existence," and his diagnosis consists in taking *possible* courses of event into account. What is really meant by an assertion of (27), Peirce claims, is that "[t]here is some *one* married woman who under all possible conditions would commit suicide or else her husband would not have failed."

Interestingly, our pragmatic outlook upon the use of indefinite noun phrases gives us precisely this. For someone's information state to support an utterance of (27), and not for an utterance of (28), the speaker must have an individual in mind which, in all possibilities which the speaker conceives possible, commits suicide if her husband fails. Various pragmatic principles contribute to making such an utterance non-trivial only if the speaker indeed has a conception of a person about which she believes such a dependency to be true.

The point about Peirce's example in the present context is this. If the indefinite gives only existentially quantified information then indeed, as Peirce observes, example (27) is in danger of conveying nothing more than example (28). However, since it is assumed that such indefinites should be supported by subjects in the speaker's information state, and because what is predicated of them must be non-trivial in a fully Gricean sense, an utterance of (27) gets its special bite. About the person which the speaker has in mind the speaker may, normally, not have pertinent information that that person's

husband actually is, or is not, going to fail in business, or that that person will, or will not, commit suicide anyway.

As implicated by Peirce himself, and as Read has made explicit, a basically similar analysis of (27) and (28) can be obtained by reading ← as a strict (not material) implication. Upon such an analysis the overtones of (27) would be properly semantic, in stead of systematically pragmatic, as they are on our account. It is hard to decide between the two analyses, because the results are the same and each of them comes with its own independent motivation. Notice, however, that, first, our pragmatic analysis is consistent with a strict reading of conditional sentences, so it is not in conflict with any motivation for that. Second, our analysis directly carries over to non-conditional variants of Peirce's examples, which a strict conditional analysis does not. As Gillon observed, the examples (28–29) can be given a disjunctive formulation, as in:

(30) Someone wins $1,000 or he does not take part.
(31) Someone wins $1,000 or someone does not take part.

Indeed, one could explain the overtones of (30) by assuming an intensional analysis of the disjunction; however, no such move need be made if they are attributed to the very same pragmatic principles I have argued for above.[15]

Another intriguing type of examples are those with indefinites on so-called 'scope islands'. 'Scope islands' are, for instance, phrases restricting the scope of quantifiers, antecedents of conditional sentences and other subordinate clauses. It is one of the rather persistent observations from the formal linguistic canon that quantified noun phrases do not 'escape' from there, that is, that they are unable to outscope quantified, conditional, or superordinate constructions. Indefinite noun phrases, however, seem to do what the canon forbids quantified noun phrases to do. Constructions in which an indefinite noun phrases figures on a scope island can often be paraphrased, appropriately, with a formulation in which the indefinite does have wide scope. I will not go through all the motivating data, which go back to the seventies of the last century, but simply refer to the more recent literature on the subject found in, e.g., (Abusch, 1994; Reinhart, 1997; Winter, 1997; Kratzer, 1998), see also some other contributions to this volume.

The 'aberrant' behaviour of indefinite noun phrases on scope islands has been first explained in (Fodor and Sag, 1982), where it is argued that indefinites can have a quantified and a referential interpretation, and under the last they can obtain 'wide scope' without being outscoping. Indeed, such an interpretation neatly fits in with the one argued for here, but for the fact that I do not want to deem indefinites ambiguous, but simply attribute the referential interpretation to the pragmatic fact that the indefinite is used with referential intentions. This pragmatic account is advantageous, as it also

undercuts one of the main arguments against the Fodor and Sag analysis: the existence of so-called 'intermediate' readings.

Indefinites on scope islands have an intermediate interpretation when their contribution is not global (i.e., directly referential), but also not purely local (restricted to the scope island itself). A very clear example is from (Abusch, 1994), which seems to favour what is classified as such an intermediate reading:

(36) Every one of them moved to Stuttgart because a woman lived there.

The most natural interpretation of this example is that for every person referred to, there was a woman who lived in Stuttgart, and who made up the reason for that person to move to Stuttgart. Notice that the motivating women thus escape from the *because*-scope island, without thereby entailing it was one particular person who motivated the move of all of them. Each of them may have had his own woman motivating his move to Stuttgart. The indefinite, thus, is not purely referential, it is argued, but if the only other options is that it is, therefore, quantificational, and at the same time takes scope over the *because*-clause, it would violate the scope island constraint.

Like I said, the facts about the interpretation of indefinites on scope islands perfectly fit the picture sketched in this paper. Basically, example (36) can be interpreted in three ways, depending on how we understand the *use* (not *meaning*) of the indefinite, and these three interpretations are naturally predicted. First, of course, the indefinite can be used without any referential intentions, and then we obtain the reading that every one moved to Stuttgart because Stuttgart was not 100% male.[16] But it may as well be a pragmatic fact that the indefinite *is* used with referential intentions. The speaker may have had, for instance, Dorit Abusch herself in mind, and can be taken to claim that the reason for everybody to move was the fact that Dorit lived there.[17] Finally, as we have seen in the previous section, referential intentions associated with indefinites in quantified constructions can be functional as well, so that the speaker can be taken to claim that the reason for each of them to move was the fact that his or her fiancé(e) lived there.[18]

In our multi-dimensional framework the three interpretations of example (36) are naturally obtained, arguably without violating scope island constraints. The semantic denotation of the indefinite is its possible witness (or, better, the function from possible witnesses to the propositions expressed about them) and upon each of the three readings this witness or witness-function plays its properly semantic or assertoric role on the scope island. It generates an individual which is a woman. This woman, however, can be either arbitrary, or specific, or functionally related to the group referred to by "them," depending on the way in which the (non-assertoric) contribution of

the indefinite is pragmatically understood. The uniform semantics of the *because*-operator, and the meaning of "every one of them", can flexibly interact with this pragmatic contribution, and neglect or absorb it. But this process of composition is arguably pragmatic as well, and the *pragmatic* fact that the indefinite is used with referential (possibly functional) intentions does not conflict with the scope island constraint, which is structural (*syntactic and/or semantic*). (See, again, the appendix, for some of the formal details).

6. The Use of Choice Functions

In the introduction I have mentioned a family of approaches to indefinites which employ choice functions and which are close in spirit to the one advocated here. In this section I will discuss some of the data which have given rise to such choice function interpretations and I will argue that a pragmatic approach like the one presented in this paper is actually more economical, although my findings are, I believe, quite consistent with the pragmatic choice function approach advocated in (Kratzer, 1998). Most of my observations are not really new, though, and they are close to those of (Schlenker, 1998; Kamp and Bende-Farkas, 2001). See also (Winter, 2003).

The interpretation of scope island indefinites has provided strong support for a choice function analysis (cf., e.g., [Abusch, 1994; Reinhard, 1997; Winter, 1997; Kratzer, 1998; Matthewson, 1999]). The general idea is that indefinite noun phrases on scope islands deliver their semantic contribution locally, and, thus, do not violate island constraints. The semantic contribution of an indefinite "Some A" is an individual, which is the value of a choice function f which is applied to the set denoted by "A". Upon most approaches this choice function is existentially quantified, but the locus of quantification is generally assumed to be free.[19] Thus, for instance, example (36) can be associated with three interpretations:

- ($\exists f$) Every one of them ($\exists f$) moved to Stuttgart because ($\exists f$) f(woman) lived there.

where the first, second, and third locus of quantification generate the global, intermediate and local reading of the example, respectively. Choice functions thus generate the required number of readings arguably without violating island constraints.

This basic analysis, however, is or has to be amended in three ways, each one of which makes it more into an analysis of the kind advocated here. In the first place choice functions may have to be skolemized themselves (Winter, 2003). I will not discuss this point in detail here, as it has been extensively discussed elsewhere [cf., e.g., (Schlenker, 1998; Kamp and

Bende-Farkas, 2001)]. The point is that once one adopts the possibility of skolemization, the effects of intermediate existential closure can be captured by means of global closure, as we have seen in the previous section. [20]

Two other amendments can be advanced by means of an example which has originally been used to motivate a choice function approach in the first place.

The following example from Irene Heim has been seen to raise what Tanya Reinhart labeled the 'Donald Duck' problem:

(39) If we invite some philosopher, Max will be offended.

An utterance of this example may be used to convey that the speaker does not really know (in a contextually relevant sense) which specific philosopher it is whose possible invitation would offend Max. By uttering (39) the speaker may wish to convey that she wants to know who this possibly disputable philosopher is. By the same token, the utterance does not need to convey that if we invite any philosopher, Max will be offended. Max may go along well with a lot of philosophers. Upon a first analysis one might think that the informational contribution of the indefinite is indeed local, but that it is somehow 'existentially quantified from the outside': that there is some individual x, and that if x is some philosopher which we invite, then Max will be offended. This is where Donald Duck hits in. For, to make such a statement true, it suffices to choose Donald Duck as a witness for x. Since Donald Duck is supposed not to be a philosopher, a disputable assumption by the way, the statement would be trivially true. As a matter of fact, upon this analysis (39) would become equivalent with:

(40) If we invite everybody, Max will be offended.

(Assuming there to be at least one philosopher.) Clearly, (40) is not an appropriate paraphrase of (39).

The use of choice functions partly solves this problem. The idea is that (39) conveys that there is some choice function, and that if we invite the philosopher whom that choice function assigns to the set of philosophers, then Max will be offended. Indeed, Donald Duck would not be a proper witness any longer, since a choice function cannot pick him out of the set of philosophers.[21] Even so, an essentially similar problem remains. For let us now take Jacques Derrida: Derrida is a philosopher, but we are simply not going to invite him. Derrida thus is a good choice from the set of philosophers which will make an assertion of (39) trivially true. The triviality problem is only partly dealt with for if we analyze (39) this way it is rendered equivalent with:

(41) If we invite all philosophers, Max will be offended.

(Again assuming there to be at least one philosopher.) Obviously, this is still not an appropriate paraphrase, and the discussion of Peirce's example above may indicate what fails.

In order for an utterance of (39) to be felicitous, the speaker must have a philosopher in mind for which the conditional sentence is significant and not trivial. True, someone who sincerely asserts it may not know in some quite relevant sense which philosopher it is about but it seems it has to be about some definite philosopher. Someone the speaker has heard about from somebody else, or somebody whose name she forgot. In either case, it ought to be a philosopher, which the speaker heard or learned about, whose possible invitation would offend Max. For her utterance to be non-trivial, in all Gricean respects, she should not know of that philosopher whether or not (s)he is going to be invited, or whether or not Max will be offended anyway. The example shows, again, that there are pragmatic constraints on what are appropriate witnesses for the indefinite. But if these witnesses, or the choice functions which generate them, are existentially quantified, we don't have a handle on the entities (individuals, choice functions) onto which to apply these constraints. Indeed, for this reason the example really supplies motivation for the pragmatic (free variable) choice function approach like that of (Kratzer, 1998), or more simply, for the one argued for here.

It could be objected that the problem with example (39) lies not so much in the interpretation of the indefinite, but in the material interpretation of the conditional. Thus, one might say, the actual (global) interpretation of the indefinite should be that there is some philosopher such that on all 'relevant' possibilities, the invitation of that philosopher would offend Max. Notice, first, that this paraphrase comes close to the interpretation we actually get by pragmatic means. Notice, second, that, as in the case of Peirce's example, a strict reading of conditionals is consistent with our analysis. But, third, this would again be of no help when we consider a variant of (39) in disjunctive form:

(42) Either we invite some philosopher, or Max will be offended.

I think this sentence has a sensible reading as to which Max has his favourite philosopher (not Derrida) such that if we don't invite him, Max will be offended. Upon this reading not any arbitrary philosopher which we *do* invite suffices to make the assertion true or felicitous. It should be the philosopher whose non-invitation would upset Max. Fourth, and most importantly, a choice function analysis of the indefinite in (39) would definitely have to be amended in case we use a strict analysis of the conditional. For if we start looking at different possibilities, there are possibly different sets of philosophers, and the choice function might yield a different philosopher in each of these possibilities, one that is not even actually a philosopher.[22]

The latter problem can be solved by resorting to an intensional choice function analysis like one suggested by Reinhart and Winter. Essentially

the idea then is that choice functions apply to properties (not sets), and such that on the locus of existential closure choice functions assign to any property an individual which *actually* has that property. But here the intuitive appeal of the choice function analysis starts to break down almost completely. If this is the ultimate analysis, the local semantic contribution of an indefinite "Some A" is whatever a function f assigns to the property A. Only on the superordinate level of existential closure it is required that f is some function assigning to any property in any possibility an individual which *at that level of interpretation* actually has that property. I find it hard to see why this analysis should still comply with the scope island constraint. The most important ingredients of the interpretation of indefinites are lifted out of their islands. Indeed, upon our approach, and upon all the paraphrases of the non-local interpretation of scope island indefinites, the very same thing happens. But on our account this happens on the pragmatic level, different from the structural level of syntax/semantics where the scope island constraint applies. On an intensional choice function alternative, this substantial type of lifting takes place at the level of logical form.

Let me summarize the findings of this section. Special interpretations of scope island indefinites can be accounted for by assuming they structurally obey scope island constraints. They have access to a pragmatic dimension of use, which can be used to explain their apparent escapist behaviour, and independently motivated functional interpretations can be used to explain so-called intermediate readings. Alternative analyses adopting choice function readings will arguably have to be adapted in three ways. They must allow skolemization, context dependence, and intensionalization in order to work out proper in general. This use of choice functions simply complicates matters, instead of being explanatory.

7. Conclusions

In this paper I have presented a view upon the use and interpretation of indefinite noun phrase formally inspired by the dynamic paradigms of discourse representation theory, file change semantics, and dynamic predicate logic. Conceptually it has been inspired by (Stalnaker, 1998) and other philosophers who have given their thoughts to the use and interpretation of definite and indefinite noun phrases. Both types of phrases are used with referential intentions, the difference being that the intended referents of definites are required to be determinable in principle, whereas the use of an indefinite indicates that the identity or determination of the intended referent is not relevant. Truly Gricean notions of support have been called upon to explain a couple of basic facts about the use of these expressions,

such as the anaphoric potential, and, e.g., their behaviour in Peirce's example.

Not much more has been needed to analyze the use of indefinites in constructions dealt with in alternative approaches to indefinites. We need to adopt some appropriate notion of ground and focus in order to deal with information structure, but this is nothing new. We also need to allow terms to have functional readings, but this, too, is old wisdom. All of the data discussed in this paper can thus be dealt in a rather conservative manner.

I have started from a classical satisfaction semantics and the tools which I have used are basically, those of Tarski (satisfying sequences), Jackendoff (information structure), Geach (division) and Grice (pragmatics). The main challenge, and result, of this paper has been that of finding a proper formulation of the interaction of semantic and pragmatic information. First, in order to account for the referential intentions associated with the use of indefinites, by adding possible witnesses as an additional parameter of interpretation; second, in order to account for the structure of information, by the distribution of meaning over various dimensions of interpretation. Interestingly, by considering the contribution which indefinites make as part of what their use may pragmatically convey, we get a neat account of their behaviour on scope islands, which seems to be different from that of other noun phrases, but which actually is the same.

Empirically, I believe the present account does just as well as *DRT*, E-type pronoun approaches, or a choice function treatment, and possibly also the other way around. For instance, an intensionalized version of Kratzer's pragmatic choice function treatment can empirically be virtually indistinguishable from ours. If such indeed is the case, I favour the account presented in this paper, for the reason that it is more principled and less involved.

8. Appendix

In this appendix I present and illustrate the basics of the formal architecture underlying the main claims of this paper. I first present the system of *PLA*, *Predicate Logic with Anaphora*, in which quite systematic facts about the use of indefinite noun phrases and pronouns are appended to a classical satisfaction semantics. I then lift this semantics to an update and support calculus, which incorporates basics aspects of information exchange and it is shown that this set up natural asks for a form of quantified and modal parametrization by means of which functional dependencies get accounted for in a straightforward manner. I finally show how information structure is handled flexibly, assuming a distribution of meaning over several dimensions which get correlated by being evaluated relative to the same parameters of witnesses.

PREDICATE LOGIC WITH ANAPHORA

The language of *PLA* is that of ordinary predicate logic, except for the fact that it also contains a category of pronouns $p_1, p_2, \ldots$, which can be used as terms in atomic formulas. For ease of exposition I focus on a fragment built up from atomic formulas by means of negation $\neg$, existential quantification $\exists$ and conjunction $\wedge$.[23] A model for *PLA* can be an ordinary first order predicate logical model, but with a view upon later intensional (and epistemic) applications, I adopt Kripke models $M = \langle W, (R_i), D, I \rangle$ consisting of a set of possibilities W, a (possibly empty) family of accessibility relations R_i over W (modeling the beliefs of agents i), a domain of individuals D, and a possibility dependent interpretation of the (individual and relational) constants I. For the sake of simplicity I assume D to be the same in all worlds, but nothing hinges upon this assumption.

Before I can turn to the semantics of *PLA*, I have to define the 'length' and the 'reach' of a formula. The length $n(\phi)$ of ϕ is the number of individuals it introduces: the number of existentials (indefinites) not outscoped by a negation. The reach $r(\phi)$ of ϕ is the number of individuals it presupposes: the number of existentials (indefinites) which pronouns require there to be present in discourse preceding ϕ. The scope $s(\phi)$ equals the sum of the two.

- $n(Rt_1 \ldots t_m) = 0 \quad n(\neg\phi) = 0$
  $n(\exists x\phi) = n(\phi) + 1 \quad n(\phi \wedge \psi) = n(\phi) + n(\psi)$
- $r(Rt_1 \ldots t_m) = MAX\{j \mid p_j \text{ is among } t_1, \ldots, t_n\} \quad r(\neg\phi) = r(\phi)$
  $r(\phi \wedge \psi) = MAX\{r(\phi), (r(\psi) - n(\phi))\} \quad r(\exists x\phi) = r(\phi)$

If $n(\phi) = 0$, ϕ is called closed, and if $r(\phi) = 0$, it is called resolved.

The semantics of *PLA* is formulated as a satisfaction relation among, on the one hand, a formula ϕ, and, on the other, a model M, a variable assignment g, a world w, and a sequence of individuals $\vec{e}$. The sequences $\vec{e}$ consist of the possibly intended referents of terms in ϕ. I will always, silently, assume that the sequences are 'long enough', that is, in any clause in which $\vec{e}$ is related to ϕ, it is assumed that the length of $\vec{e}$ is $s(\phi) = r(\phi) + n(\phi)$ at least. Moreover, if $\vec{e} = e_1 \ldots e_n$, then $\vec{e}_i = e_i$.

DEFINITION 8.1
(PLA Semantics)

- $[x]_{M,g,w,\vec{e}} = g(x) \quad [c]_{M,g,w,\vec{e}} = I_w(c) \quad [p_i]_{M,g,w,\vec{e}} = \vec{e}_i$
- $M, g, w, \vec{e} \models Rt_1 \ldots t_m \; iff \; \langle [t_1]_{M,g,w,\vec{e}}, \ldots, [t_m]_{M,g,w,\vec{e}} \rangle \in I_w(R)$
  $M, g, w, \vec{e} \models \neg\phi \; iff \; \neg\exists \vec{c} \in D^{n(\phi)} :$
  $M, g, w, \vec{c}\vec{e} \models \phi \quad M, g, w, d\vec{e} \models \exists x\phi \; iff \; M, g[x/d], w, \vec{e} \models \phi$
  $M, g, w, \vec{c}\vec{e} \models \phi \wedge \psi \; iff \; M, g, w, \vec{e} \models \phi \; and \; M, g, w, \vec{c}\vec{e} \models \psi \; (\vec{c} \in D^{n(\psi)})$

Apart from the possibility of there being pronouns, atomic formulas are evaluated in a totally classical way. A pronoun p_i simply picks up the i-th element of a satisfying sequence, thus establishing coreference with the i-th term before the pronoun. A negated formula $\neg\phi$ is satisfied if there is no way to find witnesses to satisfy ϕ. An existentially quantified formula $\exists x\phi$ is evaluated in the classical way, but for the fact that witnesses d for x by means of which ϕ can be satisfied are put on the stack of witnesses $\vec{e}$. Satisfaction of a conjunction $\phi \wedge \psi$ is also standard, except that it incorporates the 'pragmatic' fact that in actual use ϕ comes before ψ: ϕ is evaluated *before* ψ has contributed its $n(\psi)$ witnesses. Thus $\vec{ce}$ can be taken to satisfy ϕ in its conjunction with ψ because it satisfies ϕ plus the fact that $n(\psi)$ more terms have been used afterwards.

PLA models the interpretation of intersentential anaphoric relationships in a compositional way, without resorting to a representational format (which *DRT* does) and without changing the standard notions of scope and binding (which *DPL* does).[24] Technically, indefinites and pronouns (and upon a proper extension: definites) behave pretty similar. The different types of terms are assumed to be basically referential, and only differ in matters of use. Indefinites are assumed to be new, and they may leave the intended referent undetermined; pronouns (and definites) are assumed to be given and determinable.[25] Some of these facts show from the following observation (in which $\exists$ is short for $\exists x(x = x)$):

OBSERVATION 8.2
   (Indefinites and Pronouns)

(45) A diver found a pearl. She lost it again.
(46) A diver lost a pearl she found.
(47) There is something. It is a pearl. There is someone. She is a diver. She found it. She lost it again.

- $\exists x(Dx \wedge \exists y(Py \wedge Fxy)) \wedge Lp_1p_2 \Leftrightarrow$
  $\exists x(Dx \wedge \exists y((Py \wedge Fxy) \wedge Lxy)) \Leftrightarrow$
  $\exists \wedge Pp_1 \wedge \exists \wedge Dp_1 \wedge Fp_1p_2 \wedge Lp_1p_2))$

As these equivalences already suggest, we can do away with all variables and we can also eliminate all resolved pronouns. Resolved pronouns can be eliminated by means of a normalization procedure which draws from the following equations:

DEFINITION 8.3
   (Normal Binding Forms)

- $(Rt_1 \ldots t_m)^\bullet = Rt_1, \ldots, t_m$
  $(\neg\phi)^\bullet = \neg\phi \text{ if } \phi \text{ is in normalform}$
  $(\exists x\phi)^\bullet = \exists x\phi \text{ if } \phi \text{ is in normalform}$
  $(\exists\vec{x}\phi \wedge \exists\vec{y}\psi)^\bullet = \exists\vec{y}\vec{x}(\phi \wedge [\vec{x}]\psi) \text{ if}$

$-\phi$ and ψ are in normal form *and* closed

$-\vec{y}$ do not occur free in ϕ and $\vec{x}$ do not occur free in ψ

$-[x_1 \ldots x_n]\psi$ is obtained from ψ by replacing any pronoun p_i in ϕ by x_i if $i \leq n$ and by p_{i-n} otherwise (of course, the x_i should be free for the p_i in ψ)[26]

If embedded formulas are not in normal form, these equations have to be applied to them first (where ϕ is in normal form iff $(\phi)^\bullet \equiv \phi$). Notice that if ϕ is in normal form, then it is of the form $\exists \vec{x}\psi$, where ψ is both in normal form and closed. Computing normal binding forms is insightful for the following reasons:

OBSERVATION 8.4
(PLA, PL and DRT)

- $M, g, w, \vec{e} \models (\phi)^\bullet$ iff $M, g, w, \vec{e} \models \phi$

- let $\vec{x} = x_1 \ldots x_{r(\phi)}$ be not free in ϕ and free for $p_1 \ldots p_{r(\phi)}$ in $(\phi)^\bullet$, then

- $M, g, w \models_{CL} [\vec{x}](\phi)^\bullet$ iff $\exists \vec{e} \in D^{n(\phi)} : M, g, w, \vec{e} \models [\vec{x}](\phi)^\bullet$

- $[\vec{x}](\phi)^\bullet$ has the structure of a *DRS*, and is of the form $\exists \vec{z}\psi$ where $\vec{z}$ is a sequence of variables and ψ a a series of conjunctions of conditions (atomic formulas and negated formulas with the structure of *DRS*)

The first observation here shows the normalization procedure to be fully meaning preserving. The second shows that it produces a formula whose classical satisfaction conditions are the same in classical logic. (The additional substitution with $[\vec{x}]$ is needed to remove unresolved pronouns.) The third observation shows that these truth-conditions are adequately captured by the discourse representation structures of *DRT*. Thus it shows that *PLA*, like *DPL* and unlike *DRT*, can stick to a natural translation of natural language expressions which globally respects their syntactic structure. Even so the semantics of these expressions is equivalent with the (classical) interpretations of the corresponding representations produced in *DRT*.

UPDATE AND SUPPORT

I assume that information states of interlocutors contain information about the world and about the possible values of terms used in a discourse. They can be modeled by sets of sequences $w\vec{e}$ where the $\vec{e}$ are witnesses in worlds w conceived possible. Then we can define what it means to update an information state τ with the contribution made by an utterance of ϕ, the result written as $(\tau)[\![\phi]\!]_{M,g,\vec{e}}$, and what a speaker's state σ can be required to be like to support such an utterance, in case we will write $\sigma \models_{M,g,\vec{e}} \phi$. Since terms get an epistemic interpretation here, they are associated with (sequences of)

witnesses $\vec{\epsilon}$ which are individual concepts now, functions from the possibilities in an information state to individuals. They can be modeled as (sequences of) numbers, so that if $\vec{\epsilon}_i$ is j, it is associated with the projection function that assigns $\vec{e}_j$ to each relevant possibility $w\vec{e}$. Under an update with ϕ these sequences determine the interpretation of unresolved pronouns in ϕ only, whereas for ϕ to be supported, they must also specify the concepts supporting indefinites.[27] If $\vec{\epsilon}$ is a sequence of concepts modeled by the numbers $i_1 \dots i_n$, and if $w\vec{e}$ is a possibility, $\vec{\epsilon}(w\vec{e})$ is the sequence of n individuals $\vec{\epsilon}_1(w\vec{e}) \dots \vec{\epsilon}_n(w\vec{e})$ which, really, is the sequence $\vec{e}_{i_1} \dots \vec{e}_{i_n}$. The definitions run as follows:

DEFINITION 8.5
(Update and Support)

- $(\tau)\llbracket\phi\rrbracket_{M,g,\vec{\epsilon}} = \{w\vec{c}\vec{e} \mid w\vec{e} \in \tau \ \vec{c} \in D^{n(\phi)} \ M,g,w,\vec{c}\vec{\epsilon}(w\vec{e}) \models \phi\}$

  $\sigma \models_{M,g,\vec{\epsilon}} \phi$ iff $\forall w\vec{e} \in \sigma : M,g,w,\vec{\epsilon}(w\vec{e}) \models \phi$

If one accepts an utterance of ϕ, the information conveyed by ϕ is taken to be correct: one excludes possibilities inconsistent with its contents, and possibly intended witnesses are 'remembered'. In any remaining possibility $w\vec{c}\vec{e}$, $\vec{c}$ is conceived of as the sequence of individuals the speaker might have intended if his or her information is correct. Support can be understood in two ways. On a Gricean view it characterizes what a speaker's state must be like if she is cooperative and if she complies with the quality maxim. However, one may as well conceive of it as a qualitative characterization of the commitments the speaker makes, in the sense of Hamblin. These definitions are pragmatically well-behaved in the following sense:

OBSERVATION 8.6
(Supported Updates)

- $(\tau)\llbracket\phi\rrbracket$ and $\sigma \models \phi$ can be independently defined in a compositional way
- if $\sigma \models_{M,g,\vec{\epsilon}} \phi$ and ϕ is resolved, then $(\downarrow\sigma \cap \downarrow\tau) \subseteq \downarrow((\tau)\llbracket\phi\rrbracket_{M,g})$
  (where $\downarrow\sigma = \{w \mid \exists\vec{e}: w\vec{e} \in \sigma\}$)

The benefit of the first of these two observations is obvious, and it is substantiated in (Dekker, 2002a).[28] The second observation should be more appealing. It says that if an utterance of ϕ is resolved and supported, then the information which a hearer may get from it is supported by the distributed information he and the speaker had before the update. It means that supported updates are safe: they are reliable if the initial information was. (The requirement that ϕ be resolved arises from the fact that unresolved pronouns allow for the possibility of miss-resolution.)

Conditional sentences (as well as negated ones) also mediate nicely between update and support:

OBSERVATION 8.7
(Negations and Conditionals)

- $\sigma \models_{M,g,\vec{\epsilon}} \neg\phi$ iff $(\sigma)[\![\phi]\!]_{M,g,\vec{\epsilon}} = \emptyset$

  $\sigma \models_{M,g,\vec{\epsilon}} (\phi \rightarrow \psi)$ iff $\exists\vec{\alpha} : (\sigma)[\![\phi]\!]_{M,g,\vec{\epsilon}} \models_{M,g[+n(\phi)],\vec{\alpha}n(\vec{\phi})\vec{\epsilon}} \psi$

The first observation here is the most transparent one. Support for asserting $\neg\phi$ boils down to a veto on updating with ϕ, on the pain of inconsistency. Thus we see indeed something like a role switch at the formal level. Less transparent, but even more suggestive is the second observation. Support for asserting a conditional consists in having support for the consequent clause if one has updated with the antecedent clause. Thus a speaker's state supports $\phi \rightarrow \psi$ iff, if she updates with ϕ, her state supports ψ. Here we witness a double role-switch: if you support ϕ, I support ψ.[29]

As already indicated in the main text of the paper, our treatment of conditional sentences naturally invites two possible amendments, which are not restricted to conditional sentences though. As appears from observation (0), support for the assertion of a conditional consists in *possible* support for the consequent clause upon a hypothetical update with the antecedent clause. This type of support can be made specific, provided that it is functional upon the contents of the antecedent clause:

DEFINITION 8.8
(Conditional Satisfaction and Support)

- $M, g, w, \vec{ae} \models \phi \rightarrow \psi$ iff $\forall\vec{c} \in D^{n(\phi)} M, g, w, \vec{c}\vec{e}_{\vec{c}} \models \phi : M, g, w, \vec{a}_{\vec{c}}\vec{c}\vec{e}_{\vec{c}} \models \psi$

  $\sigma \models_{M,g,\vec{\alpha}\epsilon} \phi \rightarrow \psi$ iff $(\sigma)[\![\phi]\!]_{M,g,\vec{\epsilon}} \models_{M,g[+n(\phi)],\vec{\alpha}_{n(\vec{\phi})}n(\vec{\phi})\vec{\epsilon}_{n(\vec{\phi})}} \psi$

In this definition, $\vec{ae}$ and $\vec{\alpha}\epsilon$ indicate (sequences of) witnesses and concepts which are possibly functionally dependent on the (sequences of) witnesses $\vec{c}$ and concepts $n(\vec{\phi})$ respectively.[30] The definition generates equivalences like the following:

OBSERVATION 8.9
(Functional Support)

(19) If a book is printed with Kluwer it has an index. It can always be found at the end.

- $(\exists x\phi(x) \rightarrow \exists y\psi(y)) \wedge (\exists x\phi(x) \rightarrow \chi(\mathrm{p}_2)) \Leftrightarrow$
  $\exists f(\exists x\phi(x) \rightarrow \psi(f(\mathrm{p}_1))) \wedge (\exists x\phi(x) \rightarrow \chi(\mathrm{p}_2)) \Leftrightarrow$
  $\exists f((\exists x\phi(x) \rightarrow \psi(f(\mathrm{p}_1))) \wedge (\exists x\phi(x) \rightarrow \chi(f(\mathrm{p}_1))))$

Example (19) can be reformulated as follows: if something is a book, say d, then $f(d)$ is its index; and if something is a book, say d', then $f(d')$ can be found at its end. These are by and large the truth-conditions of the example.[31]

INFORMATION STRUCTURE

The earlier definitions will also have to be amended in order to account for the flexible way in which quantifiers and other operators act upon the information structure of embedded clauses. Material from *if-* and *only if-* clauses, from negated expressions, and from the restrictions of quantifiers, sometimes float free from their local environment, especially terms (definite and indefinite) and presuppositions. We can account for this behaviour by distributing the contents of all expressions (sentential and other) over various dimensions of interpretation [basically as in (Karttunen and Peters, 1979)], and let the main operators act upon these dimensions in flexible ways. I end this appendix with a concise sketch of how this works. I assume a fully compositional three-dimensional satisfaction semantics for a fragment dealing with nominal terms, quantifiers and presupposition (it is available in manuscript, but has not yet been published).

The basic idea is to split up meaning in a background and a focus component, where the background again is split up in an old and a new part. I will use 'presupposition,' 'contribution' and 'assertion' as mere technical labels for the old, new, and focal dimension respectively, and let these labels apply to all the relevant categories and types of our language. Thus, for instance, formulas will have a presuppositional, contributional, and assertional dimension which is stated as a satisfaction relation, $\models_p$, $\models_c$ and $\models_a$, respectively; nouns and verbs will have sets of individuals as their presuppositional, contributional or assertional denotation, and terms and quantifiers sets of sets of individuals, etc. The ground of any clause can be identified by the conjunction or intersection of its presupposition and contribution (so that $\models_b = (\models_p \cap \models_c)$) and the full interpretation is the conjunction or intersection of $\models_p$, $\models_c$ and $\models_a$. I present some examples by way of illustration.

(48) A student qualifies if she satisfies the prerequisites.
    $IF(P(\mathrm{p}_1))(\ SOME(S)(Q))$

(49) A student qualifies only if she satisfies the prerequisites.
    $ONLY\_IF(P(\mathrm{p}_1))(\ SOME(S)(Q))$

The main clause of these examples gets satisfied as follows:

- $M, g, w, d\vec{e} \models_b\ SOME(S)(Q)$ iff $d \in I_w(S)$
  $M, g, w, d\vec{e} \models_a\ SOME(S)(Q)$ iff $d \in I_w(Q)$

Operators like *IF* and *ONLY_IF* have their usual interpretation, but they can quantify (universally or generically) over the ground of embedded clauses. Such a (one- or more-place) operator O can then be topically restricted, roughly as follows:

- $M, g, w, R\vec{e} \models_b O^t(\phi)$ iff $R = \{\vec{c} \mid M, g, w, \vec{c}e \models_b \phi\}$

  $M, g, w, R\vec{e} \models_a O^t(\phi)$ iff $\forall \vec{c} \in R : M, g, w, \vec{c}e \models_a O(\phi)$

For examples (48) and (49) this would give:

- $M, g, w, \vec{e} \models$ (48) iff $\forall d \in I_w(S) : d \in I_w(Q)$ if $d \in I_w(P)$
- $M, g, w, \vec{e} \models$ (49) iff $\forall d \in I_w(S) : d \in I_w(Q)$ only if $d \in I_w(P)$

The donkey examples (12–15) can be analyzed in an entirely similar fashion.

The following example illustrates how we can deal with three dimensions:

(50) The boy bought a DVD-drive.
  $THE(BOY)(\lambda x SOME(\lambda y DVDy)(\lambda y BUYxy))$

- $M, g, w, bd\vec{e} \models_p$ (50) iff $\{b\} = I_w(BOY)$

  $M, g, w, bd\vec{e} \models_c$ (50) iff $d \in I_w(DVD)$

  $M, g, w, bd\vec{e} \models_a$ (50) iff $\langle b, d \rangle \in I_w(BUY)$

(51) He likes it very much. $HE_1(\lambda x IT_2(\lambda y LIKExy))$

- $M, g, w, b'd'bd\vec{e} \models_p$ (51) iff $b' = b$ and $d' = d$

  $M, g, w, b'd'bd\vec{e} \models_c$ (51) iff $\top$

  $M, g, w, b'd'bd\vec{e} \models_a$ (51) iff $\langle b', d' \rangle \in I_w(LIKE)$

As the reader can verify, the joint satisfaction of the presuppositions, contribution and assertion of the examples (50) and (51) is like we had it before.[32] However, by distributing the contents of such sentences over separate dimensions we can access and handle the various parts flexibly and in different ways in the process of further composition. Consider again example (5), assuming for the sake of simplicity it involves the sentential negation of (53):

(53) Onno runs a sushibar.
  $ONNO(\lambda x\ SOME(SB)(\lambda y\ RUNxy))$
(5) Onno doesn't run a sushibar.
  $NOT(ONNO(\lambda x\ SOME(SB)(\lambda y\ RUNxy)))$

As we have argued above, depending on the arguably pragmatic effects of context, intonation and focus, asserting (5) can convey various things, which can be explained in terms of the different ways in which the negation acts on the information structure of the embedded clause. The four readings mentioned there ('standard', topical / generic, specific and presupposition denying) can be obtained in the following ways. The assertion of the embedded

sentences (53) presupposes a witness o for *onno*, contributes a witness b for a sushibar, and asserts that o runs b. Thus we can have the following construals (specifying relevant clauses only):

1. $M, g, w, o\vec{e} \models_b$ (5) iff $\exists b$: $M, g, w, ob\vec{e} \models_p$ (53)
   iff $o = I_w(\textit{onno})$
   $M, g, w, o\vec{e} \models_a$ (5) iff $\neg\exists b$: $M, g, w, ob\vec{e} \models_{c\&a}$ (53)
   iff $\neg\exists b$: $b \in I_w(SB)$ $\langle o, b \rangle \in I_w(RUN)$
2. $M, g, w, oB\vec{e} \models_b$ (5) iff $o = I_w(\textit{onno})$ & $B = I_w(SB)$
   $M, g, w, oB\vec{e} \models_a$ (5) iff $\forall b \in B$: $\langle o, b \rangle \notin I_w(SB)$
3. $M, g, w, ob\vec{e} \models_p$ (5) iff $o = I_w(\textit{onno})$
   $M, g, w, ob\vec{e} \models_c$ (5) iff $b \in I_w(SB)$
   $M, g, w, ob\vec{e} \models_a$ (5) iff $\langle o, b \rangle \notin I_w(RUN)$
4. $M, g, w, \vec{e} \models_p$ (5) iff $\top$
   $M, g, w, \vec{e} \models_c$ (5) iff $\top$
   $M, g, w, \vec{e} \models_a$ (5) iff $\neg\exists ob$: $M, g, w, ob\vec{e} \models_{p\&c\&\mathrm{p}_a}$ (53)

What is nice about the multi-dimensional set up is that all these interpretations can be obtained, without postulating a semantic ambiguity in the negative element. The different readings emerge from the different pragmatic ways in which that element can be taken to act upon the information structure of the embedded clause.

   Very much the same goes for the examples in which indefinites seem to violate island constraints. I end this appendix with a concise sketch of the three ways in which example (36) can be construed, using a slightly simplified version of the example, and again displaying only the relevant lines and parameters:

(36′) They came because a man came.
   $THEY_1(\lambda x\ BECAUSE(SOME(MAN)(CAME))(CAMEx))$

- $g, w, D \models_b$ (36′) iff $\top$
  $g, w, D \models_a$ (36′) iff $\forall d \in D$: $d \in I_w(CAME)$ because
      $\exists m$: $g[x/d], w, m \models SOME(MAN)(CAME)$ i.e. because
      $(I_w(MAN) \cap I_w(CAME)) \neq \emptyset$
- $g, w, mD \models_b$ (36′) iff $\forall d \in D$: $g[x/d], w, m \models_b SOME(MAN)(CAME)$
      iff $m \in I_w(MAN)$
  $g, w, mD \models_a$ (36′) iff $\forall d \in D$: $d \in I_w(CAME)$ because
      $g[x/d], w, m \models_a SOME(MAN)(CAME)$ i.e. because
      $m \in I_w(CAME)$
- $g, w, fD \models_b$ (36′) iff $\forall d \in D$: $g[x/d], w, f(d) \models_b SOME(MAN)(CAME)$
      iff $\forall d \in D$: $f(d) \in I_w(MAN)$
  $g, w, fD \models_a$ (36′) iff $\forall d \in D$: $d \in I_w(CAME)$ because
      $g[x/d], w, f(d) \models_a SOME(MAN)(CAME)$ i.e. because
      $f(d) \in I_w(CAME)$

Notes

[1] The research for this work is supported by a grant from the Netherlands Organization for Scientific Research (NWO), which is gratefully acknowledged. I wish to thank the editors of this volume and two anonymous reviewers for very useful and instructive suggestions and comments.

[2] But see, e.g., (Egli, 2000) for what has gone before.

[3] See (von Heusinger, 2003) for a dynamic implementation of the choice function strategy.

[4] It may be noticed that such a pragmatic understanding of the anaphoric relationship can also be cashed out in E-type terms. For the pronoun could also be interpreted as 'the individual which the previously used term was actually intended to refer to,' i.e., as the individual denoted by a referentially understood definite description. Notice that this definite description, besides being referential, is also highly indexical and intensional.

[5] I assume that the phenomenon of kataphora is much more constrained than that of anaphora, and that it ought to be dealt with by separate means.

[6] Thus, p_1 is coreferential with the indefinite used last, the most prominent one, so to speak; p_2 with the one but last, etc.

[7] For those interested, these contents can really be conceived of as the interpretations of Hans Kamp's discourse representation structures, or simply as Irene Heim's satisfaction sets. The difference is that these contents are taken to be independently specified here, and that they are not obtained by 'updating' a previously given information state.

[8] Notice that this type of soundness cannot be preserved when we start using a more expressive language. Asserting "You do not know it, but Carl will cook tonight." is self-corrupting in precisely this sense.

[9] So, again, a speaker is required to have *some* concept of the individual referred, even though this does not need to mean that she can identify it in any contextually relevant sense.

[10] Alternative interpretations are easily made available, of course, by emphasizing, e.g., "Onno", or "run". I here assume the utterance to carry what may be called a neutral intonation.

[11] See also the contribution (Farkas, 2003) to this volume, for the various uses of indefinites.

[12] Notice that, while Sandu's first conjunct can be understood to be about 'the men', it contributes and focuses in on a set of men m who own a gun $f(m)$. It is this latter set which the second conjunct quantifies over.

[13] Actually, this requires a recursive definition of (i) a class of type t of sets of satisfying cases of type c, (ii) a class of type c of sequences of witnesses of type w, and (iii) a class of type w of individuals of type e, individual concepts of type $\langle s, w \rangle$, and Skolem functions of type $\langle c, w \rangle$.

[14] A more minimal pair of examples is (28–29), due to (Read, 1992) [also discussed in (Gillon, 1996)]:

(28) Someone wins $1000 if he takes part. $(\exists x (Wx \leftarrow Px))$

(29) Someone wins $1,000 if everyone takes part. $(\exists x Wx \leftarrow \forall x Px)$

Reading $\leftarrow$ as a material implication the two are equivalent. For instance, if (29) is true, and everyone takes part, then (1000) is made true by taking the entailed winner as a witness for the existential quantifier; if (29) is true and not everyone takes part, then (28) is made by true by taking any person who does not take part as a witness. Reasoning from the truth of (28) to that of (29) proceeds analogously.

[15] A quite related example is known as the 'Beers Puzzle', as has been pointed out to me by Paul Egré, and the solution is the same as that of 'Peirce's Puzzle'. Also the following formulas are equivalent in first order logic:

- $(\forall x\phi(x) \rightarrow \forall x\psi(x))$

  $\exists x(\phi(x) \rightarrow \forall x\psi(x))$

  $(\neg\forall x\phi(x) \lor \forall x\psi(x))$

  $\exists x(\neg\phi(x) \lor \forall x\psi(x))$

Now consider:

(32) If everybody has a beer, then everybody has a beer.

(33) There is someone such that if he has a beer, everybody has.

(34) Either not everybody has a beer, or everybody has a beer.

(35) There is someone such that either he doesn't have a beer, or everybody has.

The examples (32) and (34) are clearly tautologous, and so are (33) and (35), truth-conditionally speaking. However, asserting the latter two implicates something non-trivial, viz., that there is someone such that his or her decision to take a beer will influence that of all others. A proper support for such utterances indeed requires the speaker to know of some such person, hence the examples are not equivalent, pragmatically speaking.

[16] Quite an unlikely reading: many more cities are not 100% male.

[17] Not a good reason, by the way, because Dorit has moved herself.

[18] Let me emphasize again that these referential or functional uses of the indefinite do not require the speaker to be able to specify or identify the intended individuals or functions in a contextually relevant sense. She may simply have heard that there is such a specific individual or function, and not be able to say anything more about it than that "it is the individual or function which So and So must have intended when he told me this."

[19] With the exception of Kratzer's approach. For (Kratzer 1998) the choice function is the denotation of a free variable, the interpretation of which is pragmatically determined.

[20] However, as has been pointed out by (Schwarz, 2001), this does not hold for all configurations. An example is the negation (37) of the intermediate reading of (38):

(37) Not every student read every book some teacher had praised.

(38) Every student read every book some teacher had praised.

Indeed, global existential quantification over the choice function variable in example (37) will not negate the reading obtained by global existential quantification over the choice function variable in example (38). However, as Schwarz himself observes, these readings are, of course, highly context dependent. As a matter of fact the pair of examples provides a good case for a pragmatic analysis like that of Kratzer, and the one advocated here. If the pragmatically determined interpretation of the free choice function variable in both examples is the same, or if (in our terms) the indefinite in both examples is used with the same (functionally dependent) referential intention, then the specific uses of (37) and (38) contradict each other, of course, and we obtain the readings argued for.

[21] Upon the previously mentioned, questionable, assumption.

[22] One could maintain that 'relevant' possibilities are those in which the set of philosophers is the same as in the actual world. Obviously this will not help when we turn to global interpretations of counterfactuals like the following:

(43) If some philosopher had not gone into philosophy, his tutors would have been disappointed.

(44) If a certain war criminal had not committed his crimes, this region would have been a much safer place.

[23] Universal quantification, disjunction and (material) implication can of course be defined using the classical equivalences: $\forall x\phi \equiv \neg\exists x\neg\phi$, $(\phi \rightarrow \psi) \equiv \neg(\phi \land \neg\psi)$, $(\phi \lor \psi) \equiv (\neg\phi \rightarrow \psi)$.

[24] It is a proper extension of a classical semantics and does not suffer from the technical complications which hamper *DPL* precisely because discourse information is 'hung' on variables there, which automatically introduces the possibility of unwanted information loss.

[25] To which it should be added that this type of 'determinability' may be dependent on what actually is the intended referent of other terms.

[26] If the conditions on free variable occurrences are not met, we can use α-conversion to produce alphabetical variants.

[27] The interpretation of variables thus has to be adjusted so that $[x]_{M,g,w,\vec{e}} = g(x)(e)$; after an update with ϕ an assignment g has to be updated to $h = g[+n(\phi)]$ defined by $h(x)(\vec{e}) = g(x)(\vec{e} - n(\phi))$, where $\vec{e} - m$ is $\vec{e}$ with the first m elements stripped off.

[28] Except for the fact that we have a more given a more homogeneous definition here. By way of illustration:

- $(\tau)[\exists x\phi]_{M,g,\vec{e}} = \{wd\vec{e} \mid w\vec{e} \in (\tau)[\phi]_{M,g[x/d]}\}$
  $(\tau)[\phi \wedge \psi]_{M,g,\vec{e}} = ((\sigma)[\phi]_{M,g,\vec{e}})[\psi]_{M,g[+n(\phi)],\vec{n}(\phi)\vec{e}}$

where $n(\vec{\phi})$ is the sequence of numbers (projection functions) $1 \ldots n(\phi)$ and $g[x/d]$ does not really assign d to x, but the constant function from possibilities to d. Likewise:

- $\sigma \models_{M,g,\vec{e}} \exists x\phi \quad$ iff $\sigma \models_{M,g[x/\vec{e}_1],\vec{e}-1} \phi$
  $\sigma \models_{M,g,\vec{\alpha}\vec{e}} \phi \wedge \psi$ iff $\sigma \models_{M,g,\vec{e}} \phi$ and $\sigma \models_{M,g,\vec{\alpha}\vec{e}} \psi$

[29] Negation can be phrased analogously, if $\neg\phi$ is interpreted as the equivalent $\phi \rightarrow \bot$. If my state supports this conditional it amounts to saying: if you support ϕ we are over and done.

[30] Similar definitions can be given for, e.g., the satisfaction of universally quantified assertions and belief reports:

- $M,g,w,\vec{e} \models \forall x\phi$ iff $\forall d$: $M,g[x/d],\vec{e}_d \models \phi$
  $M,g,w,\vec{e} \models B_x\phi$ iff $\forall w'$ $wR_{g(x)}w'$: $M,g,w',\vec{e}_{w'} \models \phi$

[31] Similar equivalences are raised by a functional interpretation of universally quantified assertions and belief attributions.

[32] As an example of a genuinely quantified structure, consider:

(52) Most men who sent a present to Curt sent a different$_2$ present to Amelia.

$MOST (\lambda x\ MANx \wedge CURT(\lambda y SOME(\lambda z PRESz)(\lambda z SENDxyz)))$
$(\lambda x AMEL(\lambda y SOME(\lambda z\ z \neq p_2 \wedge PRESz)(\lambda z\ SENDxyz)))$

- presupposes witnesses c for Curt, a for Amelia, M for the non-empty set of men who sent a present to c, and a witness-function p which associates any man $m \in M$ with the present $p(m)$ which m sent to c
- contributes a witness-function p' which associates any man $m \in M$ with a present $p'(m)$ different from $p(m)$
- asserts that most $m \in M$ sent $p'(m)$ to a

References

Abusch D. (1994) The Scope of Indefinites. *Natural Language Semantics*, 2, pp. 83–135.

Aloni M., Beaver D., Clark B. (1999) Focus and Topic Sensitive Operators. In Dekker P. (ed.), *Proceedings of the 12-th Amsterdam Colloquium*, pp 55–60, ILLC/Department of Philosophy, Amsterdam.

Ballmer T. T. (1978) *Logical Grammar*. North Holland, Amsterdam.

Cooper R. (1979) The Interpretation of Pronouns. In Heny F., Schnelle H. (eds.), *Syntax and Semantics 10*, Academic Press, New York.

Dekker P. (2001a) Dynamics and Pragmatics of 'Peirce's Puzzle'. *Journal of Semantics*, 18, pp. 1–31

Dekker P. (2001b) On *if* and *only*. In Hastings R., Jackson B., Zvolenszky Z. (eds.), *Proceedings from SALT XI*, CLC Publications, Ithaca, NY.

Dekker P. (2002a) Meaning and Use of Indefinite Expressions. *Journal of Logic, Language and Information* 10, pp. 141–194.

Dekker P. (2002b) "She" 's Character. In Bras M., Vieu L. (eds.), *Semantics and Pragmatics of Discourse and Dialogue*, Elsevier, Amsterdam, pp. 7–28.

Dekker P. (2003) Grounding Dynamic Semantics. In Bezuidenhout A., Reimer M. (eds.), *Descriptions and Beyond: An Interdisciplinary Collection of Essays on Definite and Indefinite Descriptions and other Related Phenomena*, Oxford University Press, Oxford (in press).

Egli U. (2000) Anaphora from Athens to Amsterdam. In von Heusinger K., Egli U. (eds.), *Reference and Anaphorical Relations*, Kluwer, Dordrecht, pp. 17–29.

Egli U., von Heusinger K. (1995) The Epsilon-operator and E-type Pronouns. In Egli U. et al. (eds.), *Lexical Knowledge in the Organization of Language*, Benjamins, Amsterdam, pp. 121–141.

Evans G. (1977) Pronouns, Quantifiers and Relative clauses (1). *The Canadian Journal of Philosophy*, 7, pp. 467–536, Reprinted in: Evans, G., 1985, *Collected papers*, Dordrecht: Foris, pp 76–152.

Fodor J., Sag I. (1982) Referential and quantificational indefinites, *Linguistics and Philosophy*, 5(3), pp. 355–398.

Gawron J. M. (1996) Quantification, Quantificational Domains and Dynamic Logic. In Lappin S. (ed.), *The Handbook of Contemporary Semantic Theory*, Blackwell, Oxford, pp. 247–267.

Geach P. T. (1968) Quine's Syntactical Insights, *Synthese* 19, Reprinted in Peter T. Geach, 1972, *Logic Matters*, pp. 115–127

Geurts B. (1999) *Presuppositions and Pronouns*, Elsevier, Oxford.

Gillon B. S. (1996) Three Theories of Anaphora and a Puzzle from C. S. Peirce. In Dekker P., Stokhof M. (eds.), *Proceedings of the Tenth Amsterdam Colloquium*, ILLC, University of Amsterdam, Amsterdam, pp. 283–297

Groenendijk J., Stokhof M. (1991) Dynamic Predicate Logic. *Linguistics and Philosophy* 14(1), pp. 39–100.

Heim I. (1982) *The Semantics of Definite and Indefinite Noun Phrases*, Ph.D. thesis. University of Massachusetts, Amherst, Published in 1988 by Garland, New York.

Jackendoff R. (1972) *Semantic Interpretation in Generative Grammar*, MIT Press, Cambridge, MA, WA.

Jacobson P. (1999) Towards a Variable-free Semantics. *Linguistics and Philosophy* 22, pp. 117–84.

Kamp H. (1990) Prolegomena to a Structural Account of Belief and Other Attitudes. In Anderson C. A., Owens J. (eds.), *Propositional Attitudes*, CSLI, Stanford, pp. 27–90.

Kamp H., Bende-Farkas Á. (2001) *Indefinites and Binding: From Specificity to Incorporation*, Lectures Notes for the 13th European Summer School in Logic, Language and Information (Helsinki), Revised Version, Stuttgart, November 2001.

Kamp H., Reyle U. (1993) *From Discourse to Logic*. Kluwer, Dordrecht.

Karttunen L. (1968) *What Do Referential Indices Refer To*, Technical Report P-3854, RAND Corporation, Reproduced by the Indiana University Linguistics Club.

Karttunen L., Peters S. (1979) Conventional Implicature, In Oh C.-K., Dinneen D. A. (eds.), *Syntax and Semantics 11 – Presupposition*, Academic Press, New York, pp. 1–56.

Kratzer A. (1998) Scope or pseudoscope? Are There Wide-scope Indefinites? In Rothstein S. (ed.), *Events in Grammar*, Kluwer, Dordrecht, pp. 163–196.

Krifka M. (2001) Non-novel Indefinites in Adverbial Quantification. In Condoravdi C., de Lavalette G. R. (eds.), *Logical Perspectives on Language and Information*, CSLI Publications, Stanford.

Matthewson L. (1999) On the Interpretation of Wide-scope Indefinites. *Natural Language Semantics*, 7, pp. 79–134.

Meyer-Viol W. (1995) *Instantial Logic*, Ph.D. thesis, ILLC, University of Amsterdam.

Peirce C. S. (1906) Prolegomena to An Apology for Pragmaticism. *The Monist*, 16, pp. 492–546, also appeared in Charles Hartshorne and Paul Weiss (eds.), 1960, *Collected Papers of Charles Sanders Peirce, Volume IV, The Simplest Mathematics*, Harvard UP, Cambridge, MA, pp. 411–463.

Peregrin J., von Heusinger K.: (2003) Dynamic Semantics with Choice Functions. In Kamp H., Partee B. H. (eds.), *Context-dependence in the Analysis of Linguistic Meaning*, Elsevier, Amsterdam.

Read S. (1992) Conditionals are not truth-functional: an argument from Peirce. *Analysis*, 52, pp. 5–12.

Reinhart T. (1997) Quantifier scope: How Labor is Divided Between QR and Choice Functions. *Linguistics and Philosophy*, 20, pp. 335–397.

Schlenker P. (1998) *A Note on Skolem Functions and the Scope of Indefinites*, Poster presented at the Northeastern Linguistic Society (NELS 29).

Schubert L., Pelletier F. (1989) Generically Speaking. In Chierchia G., Partee B. H., Turner R. (eds.), *Properties, Types and Meaning*, Kluwer, Dordrecht.

Schwarz B. (2001) Two Kinds of Long-distance Indefinites. In van Rooy R., Stokhof M. (eds.), *Proceedings of the Thirteenth Amsterdam Colloquium*, ILLC/Department of Philosophy, Amsterdam, pp. 192–197.

Slater H. (1986) E-type pronouns and ε-terms. *Canadian Journal of Philosophy*, 16, pp. 27–38.

Stalnaker R. (1998) On the representation of context. *Journal of Logic, Language and Information* 7, pp. 3–19.

Vallduví E. (1992) *The Informational Component*, Garland, New York.

van der Sandt R. A. (1989) Presupposition and Discourse Structure. In Bartsch R., van Benthem J., van Emde Boas P. (eds.), *Semantics and Contextual Expression*, Foris, Dordrecht, pp. 267–294.

van Rooy R. (1997) *Attitudes and Changing Contexts*, Ph.D. thesis, IMS, Stuttgart.

von Fintel K. (1994) *Restrictions on Quantifier Domains*, Ph.D. thesis, University of Massachusetts, Amherst.

von Heusinger K. (2003) Choice Functions and the Anaphoric Semantics of Definite NPs. *Biometals*, 17, pp. 309–329.

von Stechow A. (1991) Focusing and Backgrounding Operators. In Abraham W. (ed.), *Discourse Particles*, John Benjamins, Amsterdam.

Winter Y. (1997) Choice Functions and the Scopal Semantics of Indefinites. *Linguistics and Philosophy*, 20, pp. 399–467.

Winter Y. (2003) Functional Quantification. *Biometals*, 17, pp. 331–363.

Zimmermann E. (1999) Remarks on the Epistemic rôle of Discourse Referents. In Moss L. S., Ginzburg J., de Rijke M. (eds.), *Logic, Language and Computatio, Vol II*, Kluwer, Dordrecht, pp. 346–368.

Utility of Mention-Some Questions

ROBERT VAN ROOY
ILLC/University of Amsterdam, The Netherlands (E-mail: vanrooy@hum.uva.nl)

Abstract. In this paper, I argue that the 'ambiguity' between mention-all and mention-some readings of questions can be resolved when we relate it to the *decision problem* of the questioner. By relating questions to decision problems, I (i) show how we can measure the utilities of both mention-all and mention-some readings of questions, and (ii) give a natural explanation under which circumstances the mention-some reading is preferred.

Key words: decision theory, resolving underspecified questions

1. Introduction

According to most approaches towards questions, the meaning of a question is its set of possible (complete) answers. Groenendijk and Stokhof (1982) argue that an answer to a *wh*-question should be *exhaustive*, and should *mention all* the relevant individuals, while Hamblin (1973) assumes that in answering a *wh*-question one only needs to *mention some* positive instance, which seems particularly convincing for a question like *Where can I buy an Italian newspaper?* More recently it has been argued that *wh*-questions are in general *ambiguous* between a mention-some and a mention-all interpretation. But what, then, is the contextual parameter that resolves the ambiguity?

   In this paper, I will propose that the *decision problem* of the questioner is crucial here. The intuition behind this proposal is the natural assumption that we only ask questions to receive some particular kind of information; the kind of information that would help to resolve the *decision problem* that the questioner faces. By relating questions to decision problems, I (i) show how we can measure the utilities of both mention-all and mention-some readings of questions, and (ii) give a natural explanation under which circumstances the mention-some reading is preferred.

2. Questions as Sets of Answers

According to Hamblin (1958, 1973), we answer a question by making a statement that expresses a proposition. Just as it is normally assumed that

you know the meaning of a declarative sentence when you know under which circumstances this sentence is true, Hamblin argues that you know the meaning of a question when you know what counts as an appropriate answer to the question. Taking both assumptions together, this means that the meaning of a question as linguistic object (interrogative sentence) can be equated with the set of propositions that would be expressed by the appropriate linguistic answers. This gives rise to the problem what an appropriate linguistic answer to a question is.

According to almost all formal analyses of questions it is assumed that a *yes/no*-question like *Does somebody walk?* has only two appropriate answers: *Yes*, i.e., *Somebody walks*; and *No, Nobody walks*. Although polar questions have two appropriate answers, it is clear that only one of these two answers can be *true*. This means that with respect to each world a *yes/no*-question simply expresses a proposition: the proposition expressed by the true appropriate answer in that world. If we represent a *yes/no*-question simply by a formula like ?A, where A is a first-order formula, and assume that $[[A]]_g^w$ denotes the truth value of A in w with respect to assignment function g, the proposition expressed by question ?A in world w is:[1]

$$[[?A]]_{w,g}^E = \{v \in C : [[A]]_g^v = [[A]]_g^w\}$$

We might call the above proposition the *extension* of question ?A in world w. To determine the *intension* of the *yes/no*-question, we simply abstract away from the real world:

$$[[?A]]_g^I = \lambda w.\{v \in C : [[A]]_g^v = [[A]]_g^w\}$$

Notice that this function from worlds to propositions is simply equivalent to the following set of propositions:

$$[[?A]]_g^I = \{\{v \in C : [[A]]_g^v = [[A]]_g^w\} \mid w \in C\}$$

and that this set of propositions *partitions* the worlds in C.

Given this analysis of polar interrogative sentences, the question arises what the meaning of a *wh*-question is; i.e., what counts in a world as an appropriate true answer to a question like *Who walks?* Hamblin (1973) makes the following proposal:

> [...] a question sets up a choice-situation between a set of propositions, namely, those propositions that count as answers to it. [...] we shall regard 'who walks' as denoting a set, namely, the set whose members are the propositions denoted by 'Mary walks', 'John walks', ... and so on for all individuals (p. 48).

Notice that in distinction with the above approach towards *yes/no*-questions, a *wh*-question might have more than one true appropriate answer according to Hamblin's analysis. In the above quote Hamblin talks only

about single *wh*-questions, but we obviously don't need to restrict ourselves to them, and can analyze multiple *wh*-questions in a similar way. Let us assume that if A is an (open) first order formula and $\vec{x}$ the sequence of variables $x_1, \ldots, x_n$, we will represent (multiple) *wh*-questions by formulae like $?\vec{x}A$. Its Hamblin-intension (*HI*) is then given by the following *function* from worlds to the set of propositions that correspond to the set of true answers of the question in that world, or the equivalent *set* below that, where $\vec{d}$ denotes an *n*-ary sequence of objects.

$$[[?\vec{x}A]]_g^{HI} = \lambda w. \left\{ \left\{ v \in C : [[A]]_g^v[\vec{x}/_{\vec{d}}] = 1 \ \&[[A]]_g^w\left[\vec{x}/_{\vec{d}}\right] = 1 \right\} \mid \vec{d} \in D^n \right\}$$

$$= \left\{ \left\{ v \in C : [[A]]_g^v[\vec{x}/_{\vec{d}}] = 1 \ \&[[A]]_g^w[\vec{x}/_{\vec{d}}] = 1 \right\} \mid \vec{d} \in D^n \ \& \ w \in C \right\}$$

Notice that the intention of a question according to Hamblin's analysis does *not* form a partition, because several elements of the set might overlap each other. This, of course, is due to the fact that according to Hamblin a *wh*-question might have more than one true appropriate answer in a world. But this means that a *wh*-question leaves to the answerer in several worlds a non-trivial *choice* how to answer the question. This choice will turn out to be important later.

We might call the above function the *intension* of a *wh*-question. To determine the *extension* of a question, we simply apply the function to the actual world. The extension of the *wh*-question in world w will then be $[[?\vec{x}A]]_g^{HI}(w)$, which is equal to

$$[[?\vec{x}A]]_{w,g}^{HE} = \left\{ \left\{ v \in C : [[A]]_g^v[\vec{x}/_{\vec{d}}] = 1 \ [[A]]_g^w[\vec{x}/_{\vec{d}}] = 1 \right\} \mid \vec{d} \in D^n \right\}$$

Notice that this set is the set of *true* answers to a *wh*-question, and for John to know who walks it seems reasonable to demand that the set of worlds that represents his knowledge state in w has to be a subset of an element of the extension corresponding to the question.

Groenendijk and Stokhof (1982) have argued, however, that knowing for one individual who walks *that* he walks is not enough for John to know who walks. They claim that to know the answer to the question *Who walks?* John needs to know of *each* single individual *whether* he or she walks. In general, Groenendijk & Stokhof argue that John knows in world w the answer to the question that is represented by $?\vec{x}A$ if and only if the set of worlds that represents his knowledge-state is a subset of the denotation of $[[?\vec{x}A]]_{w,g}^E$:

$$[[?\vec{x}A]]_{w,g}^E = \{ v \in C \mid [[\lambda\vec{x}A]]_g^v = [[\lambda\vec{x}A]]_g^w \}$$

where the lambda term $\lambda \vec{x} A$ denotes the following set of n-ary sequences with respect to world w and assignment function g:

$$[\lambda \vec{x} A]]_g^w = \left\{ \vec{d} \in D^n \mid [[A]]_g^w [\vec{x}/\vec{d}] = 1 \right\}$$

This above denotation might be called the *extension* of a question. To determine the corresponding *intension* we can, as always, simply abstract from the world. What results is the following *function* from worlds to propositions, or, equivalently, the *set* of propositions below:

$$[[?\vec{x} A]]_g^I = \lambda w.\{v \in C \mid [[\lambda \vec{x} A]]_g^v = [[\lambda \vec{x} A]]_g^w\}$$
$$= \{\{v \in C \mid [[\lambda \vec{x} A]]_g^v = [[\lambda \vec{x} A]]_g^w\} \mid w \in C\}$$

Notice that this set of propositions gives rise to a *partition* of the state space C. The intension of a question is a set of mutually exclusive propositions thought of as the set of all alternative *exhaustive* answers to the question.

Groenendijk and Stokhof's analysis of questions has a number of nice properties not shared by the analysis of Hamblin (1973), nor by Karttunen's (1977) that is built on it. First of all, on their assumption that the extension of a question is a proposition, they can straightforwardly explain why questions can freely be conjoined with declaratives when embedded under verbs like *know*. In particular, to account for *wh*-complements like *John knows who came to the party*, they don't need to postulate two separate verbs of knowledge, as Karttunen (1977) had to. Second, their analysis has the consequence that not only single and multiple *wh*-questions have denotations of the same category, but that also *yes/no*-questions are analyzed in the same way as *wh*-questions. This has the important consequence, third, that they can give a general definition of *entailment* between all kinds of interrogatives simply by inclusion of intension. Thus, if Q and Q' are the intensions denoted by two questions, question Q is said to entail question Q' iff $\forall q \in Q : \exists q' \in Q' : q \subseteq q'$.[2]

3. Mention-some Questions and Human Concerns

Although the partition analysis of questions has a number of satisfying features, there are also some worries with the approach. The main worry, perhaps, is that according to Groenendijk and Stokhof's (1982) *mention-all* analysis of questions, it is predicted that each question has at most one true and appropriate answer in a world. Groenendijk and Stokhof argue that this is not so problematic for *choice readings* of questions, or for questions *coordinated* by a *disjunction*, for they just express more than one question.

Unfortunately, however, this way out doesn't seem to work satisfactorily for other uses of interrogative sentences that can be truly and appropriately answered in more than one way. These uses of interrogative sentences are the sentences that get interpreted as *mention-some* questions.[3]

First, there are the examples discussed recently by Beck and Rullmann (1999) which contain expressions that explicitly mark non-exhaustivity:

(1) Who, for example, came to the party?

It is clear that you can completely answer this question without giving the exhaustive list of people who came to the party.

Second, there are questions like (2)–(4) that typically get a mention-some reading, although they are not explicitly marked as such:

(2) Who has got a light?
(3) Where can I buy an Italian newspaper?
(4) How can I get to the station?

Just like (1), also these questions can intuitively be answered appropriately by mentioning just one individual, place, or manner, i.e., you don't have to give an exhaustive list of persons that have got a light, place where you can buy an Italian newspaper, or way to go to the station, respectively.

On first thought it might seem that also if interrogative sentences on their mention-some reading just express one question, they are not really problematic for a mention-all analysis of interrogative sentences. The reason is that one can claim that although mentioning all relevant individuals would completely answer the *wh*-question, in practice it normally suffices to give only a *partial* answer.

However, this proposal is unsatisfactory, because even if I just mention one individual, place, or way, I have intuitively resolved the question, i.e. satisfactorily answered question (1)–(3) or (4). Moreover, as argued for by Groenendijk and Stokhof (1984, p. 532), it seems that not all partial answers to a question like (2) intuitively count as satisfactory answers. Although

(5) John hasn't got a light

would be a partial answer to (2) when it has a mention-all interpretation, the answer does intuitively not resolve the question, in case the *wh*-phrase ranges over more than two individuals.

A further fact that suggests that the appropriateness to answer questions like (2)–(4) by mentioning just one individual/place/manner should not be explained by suggesting that giving a *partial* answer normally suffices is the fact that also when (2)–(4) are embedded, like in

(6) John knows where he can find an Italian newspaper

the *wh*-phrase has still typically the mention-some interpretation. That is, John needs to know only one (relevant) place where he can find an Italian newspaper in order for the sentence to be true.

We can conclude that it doesn't seem to be a good strategy to explain mention-some answers to *wh*-questions by saying that in linguistic practice partial answers normally suffice. The natural question that arises is whether we can say something more about the kind of circumstances under which a mention-some interpretation of a *wh*-question arises.

It has also been noted by several authors that mention-some interpretations of a question like *Where is a P?* typically arise only in peculiar situations. Situations where the questioner has a problem, or goal, and learning one (relevant) place where a *P* is would already suffice to solve the problem how to reach the goal. Question (3), for instance, is typically asked by an Italian tourist in Amsterdam with the goal of getting an Italian newspaper in mind. The tourist doesn't really mind where he can buy one, all he is interested in is where he should go to buy one. Mentioning just one element of the set of alternative 'equally best' places will perfectly resolve the question.

Whether a *wh*-question has a mention-some or a mention-all reading thus seems to depend on whether a, and what kind of, *human concern* lies behind the fact that the question was asked. Groenendijk and Stokhof (1984) even argue that *wh*-questions typically have a mention-all reading, and that they *can* only get a mention-some reading when some particular human concerns are at stake. They notice that when we embed (3) under verbs like *wonder*, *ask* or *know*:

(7) John *wonders/asks/knows* where he can buy an Italian newspaper

the embedded question can, and typically will, have a mention-some reading. However, *wh*-complements embedded under verbs which are not related to human concerns only seem to allow for a mention-all interpretation:

(8) Where you can get gas *depends* on what day it is.

(9) Who will come is partly *determined* by who is invited

In this section, we have seen that *wh*-questions can have a mention-some reading, and in particular when human concerns, or goals, are at stake. Whether a mention-some answer suffices to resolve the question or not depends on how *useful* the answer is. The usefulness of the answer, in turn, should be related to the goals of the questioner (cf. Ginzburg, 1995). In Section 4, I propose to make this precise by using tools of a well developed theory of rational behavior: Bayesian decision theory.

4. Utilities of Questions

4.1. Utilities of Questions Represented by partitions

In Savage's (1954) classical formulation of Bayesian decision theory, a distinction is made between states of the world, acts, and consequences; states

of the world together with acts determine the consequences, each act-world pair has exactly one consequence, and the consequence of an act includes all features that are relevant to the decision maker's values. If we assume that the utility of doing action a in world w is $U(w, a)$, we can say that the *expected utility* of action a, $EU(a)$, with respect to probability function P is

$$EU(a) \quad = \quad \sum_{w} P(w) \times U(w, a)$$

Let us now assume that the agent faces a *decision problem*, i.e., he wonders which of the alternative actions in A he should choose. A decision problem of an agent can be modeled as a triple, $\langle P, U, A \rangle$, containing (i) the agents probability function, P, (ii) his utility function, U, and (iii) the alternative actions he considers, A. In case the set of worlds and the set of actions are finite, we might represent such a decision problem as a decision table like the one below:

World	Prob.	Actions		
		a_1	a_2	a_3
u	1/3	4	-2	0
v	1/3	1	7	1
w	1/3	1	4	4

In this decision problem there are three relevant worlds, u, v, and w, and three relevant actions, a_1, a_2, and a_3. For each of these actions we can now determine its expected utility. The expected utility of action a_1, for instance, is $(P(u) \times U(u, a_1)) + (P(v) \times U(v, a_1)) \times (P(w) \times U(w, a_1)) = (1/3 \times 4) + (1/3 \times 1) + (1/3 \times 1) = 4/3 + 1/3 + 1/3 = 6/3 = 2$. In a similar way we can see that the expected utility of action a_2 is 3, while action a_3 has a utility of 5/3.

The problem that the agent faces is which action he should perform. You might wonder why we call this a decision *problem*; should the agent not simply choose the action with the greatest expected utility, i.e., action a_2? Yes, he should, if he *chooses now*. But now suppose that our agent doesn't have to choose now, but has the opportunity to first receive some useful information by *asking question Q*.

Before we can determine the utility of Q, we first have to say how to determine the expected utility of an action conditional on learning some new information. For each action a_i, its conditional expected utility with respect to new proposition C, $EU(a_i, C)$ is

$$EU(a_i, C) = \sum_{w} P(w/C) \times U(a_i, w)$$

When John learns proposition C, he will of course choose that action in A which maximizes the above value. Then we can say that the utility value of making an informed decision conditional on learning C, UV(Learn C, choose later), is the expected utility conditional on C of the action that has highest expected utility:

$$UV(\text{Learn } C, \text{ choose later}) = \max_i EU(a_i, C)$$

In terms of this notion we can determine the value, or *relevance*, of the assertion C.[4] Referring to a^* as the action that has the highest expected utility according to the original decision problem, $\langle P, U, A \rangle$, i.e., $\max_i EU(a_i) = EU(a^*)$, we can determine the *utility value* of the *assertion* C, $UV(C)$, as follows:

$$
\begin{aligned}
UV(C) &= UV(\text{Learn } C, \text{ choose later}) - UV(\text{Learn } C, \text{ still do } a^*) \\
&= \max_i EU(a_i, C) - EU(a^*, C)
\end{aligned}
$$

This value, which in statistical decision theory (cf. Raiffa and Schlaifer, 1961) is known as the *value of sample information C*, $VSI(C)$, can obviously never be negative. In fact, it predicts that an assertion only has a positive utility value in case it influences the action that the agent will perform. And indeed, it doesn't seem unnatural to say that a cooperative participant of the dialogue makes a *relevant* assertion just in case it makes our agent *change* his mind with respect to which action he should take. In our above example, for instance, we can see that proposition $\{v\}$ has a utility value of 0, because the best action to perform after learning that v is the case is the same action as the one that would have been performed with respect to the original decision problem. Proposition $\{u, w\}$, on the other hand, has a positive utility, because if the agent would learn this proposition, he would change his mind and would perform action a_1 instead of action a_2. The utility value of $\{u, w\}$ would be $(1/2 \times (4 - (-2)) + (1/2 \times (1 - 4)) = 6/2 + (-3/2) = 3/2$. It seems not unreasonable to claim that in a cooperative dialogue the assertion that expresses $\{u, w\}$ is 'better' than the assertion that expresses $\{v\}$, because the former has a higher utility value.[5] In general, we can say that one assertion, A, is 'better' than another, B, just in case the utility value of the former is higher than the utility value of the latter, $UV(A) > UV(B)$.

In terms of the utility value of assertions/answers, we can now determine the utility values of *questions*. Suppose that question Q is represented by the partition $\{q_1, \ldots, q_n\}$. Then we can determine the *expected* utility value of a question, $EUV(Q)$ as the *average* utility value of the possible answers:

$$EUV(Q) = \sum_{q \in Q} P(q) \times UV(q)$$

Suppose that for our above example a question was asked that could be represented by the partition $\{\{u, w\}, \{v\}\}$. The expected utility value of the question would then be $(P(\{u, w\}) \times UV(\{u, w\})) + (P(\{v\}) \times UV(\{v\})) = (2/3 \times 3/2) + (1/3 \times 0) = 1$. Notice that because the utility values of assertions can never be negative, the above determined expected value of a question, which in statistical decision theory is known as the *expected value of sample information*, $EVSI$, can also never be negative. In fact, the value will only be 0 in case a^* *dominates* all other actions in A with respect to the question. An action dominates the other actions in A with respect to the question in case no answer to the question would have the result that the agent will change his mind about which action to perform, i.e., for each answer q it will be the case that $\max_i EU(a_i, q) = EU(a^*, q)$. In these circumstances the question really seems irrelevant and, assuming that asking questions is *cost free*, it seems natural to say that question Q is *relevant* just in case $EUV(Q) > 0$. It should be obvious that this measure function totally orders all questions with respect to their expected utility values.[6]

4.2. UTILITY OF MENTION-SOME QUESTIONS

We have seen above that mention-some interpretations leave a choice to the answerer in several worlds how to answer the question, because several answers in the intension of a question might overlap each other. Let us make this a bit more concrete by defining for a particular question whose intension can be represented by $\{\{u, w\}, \{v, w\}\}$ the different *answer rules* that represent the different ways the answerer could answer this question. Notice that in this simple example the answerer has a non-trivial choice only in world w, and, thus, there are only two answer rules relevant. According to the first answer rule, f, the answerer answers in both u and w by a sentence that expresses $\{u, w\}$, and in v he answers by a sentence that would express $\{v, w\}$. According to the second answer rule, g, on the other hand, the answerer answers only in u by a sentence that expresses $\{u, w\}$, but answers in both v and w by a sentence that expresses $\{v, w\}$. Notice that although the question represented by $\{\{u, w\}, \{v, w\}\} = Q$ is not a partition, if we look for each answer rule at the set of worlds in which a particular answer is given, this latter set will form a *partition*. For answer rule f, for instance, this latter set will be $\{f^{-1}(q)| \ q \in Q\} = \{\{u, w\}, \{v\}\}$.

For our question above only two answer rules were possible, but depending on how much overlap there exists between the possible answers a mention-some reading of a question can get, many more answer rules can be relevant. For a particular question Q', let us denote this set by F. Because the answerer might use any element in F, the questioner doesn't know which answer rule the answerer will actually use. Let us temporarily assume, however, that he *does* know which f will be

used. In that case, the utility of choosing after he learned the answer q, UV_f(Learn q, choose later), should be determined as follows:

$$UV_f(\text{Learn } q, \text{ choose later}) = \max_i EU(a_i, f^{-1}(q))$$

In terms of this notion, we can now also define the utility of answer q, $UV_f(q)$:

$$UV_f(q) = UV_f(\text{Learn } q, \text{ choose later}) - EU(a^*, f^{-1}(q))$$
$$= \max_i EU(a_i, f^{-1}(q)) - EU(a^*, f^{-1}(q))$$

This value will never be negative.

The problem that we want to solve is how to determine the utility of question Q that is not represented by a partition. I will do this in terms of the notion of $UV_f(q)$, and thus indirectly in terms of answer rules. Intuitively, to determine the utility of question Q we want to find out for each answer q in Q the probability that it will be given, i.e., $P(\text{get } q)$. The utility of the question Q is then equal to

$$EUV(Q) = \sum_{q \in Q} P(\text{get } q) \times UV('\text{get } q')$$

where $UV('\text{get } q')$ is the utility value of the proposition corresponding to the worlds in which answer q is given. If it is clear what the relevant answer rule is, f for instance, it is clear how to determine this utility: $UV('\text{get } q') = UV_f(q)$, and probability: $P(\text{get } q) = P(f^{-1}(q))$, i.e., the utility and probability of the set of worlds in which answer q will be given according to answer rule f. Because it is unclear, however, which answer rule is used, the probability that answer q will be given, $P(\text{get } q)$, cannot be set equal to $P(f^{-1}(q))$, but must rather be equated with $\sum_{f \in F} P(f) \times P(f^{-1}(q))$, assuming that the questioner's uncertainty about the answer rule that will be used can be quantified by probability function P.

If we agree on the proposal that the probability that answer q will be given, $P(\text{get } q)$, should be equated with $\sum_{f \in F} P(f) \times P(f^{-1}(q))$, the utility of our question Q with respect to the answer rules in F can be determined as follows:

$$EUV_F(Q) = \sum_{q \in Q} \sum_{f \in F} P(f) \times P(f^{-1}(q)) \times UV_f(q)$$

This formula looks rather complicated, but can, fortunately, be simplified considerably. First, notice that the probability that answer rule f will be chosen, $P(f)$, does not depend on any particular element of Q. This means that the above formula is equal to

$$EUV_F(Q) = \sum_{f \in F} P(f) \times \sum_{q \in Q} P(f^{-1}(q)) \times UV_f(q)$$

Remember now that $UV_f(q)$ is the same as $UV(f^{-1}(q))$, and that for each answer rule f, the set $\{f^{-1}(q) \mid q \in Q\}$ is a partition, even if Q itself is not. Let us call this partition Q^f. This partition can be thought of as the denotation of a mention-all question and has an expected utility value: $EUV(Q^f)$. Because this value $EUV(Q^f)$ is the same as $\sum_{q \in Q} P(f^{-1}(q)) \times UV_f(q)$, we can now redefine the value of $EUV_F(Q)$ also as

$$EUV_F(Q) = \sum_{f \in F} P(f) \times EUV(Q^f)$$

This redefinition is not only simpler to write down than the one we started out with, it also makes clear that we can easily compare the utilities of the mention-all and mention-some readings of *wh*-questions. This comparison is based on the easy to prove fact that if Q and Q' are the partitions denoted by two questions such that $Q \sqsubseteq Q'$, i.e., $\forall q \in Q : \exists q' \in Q' : q \subseteq q'$, the expected utility of Q will be at least as high as the expected utility of Q', $EUV(Q) \geq EUV(Q')$. Notice that this means that the question that is represented by partition Q has a utility at least as high as the perhaps non-partitional question Q', when answer rule f is used, if the following condition is fulfilled:

$$\forall q \in Q : \exists q' \in Q' : q \subseteq f^{-1}(q')$$

It is not difficult to see, fortunately, that for any answer rule this relation is guaranteed to exist between the partition induced by a mention-all reading of question $?\vec{x}A$ and the intension of the question on its mention-some reading, when for each sequence of individuals $\vec{d}$ in the relevant domain of 'quantification' of the sentence represented by $?\vec{x}A$, there exists a cell in $[[?\vec{x}A]]_g^I$ that denotes the set of worlds where $\vec{d}$ is the only element of $[[\lambda\vec{x}A]]_g$. Notice that when questions are interpreted with respect to an 'empty' context, this will always be the case.

Because the above fact holds for any arbitrary answer rule $f \in F$, also the *average* utility of the mention-some reading of the question, $EUV_F([[?\vec{x}A]]^{HI})$, can never be higher than the utility on the corresponding mention-all reading. From this we can conclude that under the above mentioned condition *the utility of a mention-some reading of a question can never be higher than the expected utility of the corresponding mention-all reading*.

This result is obviously relevant to understanding in which situations a *wh*-question has a mention-all or a mention-some reading. On the assumption that the questioner is rational and the answerer cooperative and knows the decision problem of the questioner, this suggests that a *wh*-question will usually get a mention-all interpretation, because usually the question has a utility that is strictly higher on this interpretation.

To see how things work, consider our earlier discussed example, again, where the alternative actions are a_1–a_3, and where the probabilities of the worlds u, v, and w, and the utilities of the actions in these worlds are given in the table below:

$Sick(x)$	World	Prob.	Actions		
			a_1	a_2	a_3
Only C	u	1/3	4	−2	0
Only D	v	1/3	1	7	1
C & D	w	1/3	1	4	4

Notice that this time I have assumed that in the three different worlds the property *being sick* has a different extension: in world u only individual C is sick, in world v only D, while in w both are sick. This means, obviously, that the *wh*-question *who is sick?*, represented by the formula $?x\,Sick(x)$, should on its mention-all interpretation be represented as $\{\{u\}, \{v\}, \{w\}\}$. The expected utility of the question on this interpretation can then be calculated as $\sum_q P(q) \times UV(q) = (1/3 \times (4 - (-2)) + (1/3 \times 0) + (1/3 \times 0) = 6/3 = 2$. In section 4.1 we have determined the utility of question $\{\{u, w\}, \{v\}\}$ with respect to the same decision table, and we found that this question has a utility of 1. Notice that the question represented by $\{\{u\}, \{v\}, \{w\}\}$ is *finer-grained* than the question represented by $\{\{u, w\}, \{v\}\}$, and that—in accordance with what we have said above—the former question has indeed a *higher utility*, i.e., 2 versus 1.

Whereas the question represented by $?x\,Sick(x)$ should be represented by the partition $\{\{u\}, \{v\}, \{w\}\}$ on its mention-all reading, on a mention-some interpretation the question can be represented by the following set of propositions: $\{\{u, w\}, \{v, w\}\}$. This set of propositions does not form a partition, because the answers overlap each other. Because the answers overlap each other, we should analyze the utility of this question in terms of *answer rules*.

Notice that just as in the example discussed in Section 4.2, the answerer has a non-trivial choice only in world w, and, thus, there are only two answer rules relevant. According to the first answer rule, f, the answerer answers in both u and w by a sentence that expresses $\{u, w\}$, and in v he answers by a sentence that would express $\{v, w\}$. According to the second answer rule, g, on the other hand, the answerer answers only in u by a sentence that expresses $\{u, w\}$, but answers in both v and w by an sentence that expresses $\{v, w\}$. We have seen above that although the question represented by $\{\{u, w\}, \{v, w\}\} = [[?x\,Sick(x)]]^{HI}$ is not a partition, if we look for each answer rule at the set of worlds in which a particular

answer is given, these latter sets will form partitions: for answer rule f this will be $\{\{u, w\}, \{v\}\}$, and for answer rule g it is $\{\{u\}, \{v, w\}\}$. We have determined already that the former partition has a utility of 1, and the utility of the latter partition is $(P(\{u\} \times UV(\{u\})) + (P(\{v, w\}) \times UV(\{v, w\})) = (1/3 \times (4 - (-2))) + (2/3 \times ((1/2 \times 0) + (1/2 \times 0))) = 1/3 \times 6 = 2$. Because each answer rule is equally likely, the *average* expected value of the question is $(1/2 \times EUV(\{\{u, w\}, \{v\}\})) + (1/2 \times EUV(\{\{u\}, \{v, w\}\})) = (1/2 \times 1) + (1/2 \times 2) = 3/2$. This value of question *Who is sick?* on its mention-some interpretation is *lower* than the corresponding utility value of the question on its mention-all interpretation. This is in accordance with our earlier findings where we saw that a *wh*-question on its mention-some reading will never have a higher utility than the question on its mention-all interpretation.

5. Decision Problem as Contextual Parameter

Notice, however, that our above result does not rule out the possibility that in some particular situations the utility of the mention-some interpretation of a question will be equal to the utility of the corresponding mention-all reading. I claim that due to pragmatic reasoning, exactly in these circumstances the interrogative sentence will get a mention-some reading. The reason is that providing a mention-some answer causes less *effort* than providing a mention-all answer.[7]

Pragmatics can be seen as the study of the interaction between context and utterance. A context should represent enough information to be able to determine both *what* is said (or meant) by an utterance, and whether it was used *appropriately*. We have seen above that the decision problem of the agent who asks a question is the crucial contextual parameter that helps to determine whether the interrogative sentence was used appropriately, i.e., whether the question was *relevant* in its context of interpretation. In van Rooy (1999) I argued that the decision problem of the questioner is also the crucial contextual parameter to determine what it takes for an assertion to resolve the question. Just like for other contextual parameters, also the interaction between decision problem (i.e., the relevant contextual parameter) and interrogative used might go in two directions. If you don't know the decision problem, i.e., the *intentions* of the speaker, you might learn something (by *accommodation*) about it from the interrogative sentence used. For linguistic applications of our framework, however, we will concentrate ourselves in this paper on the other side of the interaction. If you *do* know the relevant decision problem of the questioner, you typically will be able to find out what it takes to resolve a question.

Suppose now that a question is used that allows for several interpretations. Which of those interpretations was actually intended by the questioner? The answer is simple: the interpretation with the highest utility. On

the assumption that we (more or less) know the questioner's decision problem, we can calculate for each interpretation of the question its expected utility. On the assumption that it is common ground that the speaker is rational, and thus a utility maximizer, the hearer can infer that the question—interrogative sentence—should have the most relevant/useful interpretation with respect to the questioner's decision problem.

I have claimed above that in some particular situations the utility of the mention-some and mention-all readings of *wh*-questions coincide, and that in these situations it suffices for the answerer to give only a mention-some answer. In these situations the question receives a mention-some interpretation in order to minimize effort. As a typical example where this is the case, consider the table below.

$?xBIN(x)$	world	P	n	s
Only N	u	1/3	6	0
Only S	v	1/3	0	6
N & S	w	1/3	4	4

In this example, we consider three relevant worlds where the extension of the predicate *places where you can buy an Italian newspaper* differ: world u, where you can buy an Italian newspaper only in the North, world v, where you can buy one only in the South, and world w, where you can buy one at both places. The decision problem also contains two relevant actions: action n which denotes the action of walking north; and action s which denotes the action of walking south. The decision table represents a situation where (i) the agent has no preference for learning that he can only buy an Italian newspaper in the N(orth) or only in the S(outh), because walking n(orth) and s(outh), respectively, would in those cases have an equal utility, and (ii) it's indifferent to him to walk either n(orth) or s(outh) when he learns that he can buy a paper in both parts of the city. I claim that this is the typical kind of situation in which the relevant *wh*-question can receive a mention-some reading: in those situations the question can intuitively be resolved equally well by a mention-some answer as by a mention-all answer.

It turns out that in these situations also the utilities of the mention-some and mention-all readings of the question coincide. In our example above, for instance, the utilities of the two readings of the question *Where can I buy an Italian newspaper?*, modeled in the formal language by $?xBIN(x)$, turn out to be 2 for both. This is easy to see for the utility of the question on its mention-all reading, while it can also be easily checked that the mention-some reading has an expected utility of 2 with respect to each of its two answer rules, and thus also has an *average* utility of 2. Because the mention-some answer is (known to be) equally useful

as the mention-all one, and shorter, the interrogative will get the mention-some interpretation.

I have discussed above when a mention-some reading arises of a non-embedded question in terms of the decision problem that the questioner faces. But it should be clear that the same reasoning can be used to determine when an embedded *wh*-question receives a mention-some reading. The only difference is that this time it need not be the decision problem of the questioner, or speaker, that is relevant, but it can also, and typically will, be the decision problem that the agent denoted by the subject of the embedding clause faces.

6. Conclusion

Following an idea of van Rooy (1999), I have shown in this paper that by relating questions to decision problems we can determine the utility of unambiguous questions and use it to resolve the underspecification of interrogative sentences. In this earlier paper only questions were considered that give rise to partitions, i.e., *wh*-questions that give rise to *mention-all* readings. In this paper, however, I have shown that we can also determine the utility of questions on their *mention-some* readings. To determine these latter utilities, I have made crucial use of *answer rules*, rules that determine which answer will be given in which worlds. Making use of these rules, I have shown that the utility of a mention-some reading of a *wh*-question will never be higher than the utility of the mention-all reading of the same question, but that their utilities sometimes coincide. I have argued that these facts are relevant for linguistic applications, because these expected utilities of the different readings of the same interrogative sentence might help to determine the actual question asked by an interrogative sentence, or better perhaps, might determine under which circumstances a mention-some answer suffices to answer a *wh*-question satisfactorily.

Acknowledgements

Notes

[1] Here, and elsewhere in this paper, I will assume that we analyze sentences with respect to a fixed model.

[2] Until now I have used the term 'question' only to denote interrogative sentences. From now on, however, I allow myself to be more liberal and will use it also for the intensions

interrogative sentences denote. I hope this double use of the notion will not lead to confusion.

[3] For a convincing argumentation that mention-some questions differ crucially from choice readings of questions (see Groenendijk and Stokhof, 1984).

[4] Notice that 'relevance' does not denote stochastic dependence here, as it standardly does in probability theory. The standard notion says that C is relevant for B iff $P(B/C) \neq P(B)$.

[5] It is important, however, not to think of v as a single world, but rather as a representative of lots of worlds that are similar enough to treat them as an equivalence class.

[6] In van Rooy (1999), I determined the utility of questions in a somewhat different way. It turns out, however, that the two ways of calculating the utilities of questions are equivalent (cf. van Rooy, to appear).

[7] In Van Rooy (ms.) I argue that interpretation results from balancing *relevance* with *effort*.

References

Beck S., Rullmann H. (1999) A flexible approach to exhaustivity in questions. *Natural Language Semantics*, 7, pp. 249–298.

Ginzburg J. (1995) Resolving Questions, I. *Linguistics and Philosophy*, 18, pp. 459–527.

Groenendijk J., Stokhof M. (1982) Semantic Analysis of Wh-Complements. *Linguistics and Philosophy*, 5, pp. 175–233.

Groenendijk J., Stokhof M. (1984) Studies in the Semantics of Questions and the Pragmatics of Answers. PhD Thesis, University of Amsterdam.

Hamblin C. L. (1973) Questions in Montague English. *Foundations of Language*, 10, pp. 41–53.

Karttunen L. (1977) Syntax and semantics of questions. *Linguistics and Philosophy*, 1, pp. 3–44.

Raiffa H., Schlaifer R. (1961) *Applied Statistical Decision Theory*. MIT Press, Cambridge, MA.

van Rooy R. (1999) Questioning to resolve decision problems. In: Dekker, P. (ed.), *Proceedings of the Twelfth Amsterdam Colloquium*, ILLC, Amsterdam.

van Rooy R. Comparing questions and answers: a bit of logic, a bit of language, and some bits of information. In: Groenendijk et al. (eds.), *Sources and Streams of Information*. CSLI, Stanford, to appear.

van Rooy R. Implicatures and Relevance in Bidirectional OT. University of Amsterdam, The Netherland. (MS.)

Savage L. J. (1954) *The Foundations of Statistics*. Wiley, New York.

Conditionals as Definite Descriptions
(A Referential Analysis)[1]

PHILIPPE SCHLENKER
Department of Linguistics, UCLA & Institut Jean–Nicod, CNRS, Paris, France (E-mail:
schlenke@humnet.ucla.edu)

Abstract. In *Counterfactuals*, David Lewis noticed that definite descriptions and conditionals display the same kind of non-monotonic behavior. We take his observation literally and suggest that *if*-clauses are, quite simply, definite descriptions of possible worlds [related ideas are developed in M. Bittner (2001) Proceedings from SALT XI, CLC, Cornell University, Ithaca, pp. 36–55]. We depart from Lewis's analysis, however, in claiming that *if*-clauses, like Strawsonian definite descriptions, refer. We develop our analysis by drawing both on Stalnaker's Selection Function theory of conditionals and on von Heusinger's Choice Function theory of definiteness, and by generalizing their analyses to plural Choice/Selection Functions. Finally, we explore some consequences of this referential approach: being definites, *if*-clauses can be topicalized; the word *then* can be analyzed as a pronoun that doubles the referential term; the syntactician's Binding Theory constrains possible anaphoric relations between the *if*-clause and the word *then*; and general systems of referential classification can be applied to situate the denotation of the descriptive term, yielding a distinction between indicative, subjunctive and 'double subjunctive' conditionals.

1. Introduction

After developing his logic of counterfactuals (Lewis 1973), David Lewis noted – almost in passing – that a weakened form of his system could be applied to definite descriptions, *modulo* a change of domain [possible worlds were replaced with individuals, and *if p, q* was replaced with (*the P*) *Q*]. The crucial observation, which I henceforth call 'Lewis's Generalization', was that the non-monotonic behavior of natural language conditionals is shared by definite descriptions, in a way that is not predicted by standard Strawsonian or Russellian treatments. For many years, Lewis's observation went largely unnoticed. The present paper is an attempt to revive it, and to take it quite literally: we suggest that *if*-clauses are simply definite descriptions of possible worlds [similar ideas were developed independently by Maria Bittner, who built upon important work on modality by Matthew Stone (Stone, 1997; Bittner, 2001); earlier versions of the present theory were

presented in Schlenker (1999, 2000, 2001). The analysis is also related to that found in Lycan (2001) and Schein (2001). I do not attempt a systematic comparison in this article].

Following Lewis's intuition, we will suggest that *if* should be seen as the form taken by the word *the* when it is applied to a description of possible worlds. However we depart from Lewis in claiming that both definite descriptions and *if*-clauses refer, something that Lewis's system was designed to avoid. Following Klaus von Heusinger's recent work (e.g. von Heusinger, 1996), we analyze definite descriptions in terms of Choice functions, and show that the latter are a simple variant of Stalnaker's Selection functions, which were originally designed to handle conditionals. In fact, Stalnaker's system can be used to improve on von Heusinger's analysis to ensure that his Choice Functions really do what they are designed to, namely model a notion of 'maximal salience' in discourse (further axioms are needed to achieve this, which can be found in Stalnaker's theory but not in von Heusinger's). By treating *if*-clauses as *plural* definite descriptions, we then obtain a generalization of Stalnaker's system, analogous to the 'class selection function' analysis of conditionals discussed in Nute (1980), or to Lewis's Logic *with* the 'Limit Assumption'. We then use this analysis to revisit the syntax and semantics of conditionals: being definites, *if*-clauses can be topicalized (Bittner, 2001; Bhatt and Pancheva, 2001); the word *then* can be analyzed as a pronoun which doubles the referential term (Iatridou, 1994; Izvorski, 1996); the syntactician's Binding Theory constrains possible anaphoric relations between the *if*-clause and the word *then*; and general systems of referential classification can be applied to situate the denotation of the descriptive term, yielding a distinction between indicative, subjunctive and 'double subjunctive' conditionals.

2. Lewis's Generalization

2.1. BASIC IDEA

Lewis's Generalization arose from a critique of traditional analyses of *if*-clauses and definite descriptions, which predicted 'monotonic' patterns of reasoning that turned out to be empirically incorrect. For instance, if a conditional is analyzed as a material or as a strict implication, it is predicted that *If p, q* should entail: *If p & p′, q* (Strengthening of the Antecedent). But *if*-clauses do *not* obey this pattern. As a result, it is no contradiction to assert (1)a (Lewis, 1973, p. 10), which is of the form in (1)b:

(1)   a. If Otto had come, it would have been a lively party; but if both Otto and Anna had come it would have been a dreary party; but if Waldo had come as well, it would have been lively; but...

    b. If O, L; but if O & A, ¬ L; but if O & A & W, L; but ...

By the same token, it appears that the following argument is not valid:

(2)   If John came, Mary would be happy. Therefore, if John came and he
        was drunk, Mary would be happy.

Lewis (1973, 1979) makes an entirely similar observation about definite
descriptions. While he starts from a Russellian analysis (which he criticizes),
it is somewhat easier to introduce the problem in its Strawsonian guise (the
Russellian variant is discussed below). The problem is that the same pattern
of strengthening is predicted to hold *whenever all descriptions can be used
felicitously*: from *[the P] Q*, one may infer *[the (P & P')] Q* (for if *[the P]* and
[the (P & P')] can be uttered felicitously in a given context, they must denote
the same individual, hence the entailment). But such patterns fail in natural
language, as shown by the following, uttered in a piggery (Lewis, 1973):

(3)   a. The pig is grunting; but the pig with floppy ears is not grunting; but
           the spotted pig with floppy ears is grunting; but...
        b. [The P] G; but [the (P& F)] ¬G; but [the (P & F & S)] G; but ...

For the same reason, the following reasoning is invalid:

(4)   The pig is grunting, therefore the pig with floppy ears is grunting.
        We may also ascertain, for future reference, that the problem arises
        with plural descriptions as well:

(5)   (Uttered in Los Angeles)
        a. The students are happy, but the students at the Sorbonne are not
           (non-contradictory)
        b. The students are happy, therefore the students at the Sorbonne are
           happy *(invalid)*

The logical patterns appear to be similar in the case of definite descriptions
and conditionals, in the (relatively weak) sense that a monotonic behavior
that one might expect systematically fails. Interestingly, the same kind of
solution has been offered for both cases, though under different names. In
order to handle the non-monotonicity of conditionals, Stalnaker (1968)
introduced the device of *selection functions*. Intuitively, the expression *if p*
evaluated at a world w was to select among the worlds that satisfy *p* the
one that is most similar to w. This presupposed that worlds could always be
completely ordered according to their degree of similarity to a given world
of reference. A weakened version of the same device, *choice functions*, has
been used by von Heusinger to handle the non-monotonic behavior of
definite descriptions. It posits that *the P* uttered in a context c selects
among the things that satisfy *P* the most salient one. Thus the notion of
salience in the domain of individuals plays a role analogous to that of
similarity in the domain of possible worlds.

Although Lewis (1973) was the first to state the connection between
if-clauses and definite descriptions, he did *not* resort to Choice functions or to
Selection functions to account for their non-monotonic behavior. Rather, he
designed a more general and more complicated system to address some al-
leged weaknesses in Stalnaker's original analysis of conditionals. The result

was that on Lewis's final analysis definite descriptions and *if*-clauses are taken *not* to refer, nor even to 'denote' in an extended sense, unless one makes an additional assumption (the 'Limit Assumption'). We will see that there is no empirical motivation for the greater generality afforded by Lewis's system, and that a simpler and more intuitive analysis can be preserved, both for *if*-clauses and for definite descriptions. In a nutshell, we will argue that (i) in case no object (or too many objects) satisfy the restrictor P, a referential failure occurs, be if for *if P* or *the P* – as is assumed for the latter by Strawsonian analyses; and (ii) otherwise, *if P*/*the P* picks out the worlds/ individuals that satisfy P and are highest on a scale of salience (*the P*) or similarity (*if P*). Thus, by analyzing *if*-clauses as *plural* definite descriptions, we will reduce the analysis of conditionals to the theory of plurals and definiteness. The result is a strengthened version of Lewis's system which is discussed but dismissed in Lewis (1973). It is also the system which is typically used in linguistic semantics for reasons of simplicity (see e.g. Heim, 1992 and Fintel, 1999).

2.2. THE NON-MONOTONIC BEHAVIOR OF CONDITIONALS AND DEFINITE DESCRIPTIONS

2.2.1. *If-clauses*

Let us remind ourselves of the problem faced by monotonic theories of conditionals. Both in traditional modal logic and in theories based on generalized quantification over possible worlds, *if*-clauses are analyzed in terms of universal quantification over possible worlds, as in (b) (modal analysis) and in (c) (generalized quantification; w^* denotes the world of utterance):

(6)   a. If I strike this match, it will light
          b. $\Box$ (I-strike-this-match $\Rightarrow$ it-will-light)
          c. $[\forall w: wRw^*$ & I-strike-this-match(w)] [it-will-light(w)]

Both representations give rise to the same truth-conditions, and predict patterns of monotonic reasoning that are also shared by material implications. Here are three major properties, which simply derive from the logic of universal quantification, as in: 'every accessible ϕ-world is a ψ-world':

(7)   a. Strengthening of the Antecedent: If *If ϕ, ψ* then *If ϕ & ϕ',ψ*.
          b. Contraposition: If *If ϕ, ψ*, then *If $\neg \psi$, $\neg \phi$*
          c. Transitivity: If *If ϕ, ψ* and *If ψ, χ*, then *If ϕ, χ*

But these properties do not hold of natural language conditionals, which appear to be 'non-monotonic'. Consider the following examples, each of which refutes one of the properties mentioned in (7):

(8)   a. Failure of Strengthening of the Antecedent:

> If this match were struck, it would light, but if this match had been soaked in water overnight and it were struck, it wouldn't light (modified from Stalnaker 1968)

b. Failure of Contraposition

(Even) if Goethe had survived the year 1832, he would be dead by now $\neq$ > If Goethe were not dead by now, he would not have survived the year 1832 (Kratzer)

c. Failure of Transitivity

If Jones wins the election, Smith will retire to private life. If Smith dies tomorrow, Jones will win the election

$\neq$ > If Smith dies tomorrow, Smith will retire to private life.

These properties have led to the development of special, 'non-monotonic' logics for conditionals. We now argue that the same properties hold of definite descriptions as well.

2.2.2. *Definite Descriptions*

A Strawsonian account predicts that the patterns in (7) should hold of definite descriptions when *if* is replaced with *the*, at least when all the definite description(s) involved can be used felicitously (i.e. when their presuppositions are satisfied):

(9)   a. If *The* ϕ, ψ, then *The* ϕ & ϕ', ψ

b. If *The* ϕ, ψ, then *The* $\neg \psi$, $\neg \phi$

c. If *The* ϕ, ψ and *The* ψ, χ, then *The* ϕ, χ

Let us first consider (9a). If *The* ϕ can be used felicitously, there is exactly one ϕ-individual in the domain of discourse. Hence if *The* ϕ & ϕ' can also be used felicitously, it must denote the same individual, and therefore the entailment should hold. The same reasoning applies to (9b): if both *The* ϕ and *The* $\neg\psi$ can be used felicitously, there is exactly one ϕ-individual and one $\neg\psi$-individual in the domain of discourse. If the former has property ψ, then it must be distinct from the latter, which thus couldn't have property ϕ (or else there would be two ϕ-individuals in the domain, contrary to hypothesis). Third, turning to (9c), if *The* ϕ and *The* ψ can both be used felicitously, then there is exactly one ϕ-individual and one ψ-individual in the domain of discourse. Thus if the first one has property ψ, it must be identical to the second, hence the entailment. Let us note, finally, that the same predictions hold of plural descriptions if these are analyzed in terms of maximality operators. For if *The* ϕ denotes the maximal ϕ-set in the domain of discourse, and it is included in a ψ-set, then: (i) *a fortiori* the same holds for the maximal ϕ & ϕ'-set, which derives (9a); (ii) the maximal $\neg \psi$-set cannot contain any ϕ-elements (or else the maximal ϕ-set would contain these elements too, and would thus fail to be included in a ψ-set); this, in turn, derives (9b). (9c) is derived in similar fashion: if the maximal ϕ-set s_1 is included in a ψ-set s_2 and

the maximal ψ-set (which must include s_2) is contained in a χ-set s_3, then of course s_1 must be included in s_3.

In natural language, however, all of these patterns fail, just as they do with conditionals. Thus the following inferences are not valid (note that either singular or plural descriptions can be used to make this point):

(10)   Invalid inferences
        a. The dog is barking, therefore the neighbors' dog is barking.
        a′. The pig is grunting, therefore the pig with floppy ears is grunting
        a″. (Uttered in Los Angeles)
The students are happy, therefore the students at the Sorbonne are happy
        b. The professor is not Dean, therefore the Dean is not a professor
        c. The students are vocal. The undergraduates in Beijing are students.
           Therefore the undergraduates in Beijing are vocal.
For the same reason, the following are non-contradictory:

(11)   Non-contradictory statements
        a. The dog is barking, but the neighbors' dog is not barking.
        a′. The pig is grunting, but the pig with floppy ears is not grunting
           (Lewis, 1973)
        a″. (Uttered in Los Angeles)
The students are happy, but the students at the Sorbonne are not
        b. The professor is not Dean, but of course the Dean is a professor.
        c. The students are vocal, and of course the undergraduates in Beijing
           are students, but the undergraduates in Beijing are certainly not
           vocal at the moment.

Superficially it would seem that Russell fares slightly better than Strawson, since on a Russellian analysis (9a) and (9b) do *not* come out as valid. This is because one of Strawson's definedness conditions could be violated in the consequent, which on Russell's analysis leads to falsity rather than undefinedness. Here are the relevant abstract examples:

(12)   a. Refutation of (9a): Suppose that $[\![\phi]\!] = [\![\psi]\!] = \{d\}, [\![\phi \& \phi']\!] = \emptyset$
        b. Refutation of (9b): Suppose that $[\![\phi]\!] = [\![\psi]\!] = \{d\}, [\![\neg\psi]\!] = \emptyset$

By contrast, the pattern in (9c) *is* predicted to be valid by Russell, and in this respect he fares no better than Strawson:

(13)   Proof of (9c): From *The* ϕ, ψ, we obtain: $|[\![\phi]\!]| = 1$ and $[\![\phi]\!] \subseteq [\![\psi]\!]$.
        From *The* ψ, χ, we obtain: $|[\![\psi]\!]| = 1$ and $[\![\psi]\!] \subseteq [\![\chi]\!]$. Taken together,
        these conditions entail: $|[\![\phi]\!]| = 1$ and $[\![\phi]\!] \subseteq [\![\chi]\!]$

But even for 10(a′) and 10(b′), where it would superficially appear to work, Russell's analysis is utterly implausible because it saves the coherence of the sentences only by giving the negation *wide* scope in the clause where it appears, as in (14b) and (15b). By contrast, giving the negation *narrow* scope would immediately yield falsehoods, as in (14c) and (15c):

(14)   *Situation:* there is one (salient) pig without floppy ears and one (less salient) pig with floppy ears in the domain of discourse.

a. The pig is grunting, but the pig with floppy ears is not grunting (Lewis, 1973)

b. [The P] G & ¬ [the (P&F)] G (can be true in this situation)

c. [The P] G & [the (P&F)] ¬ G (cannot be true in this situation)

(15) *Situation:* there is one (salient) professor who isn't a Dean and one (less salient) professor who is Dean in the domain of discourse.

a. The professor is not Dean, but of course the Dean is a professor.

b. ¬ [the P] D & [the D] P (can be true in this situation)

c. [The P] G & [the (P&F)] ¬ G (cannot be true in this situation)

Unfortunately for the Russellian, these sentences may remain intuitively true even when the negation clearly has narrow scope:

(16) a. The pig is grunting, but the pig with floppy ears is doing something other than grunting

   b. The professor is something other than Dean, but of course the Dean is a professor

I conclude that the Russellian's advantage is only apparent, and that on closer inspection Lewis's examples are as problematic for Russell as they are for Strawson.

2.3. POSSIBLE ANALYSES

The foregoing observations can be analyzed in at least two ways. One line, which has become standard for conditionals and which we shall extend to definite descriptions, is to posit a non-monotonic semantics, as we do in Section 3 using Choice Functions. The alternative is to deny that the semantics is non-monotonic, and to blame the appearance of a non-monotonic behavior on changing domain restrictions on world and individual quantifiers (Fintel, 1999, 2001). By itself, the debate between these two lines of analysis need not affect our general point, as long the same analysis – be it based on a non-monotonic semantics or on changing domain restrictions – is applied uniformly to definite descriptions and conditionals. Thus we argue dialectically, and try to show that, in either case, *if*-clauses can and should be analyzed as definite descriptions. We do discuss, however, one argument (based on Negative Polarity Item licensing) which might suggest that definite descriptions and conditionals do not display the same logical behavior (see Section 2.3.2, where we also provide some counter-arguments).

2.3.1. *Non-monotonic Semantics with Constant Domain Restrictions*

Why should *if*-clauses and definite descriptions display a non-monotonic behavior? Stalnaker 1968, who was solely concerned with conditionals, suggested that their semantics contains a superlative component. According to him, *if it rains tomorrow* refers to that world in which it rains tomorrow

which is *most similar* to the actual world. Von Heusinger (1996), who was solely concerned with definite descriptions, analyzed *the dog* as referring to that dog in the domain of discourse which is *most salient* for the speaker. There are many variations on these theories. A natural extension is to allow both *if*-clauses and definite descriptions to refer to a plurality of individuals or worlds – a measure that is undoubtedly necessary for the analysis of plural definite descriptions. A less natural extension (advocated by Lewis, 1973) is to allow conditionals to have truth conditions even in case there is no world or even group of worlds that counts as 'the closest' to the actual world. These variations are discussed later in the paper, but they all share the main features of the 'superlative' analysis we just sketched.

2.3.2. *Monotonic Semantics with Changing Domain Restrictions*[2]

Non-monotonic analyses were challenged in the recent literature on conditionals, esp. in Fintel (1999, 2001). The general idea is that bare *if*-clauses should be analyzed as restrictors of covert universal quantifiers over possible worlds. Conditionals *appear* to have a non-monotonic behavior because the implicit domain restriction of the universal quantifier need not remain fixed, even within the confines of a single sentence – according to Fintel (1999, 2001), the domain may be expanded, though not contracted. We discuss three arguments in favor of von Fintel's theory. The first two apply in the same way to *if*-clauses and definite descriptions. Thus if they are correct, they cast doubt on the non-monotonic analysis we develop in Section 3, but not necessarily on our general point that *if*-clauses can and should be analyzed as definite descriptions. The third argument is more complex, and might cast doubt on this assimilation, though we give counter-arguments against this conclusion.

(i) *The argument from changing domain restrictions.* Suppose that the implicit domain restriction on world or individual quantifiers were allowed to change in sentences such as (8a) or (14a). If conditionals were analyzed as structures of universal quantification over possible worlds, (8a) would then have a logical form such as $[\forall w : C(w)\&\phi]\psi$ *but* $[\forall w : C'(w)\&\phi'\&\phi]\neg\psi$, where C and C' are implicit restrictions on the quantifier domains. Similarly (14a) would be of the form $[\iota x : C(w)\&\phi]\psi$ *but* $[\iota x : C'(x)\&\phi'\&\phi]\neg\psi$. As long as the extension of C' is not a subset of the extension of C, there is no reason the resulting logical forms should be contradictory, which would obviate the need for a non-monotonic analysis. In his discussion of conditionals, Lewis, (1973) explicitly considers this possibility, and rejects it.[3] Commenting on examples such as (8a), he writes that 'our problem is not a conflict between counterfactuals in different contexts, but rather between counterfactuals in a single context. It is for this reason that I put my examples in the form of a

single run-on sentence, with the counterfactuals of different stages conjoined by semicolons and 'but' (Lewis, 1973, p. 13). Unfortunately Lewis assumes a premise which is false, namely that quantifiers that are uttered in the same context share the same domain restriction. This is clearly incorrect, as is shown by the following example, which was pointed out to me by A. Szabolcsi (it is originally due to D. Westerstahl):

(17)   [*Situation:* A committee must select some applicants. Some of the applicants are Italian, and there are also Italians on the committee, though of course they are not the same.] Every Italian voted for every Italian.

The intended reading is that every Italian *on the committee* voted for every Italian *among the applicants*. This means that the domain restriction is not the same for the first and the last noun phrase, although both occur in the same sentence. Lewis's premise is thus incorrect.

Of course this does not mean that his conclusion is, only that a stronger argument is needed to establish it. Peter Svenonius (p.c.) suggests that Lewis's conclusion can be established on the basis of the following contrasts:

(18)   [There are ten girls and ten boys in the class. Three girls raise their hands. Talking to the teacher, I say:]
        a. Wait, the girls have a question!
        b. Wait, the three girls have a question!
        c. ?Wait, the girls each have a question!
        d. #Wait, every girl has a question!
        e. #Wait, all girls have a question!
        f. ?Wait, all the girls have a question!
        g. ?Wait, each of the girls has a question!

On von Heusinger's salience-based theory, the contrast between (18a) and (18d) is entirely expected. The extension of the predicate *girl* is presumably the set of the ten girls in the class. The description *the girls* picks out the most salient group among those; a good candidate is the group of the three girls that have raised their hands, thus increasing their their salience. By contrast, *every girl* and *all girls*, not being definite descriptions, must quantify over all the girls in the domain of discourse, i.e. over the ten girls in the class. Now one could *try* to argue that *the girls* denotes the three girls rather than all ten because of an implicit domain description. But if such a domain restriction is available for the definite description, why is it not also available for other quantifiers such as *every* and *all*? It would seem that von Heusinger's theory is at a clear advantage to explain these data. I also note for completeness that *all the girls* or *each of the girls* may not quite behave in the way that one might expect given (18a), since for some speakers they do not just quantify over the three most salient girls but rather over all the girls in the domain of discourse. This is presumably a problem for every compositional analysis, since *all the girls* or *each of the girl* do not have the meaning that would be

obtained by applying the meaning of *all* or *each* to the meaning of *the girls*. We could stipulate that in such contexts the salience hierarchy becomes trivial, so that all girls in the domain are equally salient, and hence that the group of the most salient girls is the maximum set of girls in the domain. But obviously this only describes the problem, which I have to leave open in this article. Be that as it may, it is interesting to note that the same problem seems to arise with conditionals. *Necessarily, if p, q* appears to display a monotonic behavior, while *if p, q* does not, as is suggested by the following contrast:

(19)   a. #Necessarily, if the United States threw its weapons into the sea, there would be war. However (necessarily) if the United States and all other nuclear powers threw their weapons into the sea, there would be peace.

       b. If the United States threw its weapons into the sea, there would be war. However, if the United States and all other nuclear powers threw their weapons into the sea, there would be peace.

To my ear, the first sentence sounds contradictory, but the second clearly does not. A compositional solution is far from obvious, but the problem seems to be structurally analogous to the one we observed with respect to (18a) vs. (18d).

While the argument we just discussed concerned the denotation of definite descriptions and universal quantifiers uttered in the same context, the following (tentative) contrasts concern patterns of entailment:

(20)   a. (Uttered in Los Angeles) Every student is happy, therefore the students at the Sorbonne are happy.

       b. (Uttered in Los Angeles) #The students are happy, therefore the students at the Sorbonne are happy.

(21)   a. Necessarily, if the United States threw its weapons into the sea, there would be war. Therefore, necessarily, if the United States and all other nuclear powers threw their weapons into the sea, there would be peace.

       b. #If the United States threw its weapons into the sea, there would be war. Therefore, if the United States and all other nuclear powers threw their weapons into the sea, there would be peace.

Suppose that changing domain restrictions were responsible for the non-monotonic behavior of *if*-clauses and definite descriptions. Still, a charitable interpreter of the sentences in (20) and (21) should try to keep the domain restrictions fixed so as to make the arguments valid. But if the judgments shown above are correct, this is quite a bit harder with *the* than with *every* and with *if* than with *necessarily if*. These facts are unexpected on the monotonic analysis. By contrast, they are explained by the non-monotonic analysis, which predicts that the arguments in (b) should not be valid *even* when the domain restrictions are fixed. Unfortunately, it must be granted that the data are not entirely clear, and this argument should be considered as speculative at this point.[4]

(ii) *The argument from the order of presentation.* Another argument [attributed variously to I. Heim (in Fintel, 2001) or to Frank 1997] has been brought up against the non-monotonic analysis of conditionals. In the version discussed by Fintel (2001), the observation is that the non-monotonic behavior of conditionals is highly sensitive to the order in which the test sentences are presented, in a way which is not predicted by Stalnaker's or Lewis's analyses:

(22) a. If the USA threw its weapons into the sea tomorrow, there would be war; but if the USA and the other nuclear powers all threw their weapons into the sea tomorrow, there would be peace.
      b. ??If all nuclear powers threw their weapons into the sea tomorrow, there would be peace; but if the USA threw its weapons into the sea tomorrow, there would be war (Fintel, 1999, attributed to Heim)

I hasten to say that *exactly the same facts hold of definite descriptions*, as shown by the following:

(23) wa. The dog is barking, but fortunately the neighbor's dog isn't.
      b. ??The neighbor's dog is barking, but fortunately the dog isn't.

Thus although this argument might cast doubt on the non-monotonic analysis, it does not by itself argue against the treatment of *if*-clauses as definite descriptions.

It is clear that these facts are not explained by the non-monotonic analysis. As Fintel 1999 writes,

> in (22a), the two counterfactuals claimed to be consistent by Lewis are reversed in their order and the sequence does not work as before. The reason seems intuitively clear: once we consider as contextually relevant worlds where all nuclear powers abandon their weapons, we can't ignore them when considering what would happen if the USA disarmed itself. We seem to be in need of an account that keeps track of what possibilities have been considered and doesn't allow succeeding counterfactuals to ignore those possibilities. An account according to which the context remains constant throughout these examples would not expect a contrast between these two orders.

The same point could obviously be made about (23) as well.[5] Be that as it may, the monotonic analysis needs a stipulation to account for these facts, to the effect that domain restrictions can be expanded but not contracted. The non-monotonic analysis could presumably make do with a stipulation too. Consider salience first. The deviance of (23b) could be explained by postulating that an object that has been mentioned has thereby become salient, and thus that the neighbor's dog becomes the most salient dog in the domain of discourse as soon as the first conjunct of (23b) is uttered. As a result, the expression *the dog* in the second conjunct must denote the most salient dog, i.e. the neighbor's dog, and the sentence ends up being contradictory. In

order to apply the same reasoning to (22b), we need to assume that the similarity measure between worlds is affected by the salience of worlds mentioned in the previous discourse. It must be granted, of course, that there is currently no independent motivation for this assumption.

(iii) *The argument from NPI licensing.* Finally, Fintel 1999 argues that the monotonic analysis is further supported by the behavior of Negative Polarity Items (NPIs) in conditionals:

(24)   a. If John subscribes to any newspaper, he is probably well informed.
        b. If he has ever told a lie, he must go to confession.
        c. If you had left any later, you would have missed the plane.

NPIs are apparently licensed in the antecedent of conditionals. This follows on the assumption that (a) conditionals are structures of universal quantification over possible worlds, whose restrictor is provided by the *if*-clause, and that (b) NPIs are licensed in downward-entailing environments (Ladusaw–Fauconnier Generalization). By contrast, non-monotonic analyses lack an explanation of these facts, since they are designed to ensure that an *if*-clause does *not* create a downward-entailing environment.

   How serious is the problem? It all depends on one's assessment of the Ladusaw-Fauconnier generalization. As it happens, it faces important problems with other examples, such as the following, discussed in Heim 1984:

(25)   a. #Most mountaineers with any experience (still) need a guide for this tour.
        b. Most men with any brains eat rutabagas.

(25a) is entirely expected, since the restrictor position N' of *Most N' I* does not provide a downward-entailing environment. On the other hand, (25b) is unexpected on any standard analysis. Heim suggests that the relevant generalization has to do with a limited form of downward-entailingness *in the presence of certain background assumptions*. Thus the contexts in which (25b) is felicitous are ones in which it is assumed that, *a fortiori*, most men with *a lot of brains* also eat rutabagas [by contrast, one does not normally assume that most mountaineers with a lot of experience need a guide, which accounts for the deviance of (25a)]. Heim applies the same strategy to conditionals in general, and to the following contrast in particular:

(26)   a. If you read any newspaper at all, you are well informed.
        b. #If you read any newspaper at all, you remain quite ignorant.

Heim observes that *despite* the non-monotonicity of conditionals, one typically has background assumptions that ensure that *If you read (at least) n newspapers, you are well informed* entails for $n' > n$ that *If you read (at least) n' newspapers, you are well informed* – the background assumption might be something like: 'the more newspapers you read, the better informed you are'. In other words, background assumptions ensure a limited form of downward-entailingness for the antecedent of (26a). By contrast, no plausible background

assumptions entail that 'the more newspapers you read, the more ignorant you remain', and as a result the antecedent of (26b) does not even display a limited form of downward-entailingness, which explains that the NPI *any* is not licensed. The facts in (25) and (26) can thus be naturally accounted for by appealing to a refined notion of downward-entailingness. Once this move is made, the non-monotonic analysis of conditionals can in principle be made compatible with this modified version of the Ladusaw–Fauconnier Generalization (although of course the precise definite definition of the 'refined notion' of downward-entailingness should be made much more precise).

Does this analysis extend to definite descriptions? At first glance, the facts might appear to be different from those we observed in conditionals, since definite descriptions often *fail* to license NPIs, e.g. in # *The students who knew anyone had a good time at the party*. However contrasts noted by Heim for conditionals can to some extent be replicated with definite descriptions:

(27)   a. ?The Ling 1 students who understood anything at all got an A.
        b. #The Ling 1 students who understood anything at all still got bored.

(28)   a. The students who read any newspapers at all are well-informed.
        b. #The students who read any newspapers at all remain quite ignorant

I conclude that the facts of NPI licensing do not necessarily show that there is a real difference between definite descriptions and conditionals.[6]

(iv) *A monotonic analysis of if-clauses as definite descriptions.*   Let us now suppose for a minute that the monotonic analysis of conditionals is in fact correct, contrary to what we will assume from Section 3 on. The monotonic analysis could still be implemented in two ways:

(i)   By postulating (following much of the literature, e.g. Fintel, 1999, 2001) that the *if*-clause restricts a universal quantifier.

(ii)   By analyzing the *if*-clause as a *standard* (monotonic) definite description, i.e. one that picks out the maximal set that satisfies its restrictor.

Within a monotonic framework, Schein (2001) offered an independent argument for (ii) and against (i), one that crucially relied on *plural* descriptions (plurality is incorporated to our non-monotonic analysis in Section 3.2). In simple cases, introducing a plural definite description as the restrictor of a generalized quantifier would seem to be harmless, as shown in the following:

(29)   a. Each student is happy
        a′. [$\forall$x: student(x)] (happy(x))
        b. Each of the students is happy
        b′. [$\forall$x: [ιX: STUDENT(X)] Xx)] (happy(x))

The analysis in (29a′) is standard. The analysis in (29b′). is more indirect, since the restrictor of 'every' is obtained by first singling out a plural object

(capital letters are used for higher-order predicates and variables). In this example the final truth-conditions are the same in both cases. In more complicated examples, however, an important difference does arise. With an eye to what is to come, I illustrate it with preposed definite descriptions in French, which will be seen to be similar to preposed *if*-clauses:

(30)   a. Les Français, ceux que je connais sont pour la plupart sympathi-
          ques.
          The French, those that I know are for the most part nice
          'As for the French, those I know are mostly nice'
       b. $\iota X' : \text{FRENCH}(X')][\iota X : X \subseteq X' \,\&\, \text{I-KNOW}(X)]$
          $[\text{MOST x: Xx}]((\text{nice}(x))$
       c. $[\text{Most x} : \text{French}(x)\&\text{I-know}(x)](\text{nice}(x))$
          'Most Frenchmen I know are nice'

Here two definite descriptions have been stacked at the beginning of the sentence. The truth-conditions should come out as in (c), which only involves first-order quantification. But the syntax of (c) is implausible because 'mostly' had to be moved to the beginning of the sentence, while the two restrictors had to be conjoined. In this case the solution is to give a plural analysis as in (b), which derives the correct truth-conditions.[7]

However, when the problem is transposed to quantification over possible worlds, the solution becomes less obvious because English does not have overt markers of plurality for worlds. The problem is Barker's puzzle about iterated *if*-clauses (Barker, 1997), and the solution is Schein's (2001), who resorts to plural quantification over events rather than possible worlds, as is done in the present paper. Here is a simplified version of the crucial examples:

(31)   a. If John comes, if Mary comes as well, the party will probably be a
          disaster.
       b. $[\iota W' : \text{JOHN-COMES}(W')][\iota W: W \subseteq W' \,\&\, \text{Mary-comes}(W)]$
          $[\text{Most w: Ww}](\text{disaster}(w))$
       c. $[\text{Most w} : \text{John-comes}(w) \,\&\, \text{Mary-comes}(w)](\text{disaster}(w))$

As in the preceding case, the final truth-conditions should be those of (c) (at least as a first approximation). But the logical form in (c) is implausible because it involves a drastic re-arrangement of various parts of the sentence. By contrast, the analysis in (b), which crucially relies on plural definite descriptions of worlds, is syntactically natural. Thus even if the monotonic analysis of conditionals is correct, *if*-clauses should be analyzed as (monotonic) plural descriptions rather than as structures of universal quantification.[8]

3. A Choice Function Analysis

In the rest of this article we tentatively assume that the non-monotonic analysis of conditionals and definite descriptions is on the right track, and we seek to implement it and explore some of its consequences.

3.1. CHOICE FUNCTIONS ACROSS DOMAINS (THE SINGULAR CASE)

Stalnaker (1968), who assumed that there *was* something to explain about non-monotonicity above and beyond domain-restriction, introduced the device of *Selection functions* to handle the problem. Intuitively, *if* ϕ is taken to 'select' the world most similar to the actual world which satisfies ϕ (if there is no such world we will assume that a presupposition failure occurs, although this isn't Stalnaker's analysis; see below). *If* ϕ, ψ is then taken to be true just in case the world selected by *if* ϕ satisfies ψ (this is simply a case of predication). And as was mentioned above, von Heusinger's Choice functions are simply a weakened version of Selection functions. By taking literally the suggestion that *if* is *the* applied to worlds, we obtain either Stalnaker's analysis (singular definite descriptions) or a strengthened version of Lewis's system (plural definite descriptions). We start our discussion with Stalnaker's Selection Function analysis, which we apply to *if*-clauses and definite descriptions alike. We then extend this system by introducing functions that select a plurality of objects.

3.1.1. *General Format*

We write *the* and *if* as ι. When ι is followed by an individual variable, it represents *the*; when it is followed by a world variable, it represents *if*. ι always selects the element that is closest to a given element under some pre-established linear ordering. What is this 'given element'? If we didn't have to worry about embeddings, we could simply assume that, both for definite descriptions and for conditionals, it is the context of utterance. But this won't do in the general case for conditionals, which can be recursively embedded:

(32)   If John were here, if Mary were here as well, the party would be a lot of fun.

Clearly the context of utterance is the same throughout the discourse. However the second *if*-clause should not be evaluated from the standpoint of the actual world w^*, but rather from the world selected by $f(w^*, [\![John is here]\!])$. In the general case, then, Stalnaker's Selection functions must take two arguments: a world of evaluation and a set of worlds.[9]

Turning now to definite descriptions, it would seem that there is no reason to provide the Choice function with two arguments (an individual and a predicate extension) rather than just one (a predicate extension). This is because on a superficial analysis *the P* might be taken to denote an element which is salient *in the context of speech*. However when more complex examples are taken into account, this treatment can be seen to be too crude:

(33)   a. The dog got into a fight with another dog (McCawley, 1979)
        b. John said that the dog had gotten into a fight with another dog.
        c. (?) All the people I know own a dog. At some point or other, each of my friends realized that the dog had gotten into a fight with another dog.

In (33a) *the dog* may be taken to denote the most salient dog relative to the speaker. In (33b), by contrast, it is plausible that *the dog* denotes the most salient dog *relative to John* rather than relative to the speaker. (33c) makes the same point more forcefully, since for various dog-owners the most salient dog is of course not the same.[10] The data are correctly handled by binding the additional argument of the ι-operator to a quantifier, as in the following simplified representation, where I have also included a contextual restriction C that depends on the variable x:

(34). $[\forall x : \mathrm{friend}(x)] \ldots [\iota_x y : \mathrm{dog}(y)\&C(x)] \ldots$

Thus there appears to be independent motivation for providing Choice Functions that are used to model salience with an additional individual argument.[11]

3.1.2. *Stalnaker's First Three Conditions*

Let us now consider Stalnaker's Selection Functions and see whether and how they may be used to model the behavior of definite descriptions as well. Stalnaker (1968) imposed four conditions on his selection functions, which we discuss in turn.

Condition 1 [= Stalnaker's Condition (1)]: For each element d and each non-empty set E, $f(d, E) \in E$. This is of course the condition that makes Stalnaker's Selection Functions a variety of Choice Functions. For definite descriptions, this means that the individual denoted by *the* ϕ must satisfy the predicate ϕ. For conditionals, this means that the world denoted by *if* ϕ must satisfy the sentence ϕ. Both conditions are uncontroversial.

Condition 2 [equivalent to Stalnaker's Condition (4) when Condition 1 holds]: For each element d and any sets E and E′, if $E' \subseteq E$ and $f(d, E) \in E'$, then $f(d, E') = f(d, E)$. When applied to individuals under a measure of salience, this can be paraphrased as: if some element is the most salient among all the members of E, then it is also the most salient among some subset of E that includes it. Although this condition has not (to my knowledge) been discussed in the Choice Function literature, it is necessary to ensure that the Choice Function indeed models a notion of *maximal* salience. More generally, this condition must apply whenever a function is supposed to select from a set the element(s) with the greatest degree of a property P (on Stalnaker's analysis, P is the degree of similarity to the world of evaluation).

The foregoing discussion was a slight distortion of the history, however. Stalnaker (1968) discusses in fact a different version of the condition, but it turns out to be equivalent to the present version given Condition 1. Stalnaker's original condition is:

Condition 2′ [= Stalnaker's Condition (4)]: For each element d and any sets E and E′, if $f(d, E') \in E$ and $f(d, E) \in E'$, then $f(d, E) = f(d, E')$.

Claim: Condition 2 and Condition 2′ are equivalent given Condition 1.
Proof: (i) Condition 2 $\Rightarrow$ Condition 2′. Suppose f(d, E′)$\in$E and f(d, E)$\in$E′. By Condition 1, f(d, E′)$\in$E′ and hence f(d, E′)$\in$E∩E′; and similarly f(d, E)$\in$E and hence f(d, E)$\in$E∩E′. By Condition 2, f(d, E)=f(d, E∩E′)=f(d, E′).

(ii) Condition 2′ $\Rightarrow$ Condition 2. Suppose E′$\subseteq$E and f(d, E)$\in$E′. By Condition 1, f(d, E′)$\in$E′, and hence f(d, E′)$\in$E since E′$\subseteq$E. Thus f(d, E)$\in$E′ and f(d, E′)$\in$E. By Condition 2′, f(d, E)=f(d, E′).

Condition 3 [=Stalnaker's Condition (2)]: For each element and each set E of elements, f(d, E)=# iff E = $\emptyset$.

Here I will interpret # as symbolizing referential failure, as is natural for definite descriptions on a Strawsonian view. Stalnaker, who was solely interested in conditionals, did not allow for referential failure. Rather, he interpreted # (which he wrote as λ) *as* 'the *absurd world* – the world in which contradictions and all their consequences are true'.[12] In Stalnaker's system any proposition is true of the absurd world. As a result, a conditional with an impossible antecedent is deemed true no matter what the consequent is (this is because in that case *if* φ, ψ is true just in case the world selected by *if* φ i.e. λ, satisfies ψ; but λ satisfies every proposition). By contrast, on the present view a conditional with a *clearly* impossible antecedent is deemed infelicitous. Thus *if* φ fails to refer just in case there is no world whatsoever, even a particular distant one, which satisfies φ. This condition does not appear to be too far-fetched given the infelicity of sentences such as: #*If John were and weren't here, Mary would be happy*.[13]

3.1.3. *Stalnaker's Last Condition: Centering*

Stalnaker's Selection Functions were defined by a fourth condition, which does not appear to be plausible for definite descriptions:
Condition 4 [=Stalnaker's Condition (3)]: For each element d and each set E, if d$\in$E, then f(d, E)=d.

Applied to possible worlds, this condition states that *if* φ always selects the world of evaluation in case φ is true in that world. For instance, in an unembedded environment, this means that the *if*-clause must always select the actual world if it happens to satisfy the antecedent. This condition is entirely natural when the ordering of two elements is defined by their similarity to the actual world, as is the case for *if*-clauses. The condition seems less plausible for definite descriptions, where elements are ordered by their relative salience from the standpoint of a particular individual. If Condition 4 were applied in this domain, it would require that an unembedded definite description should always pick out the speaker if she happens to satisfy the restrictor of the description. Clearly, this is an overly egocentric view of communication. I may certainly use the description *the guy in the white shirt* to refer to John even if I myself happen to be wearing a white shirt. But if

Condition 4 held of descriptions of individuals, 'the guy in the white shirt' would, of necessity, denote *me*.

This, however, is a problem only so long as it is assumed that the speaker necessarily serves as the default point of evaluation for the Choice/Selection Function used to model salience. But this might well be too strong. In fact, when I utter a sentence I typically take for granted some perceptual situation which I might not be a part of (this should be clear if the perceptual situation is my own visual field, since I don't typically see myself, at least not entirely). Obviously no such perceptual condition holds in the domain of worlds (because these can't be perceived). As a result, the world of evaluation in a standard speech situation c is the one which is most relevant to c, i.e. the world of (c). If this analysis is correct, the reason 'the man in the white shirt' may fail to refer to me even if I happen to be wearing a white shirt is *not* that the centering condition fails; rather, it is that the point of reference in that situation isn't the speaker himself but whatever is at the center of the speaker's perceptual field.

One final point is in order. Although in standard cases the world of utterance provides a default value for the world argument of an unembedded *if*-clause, this needn't always be the case. Consider the following example of modal subordination:

(35)   Suppose I didn't exist. If you were the same person that you actually are, you would be much happier[14]

It is clear in this case that the actual world w* satisfies the antecedent of the conditional. Yet the *if*-clause certainly does not select the actual world, for if this were the case the sentence would be trivially false (it couldn't be that you are happier in w* than you are in w*). The natural interpretation is not that Condition 4 is suspended; rather, it is that the point of evaluation is not w* but one of the worlds in which the initial assumption (*Suppose I didn't exist*) is met. This is of course parallel to the reasoning we just made for definite descriptions.

To conclude this part of the discussion, let us see how the system works in the example of 'Strengthening of the Antecedent'. The logical forms are now as follows (I use a cross-sortal ι and overt variables for each sort; w* represents the world of evaluation – typically the world of the context, and x* represents the individual of evaluation, which will often be different from the speaker. For simplicity, I omit time/world variables in (a'). and time/individual variables in (b'):

(36)   a. The pig is grunting, but the pig with floppy ears isn't grunting.
       a'. grunting($[\iota_{x*}x: pig(x)]$) but ¬grunting($[\iota_{x*}x: pig(x)$ & floppy-ears(x)]$)
       b. If this match were struck, it would light, but if this match had been soaked in water overnight and it were struck, it wouldn't light
       b'. light($[\iota_{w*}w: strike(w)]$) but ¬light($[\iota_{w*}w: strike(x)$ & soak(w)]$)

Since the most salient *pig with floppy ears* need not be the most salient *pig*, (a′) is not a contradiction. For the same reason the closest (most similar) world in which the match is soaked in water and struck might not be the closest world in which it is struck, and thus (b′) is also predicted to be consistent.

3.2. ADDING PLURALITY

Stalnaker's analysis has a major defect: it does not allow for quantification over possible worlds. This is problematic in view of examples such as 'Necessarily, if John comes, Mary will be happy' or 'Probably, if John comes, Mary will be happy'. A major insight in the recent history of conditionals was that in these examples *if*-clauses restrict generalized world or event quantifiers (Lewis, 1975). But if *if*-clauses are construed as *singular* descriptions, there is no way this can be done. Part of the solution is to treat *if*-clauses as *plural* definite descriptions, and to analyze generalized world quantification by analogy with partitives[15] (this is only *part* of the solution because, as was discussed in Section 2.3.2, one still needs to explain why *necessarily, if p, q* appears to display a monotonic behavior even though *if p, q* does not). As was mentioned above, Schein (2001) gives independent evidence for such an analysis, based on the recursion of *if*-clauses (see also Stone, 1997 for a different use of plurality in the analysis of mood and conditionals). On the assumption that *if*-clauses are plural descriptions, we can analyze 37(a–b) by analogy with 37(a′–b′):

(37)   a. Probably, if Mary comes, John will be happy.
         b. If Mary comes, John will probably be happy.
         a′. Most of the students are happy.
         b′. As for the students, most of them are happy.

The resulting theory is analogous to a strengthened version of Lewis's Logic, one which is discussed at some length by Lewis himself in *Counterfactuals* (see also Nute, 1980).
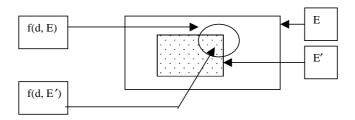
3.2.1. *Modifying Stalnaker's Conditions*

What happens, then, when plurality is incorporated to the theory? Plural Choice functions are just another name for what Nute 1980 calls 'Class selection functions'. The literature on conditionals is rife with constraints that can be imposed on these to yield various logics. I will not attempt to do justice to these suggestions, but only to point out (i) that the same constraints might in principle be imposed on salience-based theories of plural definite descriptions, and (ii) that it might be heuristically interesting to explore the possibility (which is our working assumption) that exactly the same constraints are at work for all plural Choice functions, whether they apply to *if*-clauses or to plural definite descriptions. If we wish to preserve as much as

possible of Stalnaker's original intuitions in this weaker framework, we may impose (tentatively) the following conditions:

Condition 1*: For each element d and each non-empty set E of elements, $f(d, E) \neq \#$ and $f(d, E) \subseteq E$.

Condition 2*: For each element d, each set E and each set E', if $E' \subseteq E$ and $f(d, E) \cap E' \neq \emptyset$, then $f(d, E') = f(d, E) \cap E'$.

This modification of Condition 2 is natural given that our Choice functions are supposed to select the *most salient* individuals and the *most similar* worlds in a given context. Consider the following, which represents a situation in which $f(d, E) \cap E' \neq \emptyset$:

– Clearly, if some elements of E' are the most highly ranked (with respect to salience or similarity) in the superset E, they should count as the most highly ranked among the elements of E'. This yields the inclusion: $f(d, E) \cap E' \subseteq f(d, E')$.

– Conversely, the elements of $f(d, E)$ are more highly ranked than any other element of E, hence also more highly ranked than any other element of the subset E'. Since $f(d, E) \cap E' \neq \emptyset$, any element of $E' - f(d, E)$ must be less highly ranked than the elements of $E' \cap f(d, E)$, and hence couldn't belong to $f(d, E')$. This yields the inclusion: $f(d, E') \subseteq E' \cap f(d, E)$.

As Ede Zimmermann (p.c.) points out, Condition 2* together with Condition 1* and a reinterpretation of Condition 3* turns out to be equivalent to the choice being based on a universal, centered ordering. Zimmermann's argument is laid out in Appendix II. In fact, as he suggests, the entire story could just as well be stated directly in terms of orderings, though this would obscure the historical connection with Stalnaker's Selection Function analysis and von Heusinger's Choice Function analysis of definite descriptions.

For convenience, I repeat Condition 3, which may be retained without change, and I adapt Condition 4, which will be useful for conditionals:

Condition 3* (= Condition 3): For each element d and each set E, $f(d, E) = \#$ iff $E = \emptyset$.

(We could decide to interpret # as $\emptyset$, i.e. to assume that a plural description whose restrictor is empty denotes the empty set. This is the assumption we make for technical reasons in Appendix II, though for linguistic purposes it is probably better to analyze # as a presupposition failure).

Condition 4*: For each element d and each subset E of the domain, if d ∈ E, then d ∈ f(d, E).

3.2.2. *Comparisons*

The comparison with Stalnaker's system is straightforward: the extension to plural Choice Functions allows us to leave out the requirement that similarity or salience should always be so fine-grained as to yield a single 'most salient' individual or a single 'most similar' world. We also weaken the resulting logic. In Stalnaker's initial system, *because if* φ denotes a single possible world, for any proposition you care to mention either that world is in that proposition or it isn't. As a result, *[if φ ψ]∨ [if φ¬ψ]* comes out as a logical truth (Conditional Excluded Middle). This principle fails when plural choice functions are used. If some of the worlds denoted by *if* φ satisfy ψ while others don't, it will neither be the case that *[if φ ψ]* nor that *[if φ¬ψ]*.[16]

Lewis (1973) considered a system related to the analysis based on plural Choice Functions. But he dismissed it as too strong. In our terms, the problem is that the analysis predicts referential failure whenever it is not possible to select a set of worlds that are 'the closest' among those that satisfy the antecedent. How could this situation ever arise? Consider the following example:

> Suppose we entertain the counterfactual supposition that at this point _____ there appears a line more than an inch long. (Actually it is just under an inch.) there are worlds with a line 2″ long; worlds presumably closer to ours with a line 11/2″ long; worlds presumably still closer to ours with a line 11/4″ long; worlds presumably still closer …. But how long is the line in the *closest* worlds with a line more than an inch long? If it is $1 + x''$ for any x however small, why are there no other worlds still closer to ours in which it is $1 + 1/2x''$, a length still closer to its actual length? (…). Just as there is no shortest possible length above 1″, so there is no closest world to ours among the worlds with lines more than an inch long (…)″. (Lewis, 1973, p. 20)

It isn't clear that Lewis's example is cogent; nor does his solution to the problem seem satisfactory. As pointed out by Stalnaker and others (see Harper, 1981; Nute, 1980), if such a fine-grained measure of similarity were *ever* available it would according to Lewis make the following sentence true for any value of ε (e.g. ε = 0.1″, 0.01″, 0.001″, etc.): ⌈If this line were longer than 1″ long, it would be smaller than $(1 + ε)''$ long⌉.[17] This is clearly an undesirable result, since for small enough ε's these sentences are certainly not judged to be true.

In order to address this objection, Lewis has apparently suggested that the *coarseness* of the similarity measure may vary, so that in the example at hand all worlds in which, say, the length is between 1″ and 5″ count as equally similar to one another (Nute, 1980 cites conversations with Lewis on a

different but structurally similar example). But if so the same device can be used to save the plural Choice function analysis (i.e. the class function analysis) from Lewis's purported counterexample: we may simply stipulate that the similarity measure is never as fine-grained as the difference between the length of the line at w and at w*. Coarseness may save Lewis from trouble, but it also saves the plural Choice Function analysis from Lewis.

In addition, it is highly unclear how Lewis's system can be extended to cases of generalized quantification, which require that some value be given to the *if*-clause so it can restrict the generalized world quantifier (e.g. 'probably', 'there are 30% chances', etc.). The beginning of a solution can be found if we use the plural Choice Function analysis, since a plural definite description of worlds may restrict a generalized world quantifier, just as 'the students' may restrict 'most' in 'Most of the students came'. As was mentioned above, the only remaining problem is to insure that somehow the measure of salience becomes trivial when a quantifier (e.g. *necessarily*) is applied to *if*-clause. I leave this as an open problem.

4. Consequences of the Referential Analysis (I): Topic, Focus and 'then'

4.1. TOPICALIZATION OF THE *if*-CLAUSE, FOCALIZATION OF 'THEN'

The referential analysis of *if*-clauses can now be used to derive a number of interesting syntactic and semantic facts. First, it has been observed that *if*-clauses can appear in the position of sentence topics, in a left-dislocated position. This should now come as no surprise, since referential elements can quite generally be dislocated in this fashion. By contrast, quantifiers or simple restrictors may not be:

(38)   a. *Every man, he is happy
       b. *Man, every is happy

Bittner (2001), who develops similar ideas, gives a further argument from Warlpiri. Following Hale 1976, she observes that the following sentence is ambiguous:

(39)   Maliki-rli *kaji*-ngki yarlki-rni nyuntu
       [dog-ERG ST-3SG.2SG bite-NPST you]
       ngula-ju kapi-rna luwa-rni ngajulu-rlu.
       DEM-TOP FUT-1SG.3SG shoot-NPST me-ERG
       a. Reading A. 'As for *the* dog that bites you, I'll shoot *it*.' (individual-centered)
       b. Reading B. '*If* a dog bites you, *then* I'll shoot it.' (possibility-centered)

Bittner writes:

> The dependent clause of [(39)] – with the complementizer *kaji*, glossed 'ST' for 'same topic' – introduces a topical referent of some type. On reading

(A) the topic is a contextually prominent individual, and on reading (B), a prominent possibility. In either case, the topical referent is picked up in the matrix comment by a topic-oriented anaphoric demonstrative *ngula-ju*, which is likewise type-neutral. So depending on the context, the topic of [(39)] may be either the most prominent dog which bites the addressee or the closest possibility that a dog may bite. The fact that one and the same sentence can have both of these readings suggests that they have essentially the same semantic representation, up to logical type.

Bhatt and Pancheva (2001) further observe that there is a strong syntactic similarity between conditionals and correlative constructions, which 'involve a free relative clause adjoined to the matrix clause and coindexed with a proform inside it'. In fact, in many languages *if*-clauses are overtly correlative structures themselves. Bhatt and Pancheva suggest that, quite generally, *if*-clauses are free relatives, i.e. definite descriptions of possible worlds, and that the word *then* is a world pronoun (in some languages, for instance Marathi, *then* appears to be morphologically related to other pro forms). The analysis of *then* as a world pronoun has also been proposed by van Benthem, Cresswell (1990), Iatridou (1994) and Izvorski (1996). In particular, Iatridou (1994) suggests that this could derive the semantic/pragmatic restrictions on the distribution of *then*:

(40) a. If Peter runs for President, the Republicans will lose.
      b. If Peter runs for President, then the Republicans will lose.
(41) a. If John is dead or seriously ill, Mary will collect the money.
      b. If John is dead or seriously ill, then Mary will collect the money.
      a′. If John is dead or alive, Bill will find him.
      b′. #If John is dead or alive, then Bill will find him.

Although superficially there is no difference between a sentence with and without *then* [e.g. (40a,b)], on closer inspection some subtle contrasts do arise, as illustrated by the deviance of (41b′). Iatridou's suggestion is that the presence of *then* triggers a presupposition or an implicature which she analyzes as follows, where O is a generalized quantifier over worlds (e.g. 'necessarily', 'possibly', etc.), '[p]' is its restrictor, and 'q' its nuclear scope:

(42) Iatridou's analysis of *O, if p, then q*
      a. Assertion: O[p] q
      b. Presupposition/Implicature: ¬O[¬p] q

A bare *if*-clause is analyzed as restricting a covert universal quantifier. The presupposition/implicature in (41b′) is thus that there is a possible world in which John is neither dead nor alive and in which we find him. But there are certainly no such worlds (John is either dead or alive but not both), which accounts for the deviance of this example.

Iatridou further attempts to relate this effect to the presupposition/implicature found in constructions that involve left-dislocation and doubling in

German. The following typically implicates that someone other than Hans failed to understand:

(43)   Hans$_i$, der$_i$ hat es verstanden. (German; Iatridou, 1994)
        Hans, he has it understood
        a. Assertion: P(i)
        b. Presupposition/Implicature: [∃j: j≠i]¬P(j)

Iatridou suggests that the doubled pronoun plays the same role as *then* in the preceding examples. Although these suggestions are illuminating, Iatridou doesn't really explain how *if*-clauses and dislocated noun phrases may be treated in a unified framework. Part of the problem is that on her analysis an *if*-clause is a restrictor rather than a referential element. This makes the comparison with left dislocated referential noun phrases harder to make. On the present approach, the difficulty disappears since *if*-clauses are analyzed as being referential.

Iatridou further observes, however, that the analogy between *then* and *der* in German breaks down in the following example, due to I. Heim:

(44)   Alle haben die Vorlesung verstanden. Hans hat sie verstanden. Maria hat sie verstanden. Und unser Freund Peter, der hat sie auch verstanden. (German)
        'Everybody understood the lecture. John understood it. Mary understood it. And our friend Bill, he understood it too'.

Iatridou's prediction is that (44) should be infelicitous because the implicature contradicts the assertion. But this is not the case. Iatridou leaves the problem open. The difficulty can be solved by suggesting that *then* isn't just any pronoun, but rather a *strong* pronoun. In languages that distinguish between weak and strong pronouns, Heim's facts can be replicated with the former but not with the latter. Strong pronouns thus pattern in the way predicted by Iatridou's analysis, as is illustrated by the following contrasts in French:

(45)   Everybody understood. The professors understood, the staff understood, and ...
        a. #[Les étudiants]$_i$, eux$_i$ ont compris aussi
           [The students]$_i$, them$_i$-strong$_F$ have understood too
        b. [Les étudiants]$_i$, ils$_i$ ont compris aussi
           [The students]$_i$, they$_i$-weak$_F$ have understood too

Thus the behavior of the strong pronoun *eux* appears to be similar to that of the world pronoun *then*. With *eux*, the implicature is that some non-students didn't understand, hence the infelicity of the above example (since everybody understood). In the case of *then* in *if p, then q*, the implicature is that some non-*p* worlds are non-*q* worlds, which accounts for (41b').

Izvorski (1996) explores a different line of analysis to overcome this problem. She suggests that the reason (44) is grammatical is that the focus-particle *too* has been added. Adding the same particle to Iatridou's 'then' examples makes them felicitous under the same conditions:

(46)   We will definitely play soccer. If the sun shines we will. If it is cloudy and cold we will. And if it rains (#then) we will/And ⟨also⟩ if it rains *then* too we will.

As it happens, this observation meshes well with the suggestion that *then* is a *strong* world pronoun. For in Izvorski's example *too* is added right after the word *then*. If *aussi* (also) is placed right after *eux* in (45a), the example improves. And if *too* is placed lower down in the structure in Izvorski's example, it becomes worse:

(47)   Everybody understood. The professors understood, the staff understood, and...
    a. [Les étudiants]$_i$, eux$_i$ aussi ont compris
       [The students]$_i$, them$_i$-strong$_F$ too have understood
    b. We will definitely play soccer. If the sun shines we will. If it is cloudy and cold we will. #And if it rains *then* we will too.

Let us now see whether Iatridou's and Izvorski's observations can be derived (both for individual and for world pronouns) from an independently motivated mechanism. In order to achieve this, I simplify the problem and consider only the case of a focused pronoun, disregarding the left-dislocated element. Given the referential analysis of *if*-clauses, the semantic value of *then* should be precisely that of its antecedent, which makes the simplification relatively harmless.

Let us first consider a standard example. Rooth (1996) briefly discusses the following sentence:

(48)   Situation: Steve, Paul and I took a calculus quiz (which was graded on the spot). George asked me how it went.
    a. Well, I$_F$ passed. [Rooth, 1996, 11(b)]
    b. Implicature: Steve and Paul didn't pass

Briefly, Rooth's suggestion is that a focused element introduces a set of alternatives, so that if the subject is focused in 'I$_F$ passed', a set of alternatives is evoked, of the form: {'I passed', 'Steve passed', 'Paul passed', 'Steve and I passed', 'Paul and I passed', 'Steve, Paul and I passed', etc.}. This, in turn, triggers an implicature that the sentence that was in fact asserted was the most informative true sentence among the set of alternatives. Note that the alternatives may involve plural subjects, which is crucial to derive the correct result. Rooth analyzes the alternatives as a set of propositions (i.e. as a set of sets of possible worlds). In order to simplify the treatment of focused world pronouns, I use sets of sentences instead. Rooth's analysis may then be reconstructed as follows, where 'w' is a variable that denotes the actual world (⌜ and ⌝ are Quine's quasi-quotation marks):

(49)   Situation: Steve, Paul and I took a calculus quiz (which was graded on the spot). George asked me how it went.
    a. Well, I$_F$ passed. (Rooth 1996, (11b))
    b. Ordinary value: passed(w, I)
       [Ordinary value: 'I passed']

    c. Focus value: F = {S: for some contextually given denoting (possibly plural) expression E, S = $\ulcorner$passed(w, E)$\urcorner$}
[Focus value: {'I passed', 'Steve passed', 'Paul passed', 'Steve and I passed', 'Paul and I passed', 'Steve, Paul and I passed', etc.}]
    d. Implicature: no sentence in F both (i) entails (a) and (ii) is true.
[Implicature: Among the relevant alternatives, no one but me passed]

Since the alternatives to 'I' in Rooth's example involve plural subjects, if someone other than me, X, had passed, I could have uttered: 'X and I passed' (or: 'we passed'), which would have been both true and more informative than what I in fact uttered. Since I did not utter this more informative sentence, I implicated that only I passed.

This analysis can be extended without difficulty to a case involving focalization of a plural pronoun (I use a French example because the morphological distinction between weak and strong pronouns indicates that in the following the pronoun *must* be focused):

(50)   a. ([Les étudiants]$_i$,) [eux$_i$]$_F$ ont compris [French; 'eux' is the strong pronoun] *([The students]$_i$,) [them$_i$-strong]$_F$ have understood*
    b. Ordinary value: understood(w, they$_i$)
[Ordinary value: 'they$_i$ understood']
    c. Focus value: F = {S: for some contextually given denoting (possibly plural) expression E, S = $\ulcorner$understood(w,E)$\urcorner$}
[Focus value: {'they$_i$ understood', 'I and they$_i$ understood', 'Paul, Steven and they$_i$ understood', etc}]
    d. Implicature: no sentence in F both (i) entails (a) and (ii) is true.
[Implicature: Among the relevant alternatives, no one except them$_i$ understood]

The implicature is that none of the contextual alternatives to the plural individual denoted by 'the students' did in fact understand, for otherwise a more informative sentence could have been asserted (namely: 'The students *and X* understood'). There are now two ways in which the implicature could lead to infelicity:

– If there are other (salient) people and they also understood, the implicature will simply be false.
– If there is no one else in the domain of discourse, the focus value will only contain sentences that are entailed by the sentence that was in fact asserted. This makes the implicature idle, in a way that appears to lead to deviance. Thus if there were only students in the audience of a particularly complicated thriller, it won't do to say: 'Les étudiants, eux ils ont compris' (*the students, them-strong they understood*). For this would imply that other members of the audience, non-students, didn't understand. But there are no non-students in the audience. This observation motivates the following condition:

(51)   Non triviality: some element of the Focus value should not be entailed by the asserted sentence.

Applied to conditionals, this appears to account for the deviance of (41b′), repeated as (52) and analyzed as in (53) (W is a variable over pluralities of worlds):

(52)   #If John is dead or alive, then$_F$ Bill will find him.

(53)   a. ([If p]$_i$,) [then$_i$]$_F$ q
       b. Ordinary value: q(then$_i$)
          [Ordinary value: q holds in the worlds denoted by 'then$_i$']
       c. Focus value: F = {S: for some contextually given denoting expression E, S = $\ulcorner$will-find-him(E)$\urcorner$}
          [Focus value: {'if John is dead, Bill will find him', 'if John is alive, Bill will find him', etc}]
       d. Implicature: no sentence in F both (i) entails (a) and (ii) is true.

Because the antecedent is a tautology, the condition of non-triviality is violated, hence the deviance of the example. (Note that on this analysis this sentence is not exactly parallel to the 'Hans' example given above, as was suggested by Iatridou; rather, the correct point of comparison is the thriller example we just discussed).

4.2. CONDITION C EFFECTS

The referential analysis allowed us to explain why *if*-clauses can be dislocated and doubled by the word *then*, construed as a world pronoun. But if this theory is on the right track, we would expect world pronouns and world descriptions to share other formal properties of pronominal and referential expressions. In the domain of reference to individuals, there are well-known constraints on the syntactic distribution of such elements, summarized in Chomsky's 'Binding Theory'. We now attempt to show that one of these conditions, the strong form of 'Condition C', applies to world expressions as well. This can be seen as an extension of an analysis made by Percus (2000), who suggested that some world variables must satisfy other syntactic constraints (Percus suggested that some world variables must be bound locally).

Condition C of the Binding Theory states that a proper name or definite description (R-expression) may not be bound. Typically violations of the constraint are relatively mild (and cross-linguistically unstable) when an R-expression is co-indexed with another c-commanding R-expression. By contrast, the violations are very strong (and cross-linguistically stable) when an R-expression is c-commanded by a co-referring pronoun. The latter case is illustrated by the examples in (54), whose structures are given in (55):

(54)   a. John$_i$ likes [people who admire him$_i$]
       b. *He$_i$ likes [people who admire John$_i$]
       c. [His$_i$ mother] likes [people who admire John$_i$]

(55)   a. [R-expression$_i$ [$_{VP}$ [$_{NP}$ . . . pronoun$_i$ . . .]]]
        b. * [pronoun$_i$ [$_{VP}$ [$_{NP}$ . . . R-expression$_i$ . . .]]]
        c. [[. . . pronoun$_i$ . . .] [$_{VP}$ [$_{NP}$ . . . R-expression$_i$ . . . .]]]

As is shown in (54c–55c), a pronoun may in some cases be coindexed with a referential expression that follows it (backwards anaphora). However this is impossible if the referential expression is in the scope of (c-commanded by) the pronoun, as in (54b). Exactly the same pattern can be replicated with *if*-clauses, construed as referential terms, and the word *then*, analyzed as a world pronoun:

(56)   a. [if it were sunny right now]$_i$ I would see [people who would then$_i$ be getting sunburned].
        b. *I would then$_i$ see [people who would be getting sunburned [if it were sunny right now]$_i$].
        c. Because I would then$_i$ hear lots of people playing on the beach, I would be unhappy [if it were sunny right now]$_i$

All the examples make reference only to the time of utterance, which ensures that *then* is interpreted modally, not temporally (this is because the word *now* rather than *then* must be used to refer to the time of utterance). It is plausible that *then* and an *if*-clause are adjoined somewhere below IP and above VP. This yields exactly the same schematic structures as in (55). The natural conclusion is that *if*-clauses, as other referential expressions, are subject to Condition C of Chomsky's Binding Theory.

5. Consequences of the referential analysis (II): referential classification

Referential expressions may appear with features that indicate how their denotation is situated with respect to the context of speech (context-relative classification) or with respect to some other salient entity (object-relative classification). The first case is illustrated by the contrast between *I*, which must denote the speaker, and *he*, which normally may not do so. Similarly, the word *this* must denote an entity which is close to the speaker, while the word *that* must normally denote an entity which is further away. The second case (object-relative classification) requires that we consider more exotic languages. In Algonquian, two sorts of third person agreement markers are distinguished. A 'proximate' expression must denote a salient entity which is the center of reference in a stretch of discourse. By contrast, an 'obviative' expression denotes some other entity, which is less salient than the referent of the proximate term. In English an object-relative system appears to be used in the temporal domain, where a time variable with pluperfect features must denote a moment which is prior to some other salient past moment.

In this section I suggest that the distinction between indicative and subjunctive conditionals should be analyzed by analogy with that between *this* and *that*. To simplify the analysis I assume that descriptions of worlds are

always singular, although for reasons that were discussed earlier plural descriptions should probably be countenanced as well. I suggest in this simplified framework that an indicative *if*-clause must denote a world which is in the Context Set (i.e., a world which is compatible with what the speaker and hearer presuppose), and thus counts as 'close enough' to the context of utterance. A subjunctive *if*-clause normally denotes a world which is further away. This is almost *exactly* the analysis offered in Stalnaker (1975), except that now the theory is embedded within a general system of referential classification which applies to individuals and worlds alike (similar suggestions with respect to the tense/mood analogy are made in Iatridou, 2000). I then extend this theory to examples involving two layers of subjunctive morphology (double subjunctives), which are analyzed as an instance of object-relative classification in the domain of worlds.

5.1. 'THIS' VS. 'THAT'

Consider first the difference between *this* and *that*. It appears that *this* incorporates a presupposition that its denotation should count as close to the speaker. We may be tempted to define the opposite presupposition for *that*, i.e., that *that* may *not* denote an entity which is close to the speaker. This is too strong, however. Watching a scene in the mirror, I may find out that a table I was observing is in fact the table standing right next to me. I could then utter without presupposition failure: 'That is this!'. The same point carries over to other cases. If David observes through a mirror someone whose pants are on fire, he may at some point exclaim: 'He is me!'. Although it was presupposed all along that *I* referred to the speaker, it was *not* presupposed that *he* referred to a non-speaker (or else David's utterance would have resulted in presupposition failure, contrary to fact).

In order to analyze these facts, I will make use of a device inspired by Dekker (2000). Dekker observes that there can be uncertainty as to whom I am referring to through a use of the word *he*. For instance I may be unsure whether I am referring to John or to Sam, although I know that I am not referring to Peter. This fact can be formally captured by evaluating the sentence with respect to a set of assignments $S = \{s_1, s_2, \ldots\}$ rather than with respect to a single assignment. If I know who I am referring to, all elements of S will assign the same value to the pronoun I used. If I am not entirely sure who I am referring to, S will contain some assignments which, say, assign John to *he*, while others assign Sam to the same pronoun (though no element of S assigns Peter to it, since I know that I am *not* referring to Peter.)

We can then use this system to give an asymmetric definition of the semantics of $this_x$ and $that_y$, where x and y are variables whose value is contextually supplied (by a demonstration, encoded in the assignment function with respect to which the sentence is evaluated). In the case of $this_x$,

the presupposition is that x refers to an object near the speaker. To put it more formally, for every member s of S, it must be the case that s(x) is near the speaker. For *that*$_y$, by contrast, the presupposition is simply that there is a *possibility* that the denotation of z isn't close to the speaker. This just requires that for some *s* in S, s(z) be far from the speaker. Thus no presupposition failure need arise when I utter something like: 'That is this'. This analysis is naturally cast within a dynamic semantics, as is done in Appendix I (again following Dekker's ideas). The set of assignments with respect to which the sentence is evaluated is modified in discourse, so that in uttering 'That$_y$ is this$_x$,' I eliminate from S those assignments in which s(y) ≠ s(x).

5.2. INDICATIVE VS. SUBJUNCTIVE

On Stalnaker's (1975) standard analysis, indicative mood incorporates the (default) assumption that the Selection function picks out a world within the Context Set. From the present perspective, this is just to say that indicative mood expresses a presupposition similar to that of the word *this*, but in the domain of worlds rather than of individuals. The notion 'close to the context of utterance' is rendered, following Stalnaker, as: 'within the Context Set'. The Context Set also plays a second role, which is to model the speaker's uncertainty about what the actual world is. Thus when we compute the reference of a definite description of worlds $[\iota_{w*}w: \varphi]$, we may not know with precision what the value of the point of reference w* is (in other words, there are assignments compatible with the speaker's state of knowledge which assign different values to w*). As a result, there will be uncertainty on the value of $[\iota_{w*}w: \varphi]$, which by definition denotes the closest φ-world(s) to w*. Given our very weak semantics for *that*, we expect that subjunctive marking should only indicate that there is a *possibility* that $[\iota_{w*}w: \varphi]$ denotes a world outside the Context Set. To quote Stalnaker (1975), subjunctive marking indicates that 'the selection function is one that may reach out' of the Context Set. This 'may' is now analyzed as: one of the assignments s with respect to which the sentence is evaluated is such that the value of $[\iota_{w*}w: \varphi]$ under s is not in the Context Set.

This analysis, like Stalnaker's original theory, does *not* predict that subjunctive conditionals are counterfactual. This is a welcome result in view of the following example, due to Anderson 1951 (cited in von Fintel, 1997 and also by Stalnaker):

(57)   If Jones had taken arsenic, he would have shown just exactly those symptoms which he does in fact show. [So, it is likely that he took arsenic]

The speaker wants to argue that Jones took arsenic in the actual world. If so, by Centering the *if*-clause must pick out the actual world, which is in the

Context Set. This need not cause any problem given the present theory. All the subjunctive marking indicates is that *some* of the assignments with respect to which the sentence is evaluated gives the *if*-clause a value that lies outside the Context Set. Within a dynamic framework the assertion indicates that these assignments should be thrown out of the initial information state; but this need not cause any presupposition failure.

A final point is in order. The Context Set constantly evolves in discourse, since each assertion reduces the set of assignments (and possible worlds) which are compatible with the speaker's beliefs/assertion. But since the Context Set serves to define which worlds count as 'close' to the context of utterance, the notion of closeness itself must change throughout a discourse. This is modeled in some detail in Appendix I.

5.3. 'DOUBLE SUBJUNCTIVES'

The indicative/subjunctive distinction has now been analyzed in terms of a context-dependent system of classification. As was mentioned earlier, however, there are also *object*-dependent systems of classification in natural language. We cited without much discussion the proximate/obviative distinction in Algonquian; we could also have cited the pluperfect in English, whose denotation must precede some other salient past moment rather than simply the time of utterance. As one would expect, such systems also exist in the domain of reference to worlds. The relevant facts are discussed, among others, in Dahl (1997) and Jespersen (1965). The latter observes that the pluperfect may, in colloquial speech, be used 'of the present time, simply to intensify the unreality irrespective of time'.[18] He gives the following example:

(58) If I had had money enough (at the present moment), I would have paid you.

The facts might be clearer in other languages. French displays the following minimal pairs:

59.   [John, a professional tennis player, had a terrible injury and is now at the hospital. He definitely cannot participate in the competition which is to take place tomorrow. Talking to him, I say:]
     a. Si tu avais joué demain, tu aurais gagné.
        If you had played tomorrow, you would-have won
        'If you had played tomorrow, you would have won'
     b. #Si tu jouais demain, tu gagnerais/aurais gagné
        If you played tomorrow, you would win/would-have won
        'If you played tomorrow, you would win/would have won'

In this situation, John's participation in the event is not only counterfactual (it is presupposed that it won't happen), but it is particularly remote. *Not*

using the pluperfect (interpreted as a double subjunctive) would in fact come across as insensitive, as it would disregard John's unfortunate situation.

From the present perspective an account suggests itself: in its modal uses just as in its temporal uses, the pluperfect is a device of object-relative classification. In the case at hand it indicates that the worlds picked out by the *if*-clause are more remote that some worlds which are themselves outside the Context Set. The latter are presumably worlds in which John did not have an accident. The implementation offered in the appendix simply involves a dyadic predicate ' $<$ ' (which applies across domains, i.e. to individuals, times and worlds alike). In the world domain, $\ulcorner < \urcorner$ indicates that the value of is more remote than the value of (this, in turn, can be defined in terms of Stalnaker's Selection function) (see also Stechow, 2003 for a similar analysis of temporal uses of the pluperfect; see Ippolito (2003) for a different analysis of modal pluperfects).

6. Conclusion

If the present theory is on the right track, a large part of the semantics of conditionals can be derived from independent grammatical mechanisms. *If* is simply the form taken by *the* when it applies to a description of worlds. The non-monotonicity of conditionals can thus be seen as a special case of the non-monotonicity of definite descriptions. *If*-clause can be topicalized because they are referring expression. A precise notion of co-reference between *then* and an *if*-clause can thus be developed, which accounts for Iatridou's and Izvorski's insights concerning the analogy between *then* and pronominal forms, and also predicts – correctly – binding-theoretic effects with world-denoting expressions (as anticipated in Percus, 2000). Finally, the distinction between indicative, subjunctive and double subjunctive conditionals can be analyzed in terms of a general system of referential classification which situates the value of the word(s) denoted by the *if*-clause with respect to the context or with respect to some salient world. If successful, this analysis may play a part in a general attempt to reduce the semantics of natural language to a few abstract semantic modules which apply in identical fashion to different sortal domains.

Appendix I. A Fragment

We provide a formal implementation of some of the main ideas presented in the paper. For simplicity, this system is restricted to singular descriptions, and thus adheres strictly to Stalnaker's original definition of Selection functions. The theory is stated as much as possible in a sortally-neutral fashion, i.e. whenever possible a single symbol applies to individuals, to times and to worlds (in this article no use is made of generalization to time of the relevant notions).

- *Vocabulary*

LOGICAL VOCABULARY

– Non-sortal vocabulary:
$\&, \neg, \iota, \forall, =$
– Sortal vocabulary:
for each $\xi \in \{\text{'x', 't', 'w'}\}$, for each $i \in |\text{N}$, a first-order variable ξ_i.
We say that the sortal domain of 'x' is D, the sortal domain of 't' is T and the sortal domain of 'w' is W. By extension, ξ_i is said to have the sortal domain of ξ. We write: sort('x') = D, sort('t') = T, sort('w') = W. And by extension: sort(ξ_i) = sort(ξ)

NON-LOGICAL VOCABULARY

For each m, n, p in |N, the non-logical vocabulary contains an infinite set $R^{\langle \text{:x: m, t: n, w: p} \rangle}$ of predicates taking m individual variables, n time variables and p world variables.

Proper names (individual sort): *John, Mary, Wisahkechahk , Fox* (proper names are sometimes abbreviated in what follows with their first letter only)

FEATURES

local, $\langle$local, LOCAL, $\langle$LOCAL, $<$

Note: 'local' indicates that an expression must denote a coordinate of the context (speaker, time, or world of utterance). '$\langle$local' indicates that an expression must denote an entity which is distant from the context on some measure. 'LOCAL' indicates that the value of an expression must lie in the neighborhood of the context. '$\langle$LOCAL' forces the value of an expression *not* to lie in the neighborhood of the context. Finally, ' $<$ ' is a dyadic predicate that indicates for $\lceil \alpha < \beta \rceil$ that the value of α is more remote than the value of β (from the standpoint of the context).

Parentheses and brackets: (,), [,], {, }

Note: (,) are used to indicate constituency; [,] are used to symbolize quantifiers. {, } indicate presuppositions.

- *Terms and Formulas*
- Each variable and each constant of sort s is a term of sort s.
- If $\tau_1, \ldots, \tau_m, \tau'_1, \ldots, \tau'_n, \tau''_1, \ldots, \tau''_p$ are respectively m, n and p terms of sorts D, T, W, if $R \in R^{\langle \text{x: m, t: n, w: p} \rangle}$, if φ_1 and φ_2 are formulas, and if $i \in |\text{N}$ and $\xi \in \{\text{'x', 't', 'w'}\}$, the following are formulas:

$R(\tau_1, \ldots, \tau_m, \tau'_1, \ldots, \tau'_n, \tau''_1, \ldots, \tau''_p)|\neg\varphi|(\varphi_1 \& \varphi_2)|\forall\xi_i\varphi|\tau_1 = \tau_2|\tau'_1 = \tau'_2|\tau''_1 = \tau''_2$

- If φ is a formula, if $\xi \in \{\text{'x', 't', 'w'}\}$ and if $i \in |\text{N}$, $[\iota_{\xi_k}\xi_i\varphi]$ is a term of the same sort as ξ

  – If τ, τ' are terms of sort s, the following are terms of sort s: $\tau\{local\}$ $|\tau\{<local\}$ $|\tau\{LOCAL\}$ $|\tau\{<LOCAL\}$ $|\tau\{<\tau'\}$. *Notational Convention:* For legibility, we sometimes write $R(\ldots, [\iota_{\xi_k} \xi_i \varphi], \ldots)$ as $[\iota_{\xi_k} \xi_i \varphi] R(\ldots, \xi_i, \ldots)$.
  • *Models*

A model $M = \langle D, T, W, I, f\rangle$ consists of:

(i)   three non-empty, non-intersecting sets: D, T and W (= the sortal domains or simply the 'sorts' of 'x', 't' and 'w', in the terminology introduced above)

(ii) an interpretation function I which assigns
      – a subset I(R) of $D^m \times T^n \times W^p$ to each letter R of $R^{\langle x:\ m,\ t:\ n,\ w:\ p\rangle}$
      – an element I(c) in the sortal domain of c to each constant c

(iii) a selection function from $(D \times P(D)) \cup (T \times P(T)) \cup (W \times P(W))$ into $D \cup T \cup W$ which satisfies Stalnaker's Conditions (generalized across domains):

– Condition 1 (minimum condition on Choice Functions)

$\forall S \in \{D, T, W\}\ \forall a \in S\ \forall A \subseteq S\ (f(a,A) = \# \text{ or } (f(a, A) \in A))$

– Condition 2 (condition for a Choice Function to select the 'closest' element on some measure)

$\forall S \in \{D,T,W\}\ \forall a \in S\ \forall A \subseteq S\ \forall A' \subseteq S\ (A' \subseteq A\ \&\ f(a, A) \in A' \rightarrow f(a, A) = f(a, A'))$

– Condition 3: Referential failure

$\forall S \in \{D, T, W\}\ \forall a \in S\ \forall A \subseteq S\ (f(a, A) = \# \leftrightarrow A = \emptyset)$

– Condition 4: Centering

  $\forall S \in \{D, T, W\}\ \forall a \in S\ \forall A \subseteq S\ (a \in A) \rightarrow f(a, A) = a)$

  Note 1: We can define the derived notions '$<_a$' '$\leq_a$' for each a in some sortal domain:

  $\forall S \in \{D, T, W\}\ \forall a, a', a'' \in S\ (a'' \leq_a a' \leftrightarrow_{\text{def}} f(a, \{a', a''\}) = a')$.

  $\forall S \in \{D, T, W\}\ \forall a, a', a'' \in S\ (a'' <_a a' \leftrightarrow_{\text{def}} a' \neq a''\ \&\ f(a, \{a', a''\}) = a')$.

  Note 2: $<$ is taken to represent salience for individuals, temporal remoteness in the past for times, and modal remoteness for worlds.

  Note 3: Stalnaker's conditions on f induce implausibly strong conditions on $<$, esp. for times and worlds.

  • *Information States*

– An information state is a set S of triples of the form:

  $s = \langle\langle d, t, w\rangle, g\rangle$, with $\langle d, t, w\rangle \in D \times T \times W$, and g an assignment function which assigns to each variable ξ_i for $\xi \in \{'x', 't', 'w'\}$, $i \in |N$ a value from its sortal domain.

  Note: Intuitively, an information state represents what the speaker knows about his context of speech (here identified to a triple $\langle d, t, w\rangle$), as well as about objects he might be referring to with demonstrative terms.

  Terminology: With s defined as above, we write: context $(s) = :\langle d, t, w\rangle$ and local $(D)(s) = :\ d$, local$(T)(s) = :\ t$, local$(W)(s) = :\ w$. If ξ is a variable, we write $s(\xi) = :g(\xi)$.

– We also need to determine what a speaker in a given information state considers to be the objects that count as 'close' to the context of speech.

We assume that a function LOCAL is given which associates to each element s of an information state S a triple of the form

$\langle d^+, t^+, w^+ \rangle \in P(D) \times P(T) \times P(W)$

with the stipulation that (with the notation used above) $d \in d^+$, $t \in t^+$, $w \in w^+$.

Notation: For LOCAL(s, S) defined as above, we write LOCAL(D)(s, S) =: d^+, LOCAL(T)(s, S) =: t^+, LOCAL(W)(s, S) =: w^+

Stipulation: Stalnaker's notion of Context Set

For each information state S, we stipulate that:

$\forall s \in S$ LOCAL(W)(s, S) = {local(W)(s): $s \in S$}

Let us think of S as representing the speaker's state of belief. The Stipulation has the effect of forcing LOCAL(W)(s, S) to represent Stalnaker's notion of 'Context Set' : these are the worlds which, for all the speaker knows, could be the world he lives in.

● *Reference and truth relative to an assignment and to an information state*

The definition is in two steps. First, we define reference and satisfaction under an assignment s *in an information state S*. The assignment s is not enough because S serves indirectly to determine which worlds count as 'close', which in turn determines which descriptions of worlds fail to refer (this occurs for instance when the presupposition introduced by LOCAL is violated). In a second step, we define updates on information states.

Let s be an element of an information state S.

– If ξ is a constant, $[\![\xi]\!]^{s,S} = I(\xi)$
– If ξ is a variable, $[\![\xi]\!]^{s,S} = s(\xi)$
– $[\![R(\tau_1, \ldots, \tau_m, \tau'_1, \ldots, \tau'_n, \tau''_1, \ldots, \tau''_p)]\!]^{s,S} \neq \#$ iff each of $[\![\tau_1]\!]^{s,S}, \ldots, [\![\tau''_p]\!]^{s,S}$ is $\neq \#$.

If $\neq \#$, $[\![R(\tau_1, \ldots, \tau_m, \tau'_1, \ldots, \tau'_n, \tau''_1, \ldots, \tau''_p)]\!]^{s,S} = 1$ iff $\langle [\![\tau_1]\!]^{s,S}, \ldots, [\![\tau''_p]\!]^{s,S} \rangle \in I(R)$

– $[\![\neg\varphi]\!]^{s,S} \neq \#$ iff $[\![\varphi]\!]^{s,S} \neq \#$. If $\neq \#$, $[\![\neg\varphi]\!]^{s,S} = 1$ iff $[\![\varphi]\!]^{s,S} = 0$
– $[\![\varphi_1 \& \varphi_2)]\!]^{s,S} \neq \#$ iff $[\![\phi_1]\!]^{s,S} \neq \#$ and for $S' = \{s' \in S : [\![\varphi_1]\!]^{s',S} = 1\}$, $[\![\varphi_2]\!]^{s,S'} \neq \#$.

If $\neq \#$, $[\![(\varphi_1 \& \varphi_2)]\!]^{s,S} = 1$ iff $[\![\varphi_1]\!]^{s,S} = 1$ and $[\![\varphi_2]\!]^{s,S} = 1$.

– $[\![\iota_{\xi_k}\xi_i\varphi]\!]^{s,S} \neq \#$ iff
  (a) $\forall e \in \text{sort}(\xi) : [\![\varphi]\!]^{s[\xi_i \to e],S} \neq \#$, and
  (b) $f(s(\xi_k), \{e \in \text{sort}(\xi) : [\![\varphi]\!]^{s[\xi_i \to e],S} = 1\}) \neq \#$, i.e. (given Condition 3) $\{e \in \text{sort}(\xi) : [\![\varphi]\!]^{s[\xi_i \to e],S} = 1\}) \neq \emptyset$

If $\neq \#$, $[\![\iota_{\xi_k}\xi_i\varphi]\!]^{s,S} = f(s(\xi_k), \{e \in \text{sort}(\xi) : [\![\varphi]\!]^{s[\xi_i \to e],S} = 1\})$

– If τ is a term:

$[\![\tau\{local\}]\!]^{s,S} \neq \#$ iff $\forall s' \in S [\![\tau]\!]^{s',S} \neq \#$ and $\forall s' \in S$ $[\![\tau]\!]^{s',S} = $ local (sort(τ))(s'). If $\neq \#$, $[\![\tau\{ local \}]\!]^{s,S} = [\![\tau]\!]^{s,S}$

$[\![\tau\{< local \}]\!]^{s,S} \neq \#$ iff $\forall s' \in S [\![\tau]\!]^{s',S} \neq \#$ and $\exists s' \in S$ $[\![\tau]\!]^{s',S} \neq$ local (sort(τ))(s'). If $\neq \#$, $[\![\tau\{< local\}]\!]^{s,S} = [\![\tau]\!]^{s,S}$

$[\![\tau\{LOCAL\}]\!]^{s,S} \neq \#$ iff $\forall s' \in S [\![\tau]\!]^{s',S} \neq \#$ and $\forall s' \in S$ $[\![\tau]\!]^{s',S} \in$ LOCAL(sort(τ))(s, S). If $\neq \#$, $[\![\tau\{< LOCAL\}]\!]^{s,S} = [\![\tau]\!]^{s,S}$

$[\![\tau\{< \text{LOCAL}\}]\!]^{s,S} \neq \#$ iff $\forall s' \in S[\![\tau]\!]^{s',S} \neq \#$ and $\exists s' \in S$ $[\![\tau]\!]^{s,S} \notin$
LOCAL (sort(τ))(s, S). If $\neq \#$, $[\![\tau\{\text{LOCAL}\}]\!]^{s,S} = [\![\tau]\!]^{s,S}$
$[\![\tau\{< \tau'\}]\!]^{s,S} \neq \#$ iff $\forall s' \in S[\![\tau]\!]^{s',S} \neq \#$ and $[\![\tau']\!]^{s',S} \neq \#$ and $\forall s' \in S[\![\tau]\!]^{s',S}$
$<_{\text{local}(\text{sort}(\tau))(s')} [\![\tau']\!]^{s',S}$. If $\neq \#$, $[\![\tau\{< \tau'\}]\!]^{s,S} = [\![\tau]\!]^{s,S}$
$-[\![\forall\xi_i\varphi]\!]^{s,S} \neq \#$ iff $\forall e \in \text{sort}(\xi) : [\![\varphi]\!]^{s,[\xi_i \to e],S} \neq \#$.
If $\neq \#$, $[\![\forall\xi_i\varphi]\!]^{s,S} = 1$ iff for all $e \in \text{sort}(\xi) : [\![\varphi]\!]^{s,[\xi_i \to e],S} = 1$.

- $[\![\tau = \tau']\!]^{s',S} \neq \#$ iff $[\![\tau]\!]^{s,S} \neq \#$ and $[\![\tau']\!]^{s,S} \neq \#$. If $\neq \#$, $[\![\tau = \tau']\!]^{s,S} = 1$ iff $[\![\tau]\!]^{s',S} = [\![\tau']\!]^{s',S}$
  - *Updates*

This is a standard update semantics, in which $S[\varphi]$ is the result of updating information state S with the formula φ

Let S be an information state. Then:

- $S[R(\tau_1, \ldots, \tau_m, \tau'_1 \ldots, \tau'_n, \tau''_1, \ldots, \tau''_p)] \neq \#$ iff $\forall s \in S[\![R(\tau_1, \ldots, \tau_m,$
  $\tau'_1 \ldots, \tau'_n, \tau''_1, \ldots, \tau''_p)]\!]^{s,S} \neq \#$. If $\neq \#$,
  $S[R(\tau_1, \ldots, \tau_m, \tau'_1 \ldots, \tau'_n, \tau''_1, \ldots, \tau''_p)] = \{s \in S : [\![R(\tau_1, \ldots, \tau_m, \tau'_1 \ldots, \tau'_n,$
  $\tau''_1, \ldots, \tau''_p)]\!]^{s,S} = 1\}$
- $S[\neg\varphi] \neq \#$ iff $S[\varphi] \neq \#$. If $\neq \#$, $S[\neg\varphi] = S - S[\varphi]$
- $S[(\varphi_1 \& \varphi_2)] \neq \#$ iff $S[\varphi_1] \neq \#$ and $S[\varphi_1][\varphi_2] \neq \#$. If $\neq \#$, $S[(\varphi_1 \& \varphi_2)] = S[\varphi_1][\varphi_2]$
- $S[\forall\xi_i\varphi] \neq \#$ iff $\forall s \in S \; \forall e \in \text{sort}(\xi) [\![\varphi]\!]^{s[\xi_i \to e],S} \neq \#$. If $\neq \#$,
  $S[\forall\xi_i\varphi] = \{s \in S : \forall e \in \text{sort}(\xi)[\![\varphi]\!]^{s[\xi_i \to e],S} = 1\}$
- $S[\tau = \tau'] \neq \#$ iff $\forall s \in S[\![\tau = \tau']\!]^{s,S}$. If $\neq \#$,
  $S[\tau = \tau'] = \{s \in S : [\![\tau]\!]^{s,S} = [\![\tau]\!]^{s,S}\}$
  - *Truth*

φ is true with respect to s iff $\{s\}[\varphi] = \{s\}$ (in other words, the singleton $\{s\}$ updated with φ is $\{s\}$ itself).

  - *Examples*

(I depart from the 'official' notation for predicates, for which I use English words.)

(60)   a. He$_i$ is me$_k$ (*is not a presupposition failure*)
       b. $x_i\{< \text{local}\} = x_k\{\text{local}\}$
       c. $S[b] \neq \#$ iff $\forall s \in S[\![x_i\{< \text{local}\} = x_k\{\text{local}]\!]^{s,S} \neq \#$
          iff $\forall s \in S(\exists s' \in S[\![x_i]\!]^{s',S} \neq \text{local}(D)(s') \& \forall s' \in S[\![x_k]\!]^{s',S} = \text{local}(D)(s'))$
          iff $\exists s' \in S \; s'(x_i) \neq \text{local}(D)(s') \& \forall s' \in S s'(x_k) = \text{local}(D)(s'))$
          If $\neq \#$, $S[b] = \{s \in S : [\![x_i\{< \text{local}\}]\!]^{s,S} = [\![x_k\{\text{local}\}]\!]^{s,S}\} = \{s \in S : s(x_i) = s(x_k)\}$

Note: Intuitively, the definedness condition indicates that (i) there must be a possibility that 'he$_i$' doesn't refer to the speaker, i.e. it must *not* be presupposed that 'he$_i$' refers to the speaker, and (ii) it must be presupposed that 'me$_k$' refers to the speaker.

(61)   a. That$_i$ is this$_k$ (*is not a presupposition failure*)

      b. $x_i\{<\text{LOCAL}\} = x_k\{\text{LOCAL}\}$

      c. $S[b] \neq \#$ iff $\forall s \in S[\![x_i\{< \text{LOCAL}\} = x_k\{\text{LOCAL}\}]\!]^{s,S} \neq \#$
        iff $\forall s \in S(\exists s' \in S[\![x_i]\!]^{s',S} \notin \text{LOCAL (D) } (s',S) \, \& \, \forall s' \in S[\![x_k]\!]^{s',S} \in$
        LOCAL (D) (s', S)
        iff $\exists s' \in S \, s'(x_i) \notin$
        LOCAL(D) (s',S) $\& \, \forall s' \in S \, s'(x_k) \in$ LOCAL (D)
        (s', S)) If $\neq \#$,
        $S[b] = \{s \in S : [\![x_i\{< \text{LOCAL}\}]\!]^{s,S} = [\![x_k\{\text{LOCAL}\}]\!]^{s,S}\} =$
        $\{s \in S : s(i) = s(k)\}$

(62)   a. Wisahkechahk$^{\textbf{PROX}}$ leave-behind Fox$^{\textbf{OBV}}$ *(proximate vs. obviative*
        in Algonquian)

      b. leave(W$\{<$local$\}$, F$\{< W\{<$local$\}\}$,t$_0$,w$_0$))

      c. $S[b] \neq \#$ iff $\forall s \in S[\![\text{leave}(W\{< \text{local}\}, F\{< W\{<\text{local}\}\}, t_0, w_0)]\!]^{s,S}$
        $\neq \#$iff $\exists s \in S \, I(W) \neq$ local (D) (s)$\&\forall s \in S \, I(F) <_{\text{local}(D)(S)} I(W)$
        If $\neq \#$,
        $S[b] = \{s \in S : \langle I(W), I(F), s(t_0), s(w_0)\rangle \in I(\text{leave})\}$

Note: Proximate marking is analyzed as simple third person features in English. Obviative marking on 'Fox' is analyzed as a presupposition that Fox is less salient than another individual which is denoted using a proximate expression (in this case, 'Wisahkechahk').

(63)   a. The man is coughing *(need not refer to the speaker even if the*
        speaker is male)

      b. cough($[\iota_{x_k} x_i \text{ man}(x_i, t_0, w_0)]$, t$_0$, w$_0$)

      c. $S[b] \neq \#$ iff $\forall s \in S\{e \in D : [\![\text{man}(x_i, t_0, w_0)]\!]^{s[x_i \to e], S} = 1\} \neq \emptyset$,
        iff $\forall s \in S\{e \in D :< e, s(t_0), s(w_0) >\in I(\text{man})\} \neq \emptyset$
        If $\neq \#$
        $S[b] = \{s \in S : < f(s(x_k), \{e \in D :< e, s(t_0), s(w_0) >\in I(\text{man})\})$,
        $s(t_0), s(w_0) >\in I(\text{coughing})\}$

Note: If the point of reference $s(x_k)$ is a man, then by Centering 'the man' must refer to $s(x_k)$. In particular, if $s(x_k)$ is the speaker and the speaker is a man, then 'the man' must refer to the speaker. However nothing forces x_k to denote the speaker, and thus in general 'the man' could refer to someone other than the speaker even if the latter is a man.

(64)   a. The pig is grunting, but the pig with floppy ears isn't grunting
        (need not be a contradiction)

      b. $[\iota_{x_k} x_i \text{ pig}(x_i, t_0, w_0)]$ grunting(x_i, t$_0$, w$_0$) $\& [\iota_{x_k} x_i((\text{pig}(x_i, t_0, w_0) \&$
        floppy$(x_i, t_0, w_0))]\neg$ grunting(x_i, t$_0$, w$_0$)

      c. $S[b] \neq \#$ iff $\forall s \in S\{e \in D : \langle e, s(t_0), s(w_0)\rangle \in I(\text{pig})\} \neq \emptyset$ and
        $\forall s \in S$ s.t.$\langle f(s(x_k), \{e \in D : \langle e, s(t_0), s(w_0)\rangle \in I(\text{pig})\}), s(t_0),$
        $s(w_0)\rangle \in I(\text{grunting})\}, \{e \in D : \langle e, s(t_0), s(w_0)\rangle \in I(\text{pig})\&$
        $\langle e, s(t_0), s(w_0)\rangle \in I(\text{floppy})\} \neq \emptyset$
        If $\neq \#$,
        $S[b] = \{s \in S : \langle f(s(x_k), \{e \in D : \langle e, s(t_0), s(w_0)\rangle \in I(\text{pig})\})$,

s $(t_0), s(w_0)\rangle \in$ I(grunting) & {e $\in$ D : $\langle$e, s(t_0), s(w_0)$\rangle \in$ I(pig)& $\langle$e, s $(t_0), s(w_0)\rangle \in$ I(floppy), s(t_0), s(w_0)$\rangle \notin$ I (grunting)}

Note: If the closest pig doesn't have floppy ears, the selection function f will not select the same individual for the two descriptions 'the pig' and 'the pig with floppy ears', which explains that the sentence isn't contradictory.

(65)   a. If John came, Mary would be happy, but if John came and he was drunk, Mary wouldn't be happy *(need not be a contradiction)*

  b. $[\iota_{w_0} w_i$ came(J,t_0,w_i)]happy(M,t_0,w_i) & $[\iota_{w_0} w_i$ came(J,t_0,w_i) & drunk (J,t_0,w_i)] $\neg$ happy(M, t_0, w_i)

  c. S[b] $\neq$ # iff $\forall s \in S$ {e $\in$ W: $\langle$I(J), s(t_0), e $\rangle \in$ I(came)} $\neq \emptyset$ and $\forall s \in S$ s.t. $\langle$I(M), s(t_0), f(s(w_0), {e $\in$ W: $\langle$I(J), s(t_0), e $\rangle \in$ I(came)}) $\rangle \in$ I(happy)}: {e $\in$ W: $\langle$I(J), s(t_0), e $\rangle \in$ I(came) & $\langle$I(J), s(t_0), e $\rangle \in$ I(drunk)} $\neq \emptyset$

  If $\neq$ #,

  S[b] = {s $\in$ S: $\langle$I(M), s(t_0), f(s(w_0), {e $\in$ W: $\langle$I(J), s(t_0), e $\rangle \in$ I(came)}) $\rangle \in$ I(happy) & $\langle$I(M), s(t_0), f(s(w_0), {e $\in$ W: {e $\in$ W: $\langle$I(J), s(t_0), e $\rangle \in$ I(came) & $\langle$I(J), s(t_0), e$\rangle \in$ I(drunk)}) $\rangle \notin$ I(happy)}

[For simplicity I have disregarded mood in this example. See below]

Note: If the closest world in which John comes is one in which he isn't drunk, the selection function f will not select the same world for the two descriptions 'if John comes' and 'if John comes and is drunk', which explains that the sentence isn't contradictory.

(66)   a. If John is sick, Mary is unhappy *(indicative conditional)*

  b. unhappy(Mary, t_0\{local\}, $[\iota_{w_0} w_i$ sick(J,t_0\{local\},w_i)]\{LOCAL\})

  c. S[b] $\neq$ # iff $\forall s \in S$ ({e $\in$ W: $\langle$I(J), s(t_0), e $\rangle \in$ I(sick)} $\neq \emptyset$ & s(t_0) = local(T)(s) & f(s(w_0), {e $\in$ W: $\langle$I(J), s(t_0), e $\rangle\in$ I(sick)}) $\in$ LOCAL(W)(s, S), with LOCAL(W)(s, S) = {local(W)(s): s $\in$ S} (= the Context Set), by the Stipulation introduced above.)

  If $\neq$ #,

  S[b] = {s $\in$ S: $\langle$I(M), s(t_0), f(s(w_0), {e $\in$ W:$\langle$I(J),s(t_0), e $\rangle \in$ I(sick)}) $\rangle \in$ I(unhappy)}

[Present tense and mood are treated in terms of the features 'local' and 'LOCAL']

(67)   a. If John were sick, Mary would be unhappy *(subjunctive conditional)*

  b. unhappy(Mary, t_0\{local\}, $[\iota_{w_0} w_i$ sick(J,t_0\{local\},w_i)]\{\langle LOCAL\})

  c. S[b] $\neq$ # iff $\forall s \in S$ ({e $\in$ W: $\langle$I(J), s(t_0), e $\rangle \in$ I(sick)} $\neq \emptyset$ & s(t_0) = local(T)(s))& $\exists s \in S$ (f(s(w_0), {e $\in$ W: $\langle$I(J), s(t_0), e $\rangle \in$ I(sick)}) $\notin$ LOCAL(W)(s, S)), with LOCAL(W)(s, S) = {local(W)(s): s $\in$ S} (= the Context Set), by the Stipulation introduced above.

  If $\neq$ #,

$S[b] = \{s \in S: \langle I(M), s(t_0), f(s(w_0), \{e \in W: \langle I(J),s(t_0), e \rangle \in I(sick)\}) \rangle \in I(unhappy)\}$

(68) a. If John had been sick (now), Mary would have been unhappy
       (*double subjunctive conditional*)

      b. unhappy(Mary, $t_0\{local\}$, $[\iota_{w_0}w_i$ sick$(J,t_0\{local\},w_i)]\{ < w_1\{ <$LO-CAL$\}\}))$

      c. $S[b] \neq$ # iff
         (i) $\exists s \in S$ $s(w_1) \notin$ LOCAL(W)(s, S), with LOCAL(W)(s, S) = $\{local(W)(s): s \in S\}$ (= the Context Set), by the Stipulation introduced above.
         [this is the presupposition introduced by $w_1\{ <$LOCAL$\})$]
         (ii) $\forall s \in S$ $s(t_0) = local(T)(s)$ [presupposition introduced y t_0 $\{local\}$]
         (iii) $\forall s \in S$ $\{e \in W: \langle I(J), s(t_0), e \rangle \in I(sick)\} \neq \emptyset$ [presupposition introduced by the *if*-clause]
         (iv) $\forall s \in S$ $f(s(w_0), \{e \in W:\langle I(J), s(t_0), e \rangle \in I(sick)\}) \langle_{local(W)(s)}$ $s(w_1)$ If $\neq$ #,
         $S[b] = \{s \in S: \langle I(M), s(t_0), f(s(w_0), \{e \in W:\langle I(J), s(t_0), e \rangle \in I(sick)\}) \rangle \in I(unhappy)\}$

Appendix II. Further Remarks on Plural Choice Functions

Recall the first three conditions that were imposed on our Plural Choice/ Selection Functions:

Condition 1*: For each element d and each non-empty set E of elements, $f(d, E) \neq$ # and $f(d, E) \subseteq E$.

Condition 2*: For each element d, each set E and each set E′, if E′ $\subseteq$ E and $f(d, E) \cap E' \neq \emptyset$, then $f(d, E') = f(d, E) \cap E'$.

Condition 3*: For each element d and each set E, $f(d, E) =$ # if E $= \emptyset$.

In what follows, we assume that # is the empty set $\emptyset$ (by contrast, in the rest of this article we took # to indicate referential failure, which forced us to develop the analysis in a three-valued logic; we are now going back to a bivalent system). We further assume that each subset of the domain has a name. The semantic value of a formula A is written as $[\![A]\!]$ or simply as **A**.

A. Axioms Corresponding to Condition 2* (Valentina Gliozzi)

Valentina Gliozzi (p.c.) makes the following observations. Condition 2* can be seen as the conjunction of Condition 2*a and Condition 2*b:

Condition 2*a: For each element d, each set E and each set E′, if E′ $\subseteq$ E and $f(d, E) \cap E' \neq \emptyset$, then $f(d, E') \subseteq f(d, E) \cap E'$.

Condition 2*b: For each element d, each set E and each set E′, if E′ ⊆ E and f(d, E) ∩ E′ ≠ ∅, then f(d, E′) ⊇ f(d, E) ∩ E′.

Gliozzi then remarks that, in the presence of Condition 1* and Condition 3*, Condition 2*a is satisfied if and only if the axiom (CV) holds, and similarly that Condition 2*b holds if and only if (DT) holds:

(CV) (A > B &¬(A > ¬C))⇒((A&C) > B)
(DT) ((A&C) > B)⇒A>(C⇒B)

Claim 1: Condition 2*a is satisfied iff (CV) holds

Proof: (i) Condition 2*a=>(CV)

Assume that Condition 2*a holds. Assume further that (A > B &¬(A > ¬C)) is true. This means that f(d, **A**) ⊆ **B** and not f(d, **A**) ⊆ ⟦¬**C**⟧, i.e. f(d, **A**) ∩ **C** ≠∅. Since **A** ∩ **C** ⊆ **A** and f(d, **A**) ∩ **C** ≠ ∅, by Condition 2*a we have: f(d, **A** ∩ **C**) ⊆ f(d, **A**) ∩ (**A** ∩ **C**) ⊆ **B**. This gives immediately: A&C > B

(ii) (CV)=>Condition 2*a

Assume that (CV) holds. Assume further that E′ ⊆ E and f(d, E) ∩ E′ ≠ ∅. Since we have assumed that each subset of the domain has a name, we may posit that E = ⟦A⟧, E′ = ⟦A&C⟧, and ⟦B⟧ = f(d, **A**). By Condition 1* and Condition 3*, f(d, **A**) ⊆ **A** ⊆ **B**, hence A > B holds; and f(d, **A**) ∩⟦A&C⟧ ≠ ∅, hence f(d, **A**) ∩ **C** ≠ ∅, which entails that ¬ (A > ¬C) holds. Thus the antecedent of (CV) is satisfied, and by (CV) f(d, E′) ⊆ **B** = f(d, **A**) = f(d, E). By Condition 1* and Condition 3*, f(d, E′) ⊆ E′. Thus in the end f(d, E′) ⊆ f(d, E) ∩ E′.

Claim 2: Condition 2*b is satisfied iff (DT) holds

Proof: (i) Condition 2*b=>(DT)

Assume that Condition 2*b holds. Assume further that ((A&C) > B) is true, i.e. that f(d, ⟦A&C⟧) ⊆ ⟦B⟧. Then:

–if f(d, **A**) ∩ ⟦A&C⟧ = ∅, there are two cases:

Case 1. f(d, **A**) = ∅. By Condition 3* and our assimilation of # to ∅, this entails that **A** = ∅, which in turn entails (again by Condition 3*) that A > (C⇒B) holds.

Case 2. f(d, **A**) ≠ ∅. Since by Condition 1* f(d, **A**) ⊆ **A**, it must be the case that f(d, **A**) ⊆ ⟦¬C⟧, and hence f(d, **A**) ⊆ ⟦C⇒B⟧

–if f(d, **A**) ∩ ⟦A&C⟧ ≠ ∅, by Condition 2*b f(d, **A**) ∩ ⟦A&C⟧ ⊆ f(d, ⟦A&C⟧), hence f(d, **A**) ∩ ⟦C⟧ ⊆ f(d, ⟦A&C⟧), hence f(d, **A**) ∩ ⟦C⟧ ⊆ ⟦B⟧, and thus A > (C⇒B) is true.

(ii) (DT) => Condition 2*b

Assume that (DT) holds and that E′ ⊆ E and f(d, E) ∩ E′ ≠ ∅. Since we have assumed that each subset of the domain can be named, we can posit that E′ = ⟦A&C⟧ ⊆ E = ⟦A⟧. Let ⟦B⟧ = f(d, E′) = f(d, ⟦A&C⟧). The antecedent of (DT) is (trivially) satisfied, hence so is the consequent. This means that f(d, **A**) ∩ ⟦C⟧ ⊆ ⟦B⟧, hence also f(d, **A**) ∩ ⟦A&C⟧ ⊆ ⟦B⟧, and thus that f(d, E) ∩ E′ ⊆ f(d, E′).

B. Restating Condition 1* and Condition 2* in terms of a transitive well-founded relation (Ede Zimmermann)

With the same assumptions as in A., Ede Zimmermann (p.c.) observes that for any d in the domain, the conjunction of Condition 1*, Condition 2* and Condition 3* holds if and only if (TWF) below holds:

(TWF) There is a transitive well-founded relation $\leq_d$ such that $f(d, E) = \{e \subseteq E \mid$ for all $e' \in E: e \leq_d e'\}$.

Proof:

" $<=$ "

Condition 1*: follows from the well-foundedness of $\leq_d$

Condition 2*: Suppose $E' \subseteq E$ and $f(d, E) \cap E' \neq \emptyset$, i.e. $\{e \subseteq E \mid$ for all $e' \in E: e \leq_d e'\} \cap E' \neq \emptyset$. By (TWF), $f(d, E') = \{e \subseteq E' \mid$ for all $e' \in E': e \leq_d e'\}$

Clearly, $f(d, E) \cap E' = \{e \subseteq E' \mid$ for all $e' \in E: e \leq_d e'\} \subseteq \{e \subseteq E' \mid$ for all $e' \in E': e \leq_d e'\} = f(d, E')$ (because $E' \subseteq E$)

Conversely, assume $e \in \{e \subseteq E' \mid$ for all $e' \in E': e \leq_d e'\}$. Pick any $e^* \in f(d, E) \cap E'$ (which is $\neq \emptyset$, by assumption). Then $e \leq_d e^*$ because $e^* \in E'$. For any $e' \in E$, $e^* \leq e'$ because $e^* \in f(d, E)$, and by the transitivity of $\leq_d$, $e \leq_d e'$. Hence $e \in f(d, E) \cap E'$

Condition 3*: $f(d, \emptyset) = \{e \subseteq \emptyset \mid$ for all $e' \in E: e \leq_d e'\} = \emptyset$.

" $=>$ "

Assume Condition 1*, Condition 2* and Condition 3*, and let

$\leq_d = \{\langle e, e' \rangle \in D \times D \mid e \in f(d, \{e, e'\})\}$. Then:

(i) $\leq_d$ is transitive

Assume $e \leq e'$ and $e' \leq e''$. It is enough to show that $e \in f(d, \{e, e', e''\})$: since $e \in f(d, \{e, e', e''\}) \cap \{e, e''\}$, Condition 2* implies $e \in f(d, \{e, e''\})$, i.e. $e \leq e''$.

To show that $e \in f(d, \{e, e', e''\})$, we consider two cases:

Case 1. $e' \in f(d, \{e, e', e''\})$. Hence $f(d, \{e, e', e''\}) \cap \{e, e'\} \neq \emptyset$, and by Condition 2* (since $e \leq e'$) $e \in f(d, \{e, e'\}) = f(d, \{e, e', e''\}) \cap \{e, e'\}$. Hence $e \in f(d, \{e, e', e''\})$.

Case 2. $e' \notin f(d, \{e, e', e''\})$. If $e \notin f(d, \{e, e', e''\})$, $f(d, \{e, e', e''\}) = \{e''\}$ (by Condition 1*). But by Condition 2* $f(d, \{e', e''\}) = f(d, \{e, e', e''\}) \cap \{e', e''\} = \{e''\}$, contradicting $e' \leq_d e''$. Contradiction.

(ii) $\leq_d$ is well-founded

Let E be a non-empty subset of the domain. We must show that there exists $e^* \in E$ such that for all $e \in E$, $e^* \leq_d e$. By Condition 1*, $f(d, E) \neq \emptyset$. Pick any $e^* \in f(d, E)$, and consider an arbitrary $e \in E$. By Condition 2*, $f(d, \{e^*, e\}) = f(d, E) \cap \{e^*, e\}$ since $e^* \in f(d, E) \cap \{e^*, e\}$. Thus $e^* \in f(d, \{e^*, e\})$ (again because $e^* \in f(d, E) \cap \{e^*, e\}$), i.e. $e^* \leq e$.

(iii) $f(d, E) = \{e \in E \mid$ for all $e' \in E: e \leq_d e'\}$

– If $e \in f(d, E)$, then $e \in f$, by Condition 1*. Furthermore, if $e' \in E$, $\{e, e'\} \subseteq E$ and $f(d, E) \cap \{e, e'\} \neq \emptyset$, hence by Condition 2* $f(d, \{e, e'\}) = f(d, E) \cap \{e, e'\}$. Since $e \in f(d, E) \cap \{e, e'\}$, $e \in f(d, \{e, e'\})$. Thus $f(d, E) \subseteq \{e \in E \mid$ for all $e' \in E: e \leq {}_de'\}$

– Suppose $e \in E$ and for all $e' \in E: e \leq {}_de'$. $E \neq \emptyset$, hence by Condition 1* there is some $e^* \in f(d, E) \cap E$. Since for all $e' \in E: e \leq {}_de'$, $e \leq e^*$, i.e. $e \in f(d, \{e, e^*\}) = f(d, E) \cap \{e, e^*\}$ by Condition 2*. So $e \in f(d, E)$.

Notes

[2] This section owes much to the comments of an anonymous reviewer, who provided extremely useful criticisms of an earlier version.
[3] His (short) discussion centered around the hypothesis that conditionals are strict implications with a *vague* accessibility relation and a context-sensitive mechanism of resolution. But his argument carries over to the hypothesis that the quantifier restriction rather than the accessibility relation is context sensitive.
[4] One could try other procedures to ensure that a domain restriction remains constant, for instance by resorting to the elision of a restrictor, as in the following French examples:

(i) a. Chaque Italien en a critiqué un autre.
    Each Italien EN has criticizied another
a′. Chaque Italien a critiqué un autre Italien.
    Each Italien has criticizied another Italian
b. Un Italien en a critiqué un autre.
    An Italien EN has criticizied another Italian
b′. Un Italien a critiqué un autre Italien.
    An Italien has criticizied another Italian
c. Chaque professeur en a critiqué un autre.
   Each professor EN has criticized another
c′. Chaque professeur a critiqué un autre professeur.
    Each professor has criticized another professor
d. Un professeur en a critiqué un autre.
   A professor EN has criticized another
d′. Un professeur a critiqué un autre professeur
    A professor has criticized another professor
e. Le professeur en a critiqué un autre.
   The professor EN has criticized another

e′. Le professeur a critiqué un autre professeur
   The professor has criticized another professor
   (cf. McCawley 1979: 'The dog got into a fight with another dog')

The reading obtained in (17a) appears to be harder to get in (ia) than in (ia′), in the sense that in (ia) one understands that each Italian among a group P criticized some Italian *in the same group P*. If this judgment is robust (which is not clear at all), one could hypothesize that ellipsis forces the elided restrictor to have the same semantic value (and in particular the same domain restriction) as its antecedent. This paradigm could then provide an *experimentum crucis* to decide between the monotonic and the non-monotonic approaches. On the monotonic line, (i.e.) should come out as a presupposition failure, because *the professor* presupposes that there is a single professor in the domain of discourse, while *un autre* presupposes or asserts that there are at least two. On the non-monotonic analysis, by contrast, there should be no such effect. I must grant that I am uncertain about the facts, which are rather subtle.

[5] In the case of conditionals, Fintel 1999 further notes that the following rejoinders are quite natural as uttered by the addressee of (22a)

   (i) a. But that means that if the USA threw its weapons into the sea tomorrow, there wouldn't NECESSARILY be war. (Fintel, 1999)
       b. But that means that if the USA threw its weapons into the sea tomorrow, there might NOT be war (Fintel, 1999).

Fintel (1999) writes that the possibility of such replies 'is unexpected under the standard static approach. If we go back to the simpler antecedent, the domain of quantification should shrink back to the closest worlds where just the USA disarms, ignoring the far-fetched worlds where all nuclear powers become meek. But that doesn't seem to happen.'

I am not sure this argument is convincing. Note that unlike the inference from *if p, q* to *if p and p′, q*, the inference from *Necessarily, if p, q* to *Necessarily, if p and p′, q* does not suffer obvious exceptions in natural language. Thus proponents of the non-monotonic analysis are still forced to posit a monotonic analysis for sentences involving adverbs of quantification. If so the naturalness of the rejoinders in (i) comes as no surprise at all: the initial assertion in (22) a made the point that the closest world in which the US throws its weapons into the sea is one in which there is war, but that the closest world in which the US and all other nuclear powers throw their weapons into the sea is one in which there is peace. The addressee correctly infers from this that not *all* worlds (in the domain of quantification) in which the US throws its weapons into the sea are worlds in which there is war – which is unobjectionable.

[6] Note also that if there is a significant difference between definite descriptions and conditionals, it is not clear how the monotonic analysis of Fintel, (1999, 2001) can account for it. Fintel's main analytical tool is that of *Strawson-entailment,* defined as follows:

   (i) Strawson Downward Entailingness (Fintel, 1999)
   A function f of type $\langle s, t \rangle$ is Strawson-DE
   iff for all x, y of type s such that $x \Rightarrow y$ and f(x) is defined : $f(y) \Rightarrow f(x)$ [where $\Rightarrow$ applies in the usual way to all types that 'end in t'].

Suppose we analyze definite descriptions in terms of maximality operators. On Fintel's theory *The Ns had a good time* is a Strawson DE environment, since for each N_1 and N_2 such that (a) $N_1 \Rightarrow N_2$ and (b) the presupposition of *The N_1s had a good time* is met, i.e. N1 has at least two individuals in its extension, it is the case that *The N_2s had a good time* $\Rightarrow$ *The N_1s had a good time* (example: *The students who knew many people had a good time* $\Rightarrow$ *The students who knew some people had a good time*). Thus the expectation is that plural definite descriptions should in fact license Negative Polarity Items (presumably Fintel could adopt a non-monotonic analysis of definite descriptions, while still preserving a monotonic analysis of conditionals).

[7] To make this analysis complete we would have to explain how the formula $X \subseteq X'$ can appear in the restrictor of the second quantifier. This is presumably an effect of domain restriction, but I leave this issue for future research.

[8] As Ede Zimmermann points out, however, this leaves one question open – do we ever obtain cases of collective prediction with *if*-clauses, analyzed as plural descriptions? I do not know of any such cases; if none can be found this will count as an argument against the present theory, and in favor the the analysis based on universal quantification.

[9] The second argument could just as well be taken to be a sentence rather than a set of worlds (= a proposition).

[10] Note that in each of these cases at least two dogs are relevant to each of my friends (his own and the one it got into a fight with), which shows that domain restriction alone cannot account for the data (there must be some way to select the 'salient' dog within each of these domains).

[11] In a De Se analysis of attitude reports (as in Chierchia 1989), (ia) would be analyzed along the lines of (ib), where the point of reference x of the ι-operator is bound by a λ-operator introduced by the *that*-clause:

  (i) a. John said that the dog had gotten into a fight with another dog.

    b. John said that $\lambda x\ \lambda w\ [\iota_x\ y: dog(y, w)]$ got into a fight with another dog.

[12] Stalnaker further stipulates that λ 'is an isolated element under [the accessibility relation] R; that is, no other world is possible with respect to it, and it is not possible with respect to any other world'.

[13] What about other cases? 'If round squares existed, you might get the job' – no infelicity here, except for the applicant. I would suggest that the speaker who utters this presents himself as assuming that there *is* a possible world in which round squares exist, although this is a very remote one.

[14] See for instance Roberts (1989) for an analysis of modal subordination.

[15] As an anonymous reviewer points out, Mathewson (2001) suggests that quite generally quantifiers contain a hidden referential element which selects some of the objects that satisfy the restrictor.

[16] Stalnaker (p.c.) has pointed out that a treatment of *if*-clauses as plural definite descriptions might allow us to have our cake and eat it too when it comes to the Conditional Excluded Middle, a suggestion already made in Fintel (1997). Within a semantics in which an *if*-clause denotes a plurality of worlds, *(if p, q or if p, not q)* should *not* come out as valid, since it might well happen that some of the p-worlds selected by the *if*-clause are q-worlds, while others are not. However Fintel (1997) and Löbner (1985, 1987) have argued that definite plural noun phrases have a 'homogeneity presupposition' (Fintel 1997) or satisfy the 'logical property of completeness': *If the predicate P is false for the NP, its negation not-P is true for the NP.* Fintel and Löbner observe that in a situation where all of ten children are playing, among them three boys and seven girls, (ia) and (ib) have a clear truth value, but (ic) does not, as is expected if a presupposition of homogeneity holds of definite plurals:

  (i) a. TRUE: The children are playing.

    b. FALSE: The children are not playing.

    c. ?: The children are boys." (Fintel, 1997, citing Löbner, 1987).

If the same presupposition holds of *if*-clauses, as is suggested by Fintel (1997), the Conditional Excluded Middle will appear to hold of all sentences that can be uttered felicitously.

[17] Here is why. In Lewis's sphere-based system, the truth-conditions a conditional If ϕ, ψ are as follows:

If ϕ, ψ is true at world i (according to the system of spheres $) if and only if either

(1) No ϕ-world belongs to any sphere S in $_ip$, or

(2) Some sphere S in $_1 does contain at least one ϕ-world, and $\phi \Rightarrow \psi$ holds at every world in S

Take $\phi = \ulcorner$ this line is longer than 1″$\urcorner$, $\psi_\varepsilon = \ulcorner$ this line is longer than $(1 + \varepsilon)''\urcorner$ (note that ε is a parameter, to be replaced by its value, not quoted; hence the use of Quine's quasi-quotation marks).

Let $_\varepsilon$ be the sphere that contains all the worlds in which the size of the line is smaller than $(1 + \varepsilon)''$ (by assumption such a sphere exists since the similarity measure in the context of speech orders worlds by the value of ε at those worlds). Clearly there are ϕ-world in $_\varepsilon$ if $\varepsilon > 0$. In these worlds $\phi \Rightarrow \psi$ is true. Hence the conditional should be true.

[18] Thanks to Frank Veltman for bringing these examples to my attention.

References

Anderson A. R. (1951). A Note on Subjunctive and Counterfactual Conditionals. *Analysis*, 11, pp. 35–38.

Barker S. (1997) E-type pronouns, DRT, dynamic semantics and the quantifier/variable binding model. *Linguistics & Philosophy*, 20(2), pp. 195–228.

Bhatt R., Pancheva R. (2001) *The Syntax of Conditionals*, ms., University of Texas at Austin & University of Southern California.

Bittner M. (2001) Topical Referents for Individuals and Possibilities. In Hastings R. et al. (eds.), *Proceedings from SALT* XI, CLC, Cornell University, Ithaca, pp. 36–55.

Chierchia G. (1989) Anaphora and Attitude De Se, In Bartsch R., Van Benthem J., van Emde Boas P. (eds.) *Semantics and Contextual Expressions*, Foris, Dordorecht, pp.1–31.

Dekker P. (2000) Coreference and Representationalism. In von Heusinger K., Egli U. *Reference and Anaphoric Relations*, Kluwer.

Fintel K. von (1997) Bare Plurals, Bare Conditionals, and Only. *Journal of Semantics*, 14 (1), pp. 1–56.

Fintel K. von (1999) NPI-Licensing, Strawson-Entailment, and Context-Dependency *Journal of Semantics*, 16, pp. 97–148.

Fintel K. von (2001) Counterfactuals in a Dynamic Context. In Kenstowicz, M., (ed.), *Ken Hale: A Life in Language*, MIT Press Cambridge, MA, pp. 123–152.

Frank A. (1997) *Context Dependence in Modal Constructions*. Doctoral Dissertation, University of Stuttgart.

Geurts B. (1999) *Presuppositions and Pronouns*, Elsevier.

Hale K. (1976) The Adjoined Relative Clause in Australia. In Dixon R. (ed.) *Grammatical Categories in Australian Languages*, Australian Institute of Aboriginal Studies, Canberra, pp. 78–105.

Harper W. L., Stalnaker R., Pearce G. (eds.) (1981) *Ifs*, Reidel.

Heim I. (1991) Artikel und Definitheit. In Stechow A. v., Wunderlich D. (eds.) *Handbuch der Semantik*. de Gruyter, Berlin.

Heim I. (1984) A Note on Polarity Sensitivity and Downward Entailingness. *Proceedings of NELS* 14, pp. 98–107.

Heim I. (1992) Presupposition Projection and the Semantics of Attitude Verbs. *Journal of Semantics* 9, pp. 183–221.

Heusinger K. von (1996) *Salienz und Referenz (Der Epsilonoperator in der Semantik der Nominalphrase und anaphorischer Pronomen)*.

Heusinger K. von (2000) Reference and Salience, available online at: http://ling.uni-konstanz.de/pages/home/heusinger/publications/ftp/ap71_2.pdf.

Iatridou S. (1994) On the Contribution of Conditional *Then*. *Natural Language Semantics*, 2, pp. 171–199.

Iatridou S. (2000) The Grammatical Ingredients of Counterfactuality. *Linguistic Inquiry*, 31, pp. 231–270.

Ippolito M. (2003) Presuppositions and Implicatures in Counterfactuals. *Natural Language Semantics*, 11, pp. 145–186.

Izvorski R. (1996) The Syntax and Semantics of Correlative Proforms. In Kusumoto (ed.), *Proceedings of NELS 26*.

Kratzer A. (1991) Conditionals. In von Stechow A., Wunderlich D. (eds.) *Semantik: ein internationales Handbuch der zeitgenössischen Forschung*.

Lewis D. (1973) *Counterfactuals*, Harvard

Lewis D. (1979) Scorekeeping in a Language Game. *Journal of Philosophical Logic*, 8, pp. 339–359.

Löbner S. (1985) Definites. *Journal of Semantics*, 4, pp. 279–326.

Löbner S. (1987) The Conceptual Nature of Natural Language Quantification. In Ruzsa I., Szabolcsi A. (eds.), *Proceedings of the '87 Debrecen Symposium on Logic and Language*, Akadémiai Kaidó, Budapest, pp. 81–94.

Lycan W. (2001) *Real Conditionals*. Oxford: Clarendon.

Matthewson L. (2001) Quantification and the Nature of Crosslinguistic Variation. *Natural Language Semantics* 9, pp. 145–189.

McCawley J. (1979) Presupposition and Discourse Structure. In Oh, Dinneen, (ed.) *Syntax & Semantics 11*.

Nute D. (1980) *Topics in Conditional Logic*, Reidel

Percus O. (2000) Constraints on Some Other Variables in Syntax. *Natural Language Semantics*, 8(3), pp. 173–229.

Roberts C. (1989) Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy*, 12, pp. 683–721.

Rooth M. (1996) Focus. In Lappin, (ed.), *Handbook of Contemporary Semantic Theory*, Blackwell.

Schein B. (2001) Adverbial, descriptive reciprocals. In Hastings R., Jackson B., Zvolenszky Z., (eds.) *Proceedings of Semantics and Linguistic Theory XI*. CLC Publications, Ithaca. Department of Linguistics, Cornell University. department of Linguistics, Cornell University.

Schlenker P. (1999/2000) *Propositional Attitudes and Indexicality (A Cross-Categorial Approach)*, Doctoral Dissertation, MIT (Introduction revised, August 2000).

Schlenker P. (1999) Semantic Uniformity, hand-out of a talk given at the First International Symposium on Linguistics (LICSSOL1), Lyons, France, October 12–15, 1999.

Schlenker P. (2000) Towards a Null Theory of Conditionals, hand-out of a talk given at the Conference on (Preferably) Non-Lexical Semantics, Paris, May 30, 2000

Schlenker P. (2001) A Referential Analysis of Conditionals. In *Proceedings of the Amsterdam Colloquium*.

Stalnaker R. (1968) A Theory of Conditionals. *Studies in Logical Theory, American Philosophical Quaterly*, Monograph: 2.

Stalnaker R. (1975) Indicative Conditionals, *Philosophia* 5, pp. 269–286.

Stanley J., Szabo Z. (2000) On Quantifier Domain Restriction. *Mind and Language*, 17.

Stechow A. von (2002) Binding by Verbs: Tense, Pronouns and Mood under Attitudes. In Lohnstein H., Trissler S. (eds.) *The Syntax and Semantics of the Left Periphery*, 44. de Gruyter: Berlin.

Stechow A. von (2003) Feature Deletion under Semantic Binding: Tense, Person and Mood under Verbal Quantifiers, *Proceedings of NELS 33*.

Stone M. (1997) The Anaphoric Parallel between Modality and Tense, IRCS Report 97-06. University of Pennsylvania, PA, USA.

Association with Focus and Choice Functions – A Binding Approach*

INGO REICH

Universität Tübingen, Germany (E-mail: ingo.reich@uni-tuebingen.de)

Abstract. In a discussion of the scopal properties of focus, Rooth [(1996) *The Handbook of Contemporary Semantic Theory*. B. Blackwell, Oxford, pp. 271–298] notes that association with focus, indefinites and in situ *wh*-phrases appear to form a class of island-insensitive phenomena, and suggests that this is the natural consequence of a common semantic property. In this paper, I analyse the hypothesis that this lowest common denominator is an analysis based on choice functions. Generalizing Reinhart's [1994 Technical Report OTS-WP-TL-94-003, OTS Working Papers, Utrecht University] binding analysis to association with focus, I propose that – similar to the case of indefinites and *wh*-phrases – focus introduces a choice function variable that operates on a contextually given set of alternatives, and gets bound by a coindexed focus-sensitive operator like, for example, the particle *only*. Having shown that this approach accounts for a wide range of data, it is compared to a similar one made in Wold [(1996) *Proceedings of SALT VI*. Ithaca, N.Y., pp. 311–328], and, in the light of the discussion of 'association with focus phrase', a possible extension of the analysis is taken into consideration.

Key words: association with focus, choice functions, island insensitivity

1. Introduction

It is a well-known fact that particles like *only*, *also* or *even* are sensitive to the focus background structure of their syntactic scope in that a difference in the placement of focus results in a difference in truth-conditions, cf. for example (1) vs. (2).

(1) John only introduced [BILL]$_F$ to Sue.
    (= Bill is the only person x s.t. John introduced x to Sue)
(2) John only introduced Bill to [SUE]$_F$.
    (= Sue is the only person y s.t. John introduced Bill to y)

In the relevant literature, this phenomenon is usually referred to as 'association with focus'. Since *only* (being a VP-adjunct) is not adjacent to the focus it is associated with, the obvious challenge is to derive this truth-conditional effect in a compositional way.

In this paper, I want to propose and investigate a binding analysis of association with focus based on choice functions: Focus introduces a choice function variable that operates on a contextually given set of alternatives, and gets bound by a coindexed focus-sensitive operator. It turns out that this way of approaching the syntax and semantics of association with focus not only accounts for a wide range of the relevant data, but also captures the fact that association with focus, indefinites and in situ *wh*-phrases are interrelated in various ways, in especially and probably most importantly that they appear to form a natural class of island-insensitive phenomena (cf. Rooth, 1996).

Before laying down the details of my own proposal in Section 3, I will first discuss two prominent previous approaches to the syntax and semantics of association with focus, the movement approach and so-called alternative semantics. Since the main purpose of Section 2 is to show that an adequate treatment of association with focus calls for a selective binding approach, I will focus on questions related to the island-insensitivity of association with focus.[1,2] In Section 4, then, the proposed choice function approach is compared to a similar approach developed by Wold (1996), followed by a short discussion of so-called 'association with focus phrase' in Section 5.

2. Two Approaches to Association with Focus

2.1. FOCUS MOVEMENT

The first approach to association with focus I want to discuss goes back to a proposal made by Chomsky (1976) for contrastive focus and assumes covert focus movement: The focused constituent is covertly moved to the focus-sensitive expression it is associated with. Given binary branching, there are two ways to implement this idea, cf. (3): Either the focus *Bill* adjoins to *only* and forms a constituent with it, cf. (3a), or it adjoins to its sister, cf. (3b).

(3) John only introduced $[BILL]_F$ to Sue.
    a. John $[[only\ [BILL]_F]\ \lambda t_1\ [_{VP}$ introduced t_1 to Sue]]
    b. John only $[_{VP}\ [BILL]_F\ \lambda t_1\ [_{VP}$ introduced t_1 to Sue]]

In the case of (3a), it is immediately clear that the focus particle *only* now has direct access to the focus it is associated with.

This does not hold for (3b), however: Without further assumptions, the moved constituent is simply semantically reconstructed to its base position within VP. For this reason, von Stechow (1981) proposes to interpret the F-Index carried by the moved constituent as triggering the construction of a structured property, i.e., an ordered pair $\langle \alpha, \beta \rangle$ consisting of a focus α and a

background β (where β (α) is well-formed and denotes the corresponding unstructured property); (3b), then, is represented along the following lines:

(4) $only(\langle Bill, \lambda x.\ introduced\ x\ to\ Sue\rangle)(John)$

Since *only* now operates on a structured property rather than an unstructured one, it has immediate access to both of its parts, the focus *Bill* as well as the background *being introduced to Sue*. Ignoring its presupposition, the semantics of *only* thus can be defined as follows: If, for any given individual b and any contextually given alternative x to the focus a, $P(x)$ (b) is true, then x must be identical to the focus a.

(5) $only(\langle a, P\rangle)(b) = 1$ iff $\forall x \in \text{alt}(a)(P(x)(b) = 1 \rightarrow x = a)$.

Because of its relatedness to the categorial semantics of wh-questions, I'd like to call this variant of the movement approach the 'categorial approach' to association with focus.

Unfortunately, the movement approach seems to face a serious problem. As has already been observed by Anderson (1972) and Jackendoff (1972), association with focus appears not to be subject to well-established constraints on overt as well as covert operator movement like, e.g., the Complex Noun Phrase Constraint (CNPC):

(6) Dr. Jones only rejected [the proposal that [BILL]$_F$ submitted]
(7) *[Which student]$_1$ did Dr. Jones reject [the proposal that t$_1$ submitted]
(8) #Dr. Jones rejected [the proposal that most students submitted]
    (For most students x: Dr. Jones rejected the proposal that x submitted)

Proponents of the movement approach thus have to conclude that there are at least two kinds of covert movement operations, one obeying so called island constraints [*wh*-movement and quantifier raising in (7) and (8), respectively] and one that doesn't [focus movement in (6)].

2.2. ALTERNATIVE SEMANTICS

Mainly to avoid this conclusion, Rooth (1985) developed an 'in situ' semantics for association with focus that has been labeled 'alternative semantics' in von Stechow (1991). The basic idea is roughly as follows: A focused constituent introduces a set of alternatives that, modulo semantic composition, percolates (upward) to the sister node of the focus particle, which in turn is able to retrieve the relevant information.

To implement this idea, Rooth (1985, 1992) postulates the existence of a second dimension of interpretation that computes for any expression α the set of its alternatives. In addition to the usual interpretation function $[\![\cdot]\!]$ he therefore introduces a recursive focus-sensitive evaluation function $[\![\cdot]\!]_F$ defined as follows:

(9)   a. $[\![\alpha]\!]_F = \{[\![\alpha]\!]\}$

   b. $[\![[\alpha]_F]\!]_F = \{u \in D_\tau; \tau = \text{type}(\alpha)\}$

   c. $[\![[\alpha\beta]]\!]_F = \{u; \exists a \in [\![\alpha]\!]_F, \exists b \in [\![\beta]\!]_F : u = a(b) \text{ or } u = b(a)\}$

If α is not focused, α does not introduce any alternatives, and therefore its alternative set is simply the singleton set whose only element is α itself, cf. (9a). But if α is focused, it intuitively does introduce alternatives, and therefore its alternative set is identical to the whole domain corresponding to α's logical type, cf. (9b), or, alternatively, to a contextually restricted subset. Finally, the interpretation of binary branching: The alternative set corresponding to the mother node is identical to the set of all well-formed function-argument combinations with elements of the alternative sets of the daughter nodes, cf. (9c).

If, for example, the domain of individuals is restricted to Bill, John, and Paul, and the mechanics in (9) are applied to example (3), we get the set of alternatives given in (10).

(10) $[\![\text{introduced } [\text{Bill}]_F \text{ to Mary}]\!]_F = \{$*introduced Bill to Mary, introduced John to Mary, introduced Paul to Mary*$\}$

Having defined for each LF-constituent α the set of its alternatives, the next step is to make these alternatives available to the focus particle *only*. To this effect, Rooth (1992) assumes that an operator $\sim$ together with a context variable Γ_i adjoins to the sister node of *only*, cf. (11b), and Γ_i is taken to be anaphorically related to the context variable C_i which is, by assumption, implicit in the semantics of *only*.

(11) a. John only introduced $[\text{BILL}]_F$ to Sue.

   b. John $[\text{only}(C_i)$ $[[\text{introduced } [\text{Bill}]_F \text{ to Sue}] \sim \Gamma_i]]$

The interpretation of the 'squiggle' operator $\sim$ then has two important effects: Focus is bound by stipulation, cf. (12b), and the interpretation of the contextual restriction C_i is restricted via coindexation with Γ_i to a subset of the set of alternatives of the sister node of *only*, cf. (12c).

(12) a. $[\![[\alpha[\sim \Gamma]]]\!] = [\![\alpha]\!]$

   b. $[\![[\alpha[\sim \Gamma]]]\!]_F = \{[\![\alpha]\!]\}$

   c. Presupposition of $[\alpha[\sim \Gamma]] : [\![\Gamma]\!] \subseteq [\![\alpha]\!]_F$

If, finally, *only* is given a semantics along the following lines,

(13) $\forall x, w : only_w(C)(\alpha)(x) = 1$ iff $\forall P \in C(P_w(x) = 1 \rightarrow P = \alpha)$.

the correct truth-conditions are derived. Since these mechanics do not refer to movement, no constraints on movement can be violated. As a consequence, so it seems, the assumption that there is only one kind of movement, namely island-sensitive movement, can be maintained.

But, as Kratzer (1991) points out, this impression is in fact wrong. To see this, consider the discourse sequence in (14a), followed by the elliptical construction in (14b). (14b) is a case of VP-ellipsis, and it is commonly assumed that on LF both the antecedent VP and the elided VP are in some sense 'identical' (cf., e.g., Sag, 1976); i.e., the LF-representation of (14b) can be considered to be the one in (14c).

(14) a. What a copycat you are! You went to Block Island because I did. You went to Elk Lake Lodge because I did. And you went to Tanglewood because I did.
　　b. I only [went to Tanglewood$_F$] because you did [e]
　　c. I only [went to Tanglewood$_F$] b. you [went to Tanglewood$_F$]

If (14c) is interpreted according to the mechanics developed in Rooth (1985, 1992), this results in a set of alternatives containing alternatives which are excluded by the preceding discourse, cf. (15b).

(15) a. I only [went to Tanglewood$_F$] b. you [went to Tanglewood$_F$]
　　b. $[\![(15a)]\!]_F$ = {I went to x because you went to y; x,y $\in$ {Block Island, Elk Lake Lodge, Tanglewood}}

To be more concrete: The proposition that *I went to Tanglewood because you went to Block Island* is clearly not a salient alternative in the context of (14a). In fact, the correct set of alternatives is the one given in (16b), where the instantiations of the F-marked constituents are always parallel. To derive this set of alternatives the F-marked constituent needs to be QRed out of VP, cf. (16a).

(16) a. I only [Tanglewood$_F$ λt_1 [[went to t_1] b. you [went to t_1]]]
　　b. $[\![(16a)]\!]_F$={I went to x because you went to x; x $\in$ {Block Island, Elk Lake Lodge, Tanglewood}}

In general, however, this results in exactly the same violations of island constraints that motivated the development of an in situ semantics for association with focus, cf. (17).

(17) a. You always contacted every responsible person before me.
　　b. I only contacted [the person who chairs [the zoning board]$_F$] before you did.

Kratzer (1991) therefore proposes to pursue a different, representational variant of alternative semantics, one that was already mentioned in Rooth (1985) and goes back to Jackendoff (1972). This proposal crucially relies on the following two assumptions about F-marking:

(18) a. All F-marked constituents bear an F-index i, $i \in$ IN.
　　b. No two constituents bear the same F-index in a given tree.

Substituting F-indices for F-markers is in fact all that is necessary to deal with the problematic cases of VP-ellipsis. Let us suppose that the F-marked expression in the antecedent VP carries an F-index, say F1. In this case, because of the identity condition on VP-ellipsis, the F-marked constituent in the elided VP carries exactly the same F-index F1, cf. (19b). If it is assumed – as Kratzer does – that a focused constituent α_{Fi} is represented by a corresponding focus variable v_i, both occurrences of the focused constituent are represented by exactly the same variable, and, consequently, the correct set of alternatives can be derived without moving the focus out of VP, cf. (19c).

(19) a. I only [went to Tanglewood$_{F1}$] because you did [e]
    b. I only [went to Tanglewood$_{F1}$] b. y. [went to Tanglewood$_{F1}$]
    c. $[\![(19b)]\!]_F = \{p; \exists h : p = [\![I \text{ [went to } v_1] \text{ b. y. [went to } v_1]]\!]^h\}$

This, however, is not yet the end of the story. As Krifka (1991) points out, there are cases of multiple focus in which association with focus behaves selectively, cf. (20).

(20) a. John only introduced BILL$_{F1}$ to Mary.
    b. John also$_{(F2)}$ only$_{(F1)}$ introduced BILL$_{F1}$ to SUE$_{F2}$.

In the context of (20a), (20b) is understood as 'it is also true for Sue that Bill is the only person which John introduced to her.' Thus, the additive particle *also* appears to associate with the prominent focus on *Sue* whereas the exclusive particle *only* associates with the second occurrence focus on *Bill*.[3,4]

It is not difficult to see that neither the denotational variant nor the representational variant of alternative semantics is able to cope with examples like (20) without moving the focus *Sue* out of the scope of *only*, the reason simply being that alternative semantics is unselective in nature. However, if the focus is moved out of the scope of the focus particle *only*, this again – as Rooth (1996, p. 288) showed himself – results in the violation of island constraints, cf. (21).

(21) a. We only$_{(F1)}$ recovered [the diary entries [that MARYLIN$_{F1}$ made about John]]
    b. We also$_{(F2)}$ only$_{(F1)}$ recovered [the diary entries [that MARYLIN$_{F1}$ made about BOBBY$_{F2}$]]

This suggests that, in general, neither the denotational nor the representational variant of alternative semantics is able to avoid reference to the kind of movement that motivated its development in the first place: island-insensitive focus movement.

3. A Binding Approach based on Choice Functions

As far as I can see, the previous discussion shows two different things: First, to avoid island-insensitive movement, some kind of in situ analysis is called for. Second, to deal with examples like (21), this analysis needs to be selective by nature. In this section, I propose a selective binding analysis based on choice functions, and I argue that an analysis along these lines is a good starting point to capture the fact that – as the following discussion is intended to show – focus, indefinites, and (in situ) *wh*-phrases are closely related phenomena.

3.1. INDEFINITES, *WH*-PHRASES AND FOCUS

Consider, for example, word order in German. As Lenerz (1977) observes, focused constituents and indefinites are subject to the same or at least similar restrictions: Given the basic word order 'indirect object (IO) before direct object (DO)', a focused direct object mustn't scramble over a non-focused indirect object, cf. (22), and an indefinite direct object mustn't scramble over a definite indirect object, cf. (23).

(22) a. ?(weil) er [das BUch]$_{F,DO}$ [dem Hans]$_{IO}$ gab
       (because) he [the BOOK]$_{F,DO}$ [the Hans]$_{IO}$ gave
       '(Because) he gave the book to Hans'
   b. (weil) er [dem Hans]$_{IO}$ [das BUch]$_{F,DO}$ gab
       (because) he [the Hans]$_{IO}$ [the BOOK]$_{F,DO}$ gave
       '(Because) he gave the book to Hans'

(23) a. ?(weil) er [ein Buch]$_{DO}$ [dem Hans]$_{IO}$ gab
       (because) he [a book]$_{DO}$ [the Hans]$_{IO}$ gave
       '(Because) he gave a book to Hans'
   b. (weil) er [dem Hans]$_{IO}$ [ein Buch]$_{DO}$ gab
       (because) he [the Hans]$_{IO}$ [a book]$_{DO}$ gave
       '(Because) he gave a book to Hans'

Secondly, it is well known that the property of being island-insensitive is not restricted to association with focus, but can be observed with respect to indefinites, too (cf. already Ross, 1967). In (24), for example, the indefinite *a student* allows for an intermediate reading, where *usually* has wider scope than *a student* and *a student* outscopes the definite complex noun phrase *[the first three proposals that…]*.

(24) Dr. Svenson usually rejects [the first three proposals [that a student submits]]. (= Usually: if there is a student *x*, then Dr. Svenson rejects the first three proposals that *x* submits.)

In Reinhart (1994, 1997) and Rooth (1996) a similar claim is also made with respect to in situ *wh*-phrases in English, cf. (25a). Surprisingly enough, however, its German counterpart is ungrammatical, cf. (25b).

(25) a. Tell me who rejected [the proposal [that who submitted]]. (Tell me about all pairs $\langle x, y \rangle$ such that x rejected the proposal that y submitted.)

   b. *Sag mir, wer das Papier ablehnte, das wer einreichte.

But, like many other languages, German exhibits a different close connection between *wh*-phrases and indefinites: Most pronominal *wh*-phrases in German – e.g. *wer* ('who'), *was* ('what'), *wo* ('where') – have a homonymous indefinite counterpart, cf. (26) vs. (27).
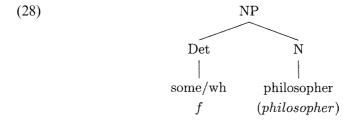
(26) Wen / Was hat Peter empfohlen?
      Who/ What has Peter recommended
      'Who /what did Peter recommend?'

(27) Peter hat wen /was empfohlen.
      Peter has someone /something recommended
      'Peter recommended someone/something'

And last but not least, *wh*-phrases, indefinites and focus all relate to the notion of 'new' information in one way or another: Whereas indefinites typically (though not exclusively) introduce new discourse referents and *wh*-phrases typically ask for 'new' information, it is the focus of a sentence that typically supplies the 'new' information asked for.

   All these similarities suggest that indefinites, *wh*-phrases, and focus form some sort of natural class of 'indefinite' or 'weak' phenomena. If this is correct, it should be reflected by a common core in their analysis.

3.2. INDEFINITES, *wh*-PHRASES AND CHOICE FUNCTIONS

Actually, Reinhart (1994, 1997) already made a proposal that goes a good way towards a unified analysis of indefinites, *wh*-phrases, and focus. To account for the island-insensitivity of indefinites and in situ *wh*-phrases (within a minimalist framework) without referring to any kind of island-insensitive movement, Reinhart (1994, 1997) proposes a binding analysis based on choice functions: Both indefinites and *wh*-phrases introduce a choice function variable that (i) operates on the restriction supplied by their complement, cf. (28),

(28)

```
                     NP
                   /    \
                 Det      N
                  |       |
               some/wh  philosopher
                  f    (philosopher)
```

and (ii) gets bound by some c-commanding operator – existential closure in the case of indefinites, a Q-morpheme in the case of in situ *wh*-phrases.[5] Technically, a choice function is any function whose domain consists of a set of non-empty sets mapping each of these sets to one of its elements, cf. (29).

(29) choice(f) = 1 iff $\emptyset \notin$ Dom(f) and $\forall X \in$ Dom(f): $f(X) \in X$.

As regards content, a choice function simply chooses an element from a given set. The observed island-insensitivity, then, apparently follows from the assumption that the choice function variable is bound by existential closure or by a Q-morpheme:

(30) Usually, $\exists_f$ Dr. Svenson rejects [the first three proposals [that f(student) submits]]
(31) Tell me who Q_f rejected [the proposal [that f(person) submitted]]

The basic idea therefore is quite parallel to that of alternative semantics: no movement, no violation of movement constraints.

3.3. CHOOSING FROM ALTERNATIVES

Having argued that indefinites, (in situ) *wh*-phrases and (association with) focus form a natural class of 'weak' phenomena, and having presented a uniform semantic analysis of indefinites and (in situ) *wh*-phrases based on choice functions, it seems reasonable to consider choice functions as the common core of this class of phenomena, and to try to extend this approach to association with focus.[6] In the following, I will first give an informal sketch of the basic idea of such an analysis, and then specify its precise semantics.

Following Rooth (1985) and many others, I take it that the notion of introducing alternatives is the basic notion when it comes to focus: Focusing a constituent evokes (the construction of) a set of contextually salient alternatives to this constituent.[7] In (32), for example, focusing the constituent *Bill* highlights the set of all the people, including Bill, who have or could have been introduced to Mary by John.

(32) John only introduced [BILL]$_F$ to Mary.

Let me introduce the notation alt(*Bill*) as a shorthand for the set of alternatives to Bill (in the relevant context). In Rooth's (1985, 1992) system, this set is made available on a level of interpretation different from the level of interpretation that calculates truth-conditions. In the following, I want to pursue the obvious alternative, namely the hypothesis that the set of alternatives is, in fact, part of the level that calculates truth-conditions. Taking the phrase 'focus introduces a set of alternatives' quite literally, let us assume that the minimal focus on the constituent Bill substitutes the set of alternatives alt(*Bill*) for the individual Bill, cf. (33).

(33) John only introduced alt(*Bill*) to Mary.

Apparently, this assumption leads to two problems. First, the resulting representation isn't interpretable, since the set alt(*Bill*) of alternatives to Bill is not a suitable argument for the predicate *introduce*. Second, we lose the information about Bill.

The solution to both of these problems is rather obvious. In an example like (33) there are, intuitively, at least two things going on: (i) Focus introduces a set of alternatives to Bill, *and* (ii) Bill is chosen from this set and enters into the predication. Since choosing from a given set is exactly what choice functions do, it would seem advisable to model this intuition by making use of exactly this kind of functions. So let us assume that, in fact, a choice function f operates on alt(*Bill*), selecting Bill from this set. Let us furthermore assume for reasons that will be clear in a moment that the set alt(*Bill*) of alternatives to Bill is the only element within the domain of f, i.e., f is a partial choice function with $\text{Dom}(f) = \{\text{alt}(Bill)\}$ and $f(\text{alt}(Bill)) = \text{Bill}$. If we call this function f_{Bill}, the resulting representation is (34).

(34) John only introduced $f_{\text{Bill}}(\text{alt}(Bill))$ to Mary.

The problem with the resulting representation is, of course, that the focus particle *only* still has no access to the information triggered by the focus on Bill. Let us therefore assume – just for the moment – along the lines of the categorial approach that the choice function f is moved to the sister of *only*, and a structured proposition consisting of the choice function f_{Bill} and a property of choice functions is generated:

(35) *only*($\langle f_{\text{Bill}}, \lambda f_i$. *John introduced* $f_i(\text{alt}(Bill))$ *to Mary*$\rangle$)

Given the representation in (35), it is rather straightforward to give a precise semantics for *only* that results in the desired truth-conditions, cf. (36): If there is a choice function f' with the same domain as f and if $\alpha(f')$ is true, then this choice function is identical to f.

(36) *only*($\langle f, \alpha \rangle$) = 1 iff $\forall f'(\text{Dom}(f') = \text{Dom}(f) \wedge \alpha(f') \rightarrow f = f')$.

In the case of (35) this is equivalent to saying that if there is a choice function f' that selects an element u from the set of alternatives to Bill so that *John introduced u to Mary* is true, then f' selects Bill.[8]

As long as movement of the focused constituent is assumed to get from (34) to (35), the sketched analysis is, of course, more or less equivalent to the categorial approach as presented in section 2.1. The central claim of this section thus is that the introduction of choice functions into the semantics of association with focus enables us to develop a variant of the categorial approach that does completely without focus movement. The analysis I want to propose is a binding analysis parallel to Reinhart's (1994) analysis of

indefinites and in situ *wh*-phrases: Focus introduces a choice function variable (operating on a set of alternatives) that is selectively bound by a coindexed binder like, e.g., the focus particle *only*, cf. (37a).

(37) a. John only$_{F1}$ introduced [BILL]$_{F1}$ to Mary.
b. *only*($\langle f_{Bill}, \lambda f_1.$ *John introduced* $f_1(\text{alt}(Bill))$ *to Mary*$\rangle$)

What needs to be shown, then, is that there is in fact a binding mechanism that gets us from an LF like (37a) to the representation in (37b), and that this mechanism can be defined in a compositional way. This is the task of the following subsection.

3.4. A SELECTIVE BINDING ANALYSIS

The following analysis rests on three sets of assumptions: assumptions about coindexation, assumptions about the interpretation of foci, and finally assumptions about the binding mechanism. Let us start with the assumptions about coindexation.

3.4.1. *Coindexation*

Following Kratzer (1991), I assume that a focused constituent carries an F-index F_i rather than a simple F-marker F and that no two expressions in a tree bear the same (kind of) F-index. Contrary to her analysis, I further assume that a focus particle carries a binder index Fi ($= \langle$ Fi, $-$p$\rangle$) which has to be distinguished from the bound index Fi ($= \langle$ Fi, $+$p$\rangle$) carried by the focus, for only the latter is subject to phonological interpretation (indicated by the feature [$+$p]).

(38) *Assumptions about F-indexing*
a. There are bound indices Fi as well as binder indices Fi.
b. No two constituents bear the same (kind of) F-index in a given tree.
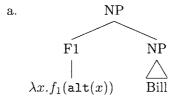
A binder index Fi and a bound index Fj are coindexed iff *i=j*. In case of (32) this results in a representation like the one given in (39).

(39) John only$_{F1}$ introduced [BILL]$_{F1}$ to Mary.

3.4.2. *The Interpretation of Focus*

Syntactically, (minimally) focusing a constituent can be conceived of as adjunction of an F-index, cf. (40a). Semantically, the focus F_i is interpreted as a kind of complex generalized skolem function, cf. (40b).[9]

(40) *Interpretation of F-indexed foci*

a.

$$\text{NP}$$

F1   NP

$$\lambda x.f_1(\texttt{alt}(x))$$ Bill

b. Let α be a branching node with daughters β and γ, where β dominates only an F-index $Fi = \langle\, Fi, +\text{p}\rangle$, then
$$[\![\alpha]\!]^w = (\lambda x.f_i(\text{alt}_c(x)))([\![\gamma]\!]^w),$$
where $\text{alt}_c(x) =$ the unique $X \in \{Y; x \in Y\}$ s.t. all and only the elements of X are salient alternatives to x in c.

The focus is interpreted as a generalized skolem function, since it maps an object of type σ to another object of type σ, and it is complex, since the mapping involves two steps: First, the alternative function alt_c maps the focused constituent x – *Bill* in (40a) – to the set of (in c) salient alternatives to x (an object of type $\langle \sigma, t\rangle$). Then, a choice function (variable) f_i chooses exactly one (arbitrary) element from this set (an object of type σ). In the case of (39) this results in the following, only partly interpreted representation:

(41) John only$_{\text{F1}}$ introduced $f_1(\text{alt}_c(\textit{Bill}))$ to Mary.

Here it is important to note that the representation of the F-index Fi involves a choice function variable f_i rather than a *constant* choice function (like, e.g., f_{Bill}). In this way, the denotation of $f_1(\text{alt}_c(\textit{Bill}))$ in (41) depends on the local variable assignment g, and the focus can be bound by a coindexed focus-sensitive operator.

3.4.3. *The Binding Mechanism*

The central element of the binding analysis is, of course, the binding mechanism itself. Following Heim and Kratzer (1998), I assume that binder indices adjoin to their sister node at LF, i.e., the (partly interpreted) representation in (41) is mapped to (42).[10]

(42) only [F1 [John introduced $f_1(\text{alt}_c(\textit{Bill}))$ to Mary]]

The crucial question then is: How can we interpret adjoined binder indices of the sort in (42)? Whatever the exact definition looks like, we know from the previous subsection that we want to end up with something like (37b), repeated here as (43), i.e., the binder index F1 maps its complement to the structured proposition consisting of the 'focus' f_{Bill} and the 'background' λf_1. *John introduced* $f_1(\text{alt}_c(\textit{Bill}))$ *to Mary*.

(43) *only*($\langle f_{\text{Bill}}, \lambda f_1.\ \textit{John introduced } f_1(\text{alt}_c(\textit{Bill})) \textit{ to Mary}\rangle$)

The 'background' of this structured proposition is simply derived by λ-abstracting over the choice function variable f_1. The compositional derivation of the 'focus' f_{Bill}, however, is somewhat more complicated: To characterize the function f_{Bill}, it is necessary to know its domain (i.e., $\{\text{alt}_c(Bill)\}$), its range (i.e., $\{Bill\}$), and which argument is mapped to which element in its range (i.e., $f_{\text{Bill}}(\text{alt}_c(Bill)) = Bill$). But even though the focused constituent is not locally accessible, this information can in fact be locally reconstructed: Suppose f is a choice function such that (λf_1. *John introduced $f_1(\text{alt}_c(Bill))$ to Mary*)(f) is defined. This is equivalent to saying that the set $\text{alt}_c(Bill)$ is an element of the domain of f. The following definition moreover ensures that the set $\text{alt}_c(Bill)$ is the only element in the domain of f:

(44) $\text{min-ch}(f, P) = 1$ iff $\text{choice}(f) \wedge f \in \text{Dom}(P) \wedge \forall g(\text{choice}(g) \wedge g \in \text{Dom}(P) \rightarrow \text{Dom}(f) \subseteq \text{Dom}(g))$

Given the minimality condition (44), f is one of the $|\text{alt}_c(Bill)|$-many choice functions mapping the set $\text{alt}_c(Bill)$ (and only this set) to one of its elements. The next question we are faced with then is: How can we guarantee on the basis of the locally available information that f chooses Bill from $\text{alt}_c(Bill)$ and not, say, Paul or John? What we know for certain is the following: If the background (λf_1. *John introduced $f_1(\text{alt}_c(Bill))$ to Mary*) is applied to f, then the resulting proposition (*John introduced f(alt$_c$(Bill)) to Mary*) is identical to the proposition that (*John introduced Paul to Mary*), if f chooses Paul; it is identical to the proposition that (*John introduced John to Mary*), if f chooses John; and it is identical to the proposition that (*John introduced Bill to Mary*), if f chooses Bill. If we make the reasonable assumption that all these propositions are pairwise distinct,[11] then there is a one-to-one mapping between those propositions and the relevant choice functions f_{Bill}, f_{John}, and f_{Paul}. Thus, to guarantee that f is in fact identical to f_{Bill}, it is sufficient to somehow state the following condition on f: The proposition in (45a) – which results from applying the 'background (λf_1. *John introduced $f_1(\text{alt}_c(Bill))$ to Mary*) to f – has to be identical to the proposition in (45b).

(45) a. *John introduced f(alt$_c$(Bill)) to Mary*
     b. *John introduced Bill to Mary*
(46) only [F1 **[** John introduced [Bill]$_{\text{F1}}$ to Mary]]

It is clear, however, that not only the proposition in (45a), but also the proposition in (45b) can be derived on the basis of the sister node γ of the binder index F1 (indicated by bold square brackets in (46)) – we just have to ignore the (bound) F-index F1 on *Bill*. Technically, this can be done by 'telling' a given interpretation function $[\![\cdot]\!]$ to ignore every (bound) F-index F1 – and only this F-index – within its scope:

(47) *'Modifying' a given interpretation function $[\![\cdot]\!]$*
    Let $[\![\cdot]\!]$ be any arbitrary function from the set $\mathcal{F}$ of all possible LF-structures into the set $\mathcal{D}$ of all possible denotations. Then $[\![\cdot]\!]_{Fi}$ is the unique function from $\mathcal{F}$ into $\mathcal{D}$ such that:
    i. $[\![\alpha_{Fi}]\!]_{Fi} = [\![\alpha]\!]$, and
    ii. $[\![\mu]\!]_{Fi} = [\![\mu]\!]$ for all $\mu \in \mathcal{F}$ that carry no index Fi.

Putting everything together, the choice function f_{Bill} can now be locally reconstructed as a definite description restricted by two conditions: f_{Bill} is the unique choice function f such that (i) f is minimally defined with respect to $\lambda f_1.[\![\gamma]\!]$, and that (ii) applying $\lambda f_1.[\![\gamma]\!]$ to f results in the same proposition that we get, if we simply ignore the (bound) F-index F1 when we interpret γ (i.e., $(\lambda f_1.[\![\gamma]\!])(f) = [\![\gamma]\!]_{F1}$), cf. definition (48).

(48) *Interpretation of adjoined binder indices Fi*
    Let α be a branching node with daughters β and γ where β dominates only an F-index F$i=\langle Fi,-p\rangle$, then
    $[\![\alpha]\!]^w = \langle \iota f[\text{min-ch}(f, \lambda f_i.[\![\gamma]\!]) \wedge (\lambda f_i.[\![\gamma]\!])(f) = [\![\gamma]\!]_{Fi}], \lambda f_i.[\![\gamma]\!]\rangle$.

According to (48), the logical form in (37a) is now in fact mapped to (37b) without making reference to any kind of focus movement, cf. (49), and is finally interpreted along the lines of (36).[12]

(49) a. John only$_{F1}$ introduced [BILL]$_{F1}$ to Mary.
    b. *only*$(\langle f_{Bill}, \lambda f_1.$ *John introduced* $f_1(\text{alt}_c(Bill))$ *to Mary*$\rangle)$

To keep things simple, I made the assumption that the sister node of the focus particle *only* is of type t. However, since *only* is a VP-adjunct, this is not necessarily the case.[13] Moreover, in examples like (50) there are good reasons to assume that *only* adjoins to the DP *Peter*.[14]

(50) [Nur$_{F1}$ PEter$_{F1}$] hat ein Buch gekauft.
    [only$_{F1}$ PEter$_{F1}$] did a book buy
    'Only Peter bought a book'

So the question emerges whether the mechanics developed so far for sisters of type t can be generalized to sisters of arbitrary conjoinable type. By using well-known techniques, this can of course be easily done: Suppose *only* adjoins to an expression γ of type $\langle \sigma, t\rangle$ and P is a variable of type σ, then simply replace every occurrence of $\lambda f_i.[\![\gamma]\!]$ in definition (48) with $\lambda f_i.[\![\gamma]\!](P)$, and $[\![\gamma]\!]_{Fi}$ with $[\![\gamma]\!]_{Fi}(P)$; finally λ-abstract over the variable P (resulting in $\lambda P.\langle \ldots \lambda f_i.[\![\gamma]\!](P) \ldots, \lambda f_i.[\![\gamma]\!](P)\rangle$).

3.5. Some central properties

Having presented the technical details of the choice function analysis proposed here, I now want to outline some of its most central properties.

First of all, and not very surprisingly, the proposed semantics predicts island-insensitivity of association with focus. The reason for that is exactly the same as it is in the case of indefinites and *wh*-phrases in Reinhart's approach: Since the analysis makes no use of movement operations, no constraints on movement can be violated.

The second issue I'd like to discuss here concerns the analysis of so-called 'free focus', i.e., constructions in which no possible binder seems to be present, cf. for example (51a). If we apply the mechanics developed so far, the translation of (51a) results in something like (51b). But since the choice function f_1 is not bound, the truth-conditions of (51b) depend on the context (i.e., the local variable assignment g), and this is certainly not true for (51a).

(51) a. John introduced [BILL]$_{F1}$ to Mary.
  b. *John introduced f_1(alt$_c$(Bill)) to Mary.*

To derive the correct truth-conditions, even so-called 'free foci' need to be bound by some covert operator. Actually, this has already been argued for in Jacobs (1984) and is known under the label 'the relational approach to association with focus.' To be specific, Jacobs proposes that 'free foci' are bound by focus-sensitive operators like assert or ask that indicate the mood of the sentence, cf. (52).[15]

(52) a. assert F1 [John introduced f_1(alt$_c$(Bill)) to Mary]
  b. assert($\langle f_{\text{Bill}}, \lambda f_1$. *John introd.* f_1(alt$_c$(Bill)) *to Mary*$\rangle$)

Strictly speaking, it is a bit sloppy to say that 'free foci' are *bound* by covert operators in Jacobs' proposal, for he assumes a categorial movement approach, and this kind of approach allows, at least in principle, an arbitrary number of focus-sensitive expressions to access one and the same focus. This may be welcome with respect to examples like (53a),[16] but in general this property seems to lead to too many non-existent readings.

(53) a. John even$_{F1}$ only$_{F1}$ introduced [BILL]$_{F1}$ to Sue.
  b. *John even [F1 [only [F1 [introduced [BILL]$_{F1}$ to Sue]]]]
  c. John [even only] F1 [introduced [BILL]$_{F1}$ to Sue]

In alternative semantics this consequence is avoided by stipulating that retrieving alternatives binds focus. Within the semantics proposed here, on the other hand, a bound focus is not available for further interpretation without further stipulations. This is simply because the interpretation of adjoined binder indices involves λ-abstraction over the choice-function variable

introduced by the focus, and another attempt to bind the same variable necessarily leads to vacuous binding. Consequently it is predicted that *even* and *only* in an example like (53a) do not 'share' their focus, but that *even* is in some sense parasitic on *only*, cf. (53c), as proposed e.g. in von Stechow (1991).

3.6. TWO PROBLEMS RECONSIDERED

Having illustrated the details of the choice function analysis and its most basic properties, it remains to be shown that this approach is able to deal with the problems outlined in section 2.2.

3.6.1. *VP-Ellipsis*

Since the assumptions about F-indexing are an extension of those assumed in Kratzer's approach, it comes as no surprise that Kratzer's Tanglewood example is handled correctly. Since both instances of the focused expression *Tanglewood* (TW) carry the same F-index after copying, identical choice-function variables are introduced which finally results in a binding effect, cf. (54).

(54) a. only [F1 [I [went to TW$_{F1}$] because you did [*e*]]]
　　 b. only [F1 [I [went to TW$_{F1}$] because you [went to TW$_{F1}$]]]
　　 c. *only*($\langle f_{TW}$, λf_1. I went to f_1(alt$_c$(*TW*)) because you went to f_1(alt$_c$(*TW*))$\rangle$)

3.6.2. *Selective Association with Focus*

Since the proposed analysis has a mechanism for coindexation at its disposal, it is also to be expected that instances of selective binding as in (56) (in the context of (55)) are treated correctly.

(55) John only introduced BILL$_{F1}$ to Mary.
(56) John also$_{F2}$ only$_{F1}$ introduced Bill$_{F1}$ to SUE$_{F2}$.

But it needs some calculation to see this. To avoid unnecessary complications, let us again assume the VP-internal subject hypothesis, and start with the logical form given in (57).

(57) also [F2 [only [F1 [John introduced Bill$_{F1}$ to Sue$_{F2}$]]]]

According to (48) (57) is interpreted as indicated in (58).

(58) *also* ($\langle \iota f$ [min-ch(f, $\lambda f_2.[\![\alpha]\!]$) $\wedge$ (($\lambda f_2.[\![\alpha]\!]$)($f$) = $[\![\alpha]\!]_{F2}$), $\lambda f_2.[\![\alpha]\!]\rangle$), where α= only [F1 [John introduced Bill$_{F1}$ to Sue$_{F2}$]].

To figure out what exactly (58) means, we need a semantics for *also* and we need to calculate $[\![\alpha]\!]$ and $[\![\alpha]\!]_{F2}$. Let's start with $[\![\alpha]\!]$. According to (48) $[\![\alpha]\!]$ is mapped to (59).

(59) *only* $(\langle \iota g[\text{min-ch}(g, \lambda f_1.[\![\gamma]\!]) \wedge ((\lambda f_1.[\![\gamma]\!])(g) = [\![\gamma]\!]_{F1}), \lambda f_1.[\![\gamma]\!]\rangle)$, where γ=John introduced Bill$_{F1}$ to Sue$_{F2}$.

Since $[\![\gamma]\!]$=*John introduced* $f_1(\text{alt}_c(Bill))$ *to* $f_2(\text{alt}_c(Sue))$ and $[\![\gamma]\!]_{F1}$=*John introduced Bill to* $f_2(\text{alt}_c(Sue))$, we know that the definite description in (59) is identical to f_{Bill} and (59) is thus reduced to (60).

(60) *only*$(\langle f_{\text{Bill}}, \lambda f_1.John\ intr.\ f_1(\text{alt}_c(Bill))\ to\ f_2(\text{alt}_c(Sue))\rangle)$

Similarly, $[\![\alpha]\!]_{F2}$ is interpreted as indicated in (61):

(61) *only*$(\langle \iota g[\text{min-ch}(g, \lambda f_1.[\![\gamma]\!]_{F2}) \wedge ((\lambda f_1.[\![\gamma]\!]_{F2})(g) = [\![\gamma]\!]_{F2,F1}), \lambda f_1.[\![\gamma]\!]_{F2}\rangle)$, where γ =John introduced Bill$_{F1}$ to Sue$_{F2}$.

Again, since $[\![\gamma]\!]_{F2}$=*John introduced* $f_1(\text{alt}_c(Bill))$ *to Sue* and $[\![\gamma]\!]_{F2,F1}$=*John introduced Bill to Sue,* we know that the definite description in (61) is likewise identical to f_{Bill}, and that (61) is reduced to (62).

(62) *only*$(\langle f_{\text{Bill}}, \lambda f_1.John\ introduced\ f_1(\text{alt}_c(Bill))\ to\ Sue\rangle)$

Because of (60) and (62) we now know that the definite description in (58) is identical to f_{Sue}, and that (58) is equivalent to (63).

(63) *also*$(\langle f_{\text{Sue}}, \lambda f_2.\ only(\langle f_{\text{Bill}}, \lambda f_1.\ John\ intr.\ f_1(\text{alt}_c(Bill))\ to\ f_2(\text{alt}_c(Sue))\rangle)\rangle)$

Given a 'standard' semantics for *also*, (63) derives the correct truth-conditions. This concludes the discussion of the central properties of the proposed choice function approach to association with focus.

4. A Comparison to Wold (1996)

Manfred Krifka and an anonymous reviewer called my attention to a proposal of Wold (1996) that bears some resemblance to the analysis proposed in section 3: Both analyses are motivated by the same set of data, and both are selective binding approaches. They differ however in the way the binding analysis is implemented: Whereas the proposal in the previous section is based on choice functions, the one in Wold (1996) is based on partial variable assignments. Consider (64).

(64) John only$_1$ introduced BILL$_{F1}$ to Sue.

As in the previous proposal the focus-sensitive particle *only* and the focused constituent *Bill* are coindexed. In contrast to the previous proposal, however, the focused constituent neither introduces a set of alternatives nor a choice function or skolem function, but is interpreted like a pronoun, if the F-index is 'old' (i.e., is an element of the domain of the local

variable assignment g), and as if it were not focused, if the F-index is 'new', cf. (65).

(65) $[\![Bill_{\text{F1}}]\!]^g = \begin{cases} g(1) & \text{if } 1 \in \text{Dom}(g), \\ Bill & \text{if } 1 \notin \text{Dom}(g). \end{cases}$

Apparently, the new/old condition on variable assignments is the counterpart to the modification of an interpretation function defined in section 3. To avoid a pronominal interpretation of the focused constituent, the F-index F1 needs to be bound by a focus-sensitive particle like *only* (or a covert operator in case of so-called 'free' focus). As regards content, Wold (1996) basically adopts the semantics for *only* proposed in Rooth (1985, 1992), cf. (66b), but gives a different derivation.

(66) a. $[\![only_1\phi]\!]^g$ defined only if $1 \notin \text{Dom}(g)$; if defined:
    b. $[\![only_1\phi]\!]^g(w) = 1$ iff $\forall p \in C(p(w) = 1 \rightarrow p = [\![\phi]\!]^g)$,
       where $C = \{[\![\phi]\!]^{g \cup \{\langle 1,x \rangle\}}; x \in D\}$.

Suppose we interpret (64), and start with the empty variable assignment $g = \emptyset$. Ignoring $only_1$ for the moment, $[\![\cdot]\!]^{\emptyset}$ interprets the focused constituent $Bill_{\text{F1}}$ as if it were not focused, since the index 1 is not an element of the domain of $\emptyset$. If however, the domain of $\emptyset$ is extended to include the index 1, and 1 is mapped to u, then $Bill_{\text{F1}}$ is interpreted as u and the proposition *John introduced u to Sue* is derived. By quantifying over all possible 'index 1 extensions' of the given variable assignment g, it is thus possible to derive the set C consisting of all and only the propositions of the form *John introduced u* to *Sue*, where u is some individual of type e. The interpretation of *only* then states that the only 'index 1 extension' in C which is true is the same proposition that we get if we ignore the F-index F1 (i.e., if we don't extend g by 1).

Despite apparent similarities, both proposals are binding analyses based on coindexation and a mechanism that makes it possible to ignore particular foci, there are also important differences with both theoretical and empirical consequences. As Krifka (1996) points out, Wold's (1996) analysis, even though it is somewhat stronger, is still a variant of Rooth's alternative semantics, for the information about the focused constituent is not accessible at the level of the focus binder *only*. As a consequence, this analysis makes it impossible to express non-existent verbs like *tolfed* (cf. the discussion in Rooth, 1996), but, on the other hand, makes it also impossible to express certain necessary restrictions on the set of alternatives (cf. the discussion of the Zimmermann example in von Stechow, 1991). Exactly the opposite is the case for the choice function approach proposed in this paper.

It is also quite informative to have a somewhat closer look at the derivation of alternative sets in Wold's analysis. As I have already mentioned above, the interpretation of the focused constituent makes no reference to the

notion of alternatives at all, but the relevant set of propositions is – modulo pronominal interpretation of the focus – exclusively derived at the level of *only* (or, more precisely, its index 1): $C = \{[\![\phi]\!]^{g\cup\{\langle 1,x\rangle\}}; x \in D\}$, where ϕ is the sister of *only* at LF. But since in general only a real subset C' of C – in (64) by assumption the set {*John introduced x to Sue; x* ∈{*John, Bill, Paul*}}– is contextually relevant, the question arises of how to restrict the set C to the relevant alternatives. The easiest way to do this is, I think, to restrict D to some contextually given set $A \subset D$, and to stipulate that $[\![\phi]\!]^g \in C_A = \{[\![\phi]\!]^{g\cup\{\langle 1,x\rangle\}}; x \in A\}$ (just to ensure that Bill is in fact one of the relevant alternatives). Technically, this seems to work. The important question to be asked now is the following: Is A given independently by some pragmatic mechanism, or is it determined 'online' as a set of *alternatives* to the focused constituent Bill? To put it somewhat differently: Is A based on an asymmetric relational notion or not? If it is, then it is necessary to refer to the focused constituent at some level during the interpretational process. This, however, is impossible within Wold's (1996) approach (even though it is part of Rooth's system). As is clear from my own analysis, I intuitively tend towards an asymmetric relational conception of alternative sets, but this matter certainly deserves more serious empirical investigation.

To conclude the comparison, I want to hint at a possible problem in Wold's approach that concerns the treatment of second occurrence expressions. As Wold (1996) shows in great detail, it is essentially no problem to deal with structures like (68) that show selective binding.

(67) John only$_1$ introduced BILL$_{F1}$ to Mary.
(68) John also$_2$ only$_1$ introduced Bill$_{F1}$ to SUE$_{F2}$.

But the analysis rests on one crucial assumption: The indices 1 and 2 have to be 'unknown' to the variable assignment g relative to which (68) is interpreted – otherwise the interpretation results in a presupposition failure, cf. (66a).[17] In the context of (67), then, the F-index on *Bill* needs to be changed to an index different from 1 (and 2). Again, this is no problem technically. But what does it mean for a focused constituent to be a second occurrence expression? And how is phonology to know that it is to interpret the F-index on a second occurrence expression differently than an F-index on first occurrence expressions? Within a selective binding approach the most straightforward answer to these questions is, I think, that (i) a second occurrence expression is simply a complete copy of its first occurrence, including its F-index, and that (ii) old indices are mapped to a different representation than new ones. This, however, is not a possible answer within Wold's (1996) analysis.

5. 'Association with Focus Phrase'

In the last part of this paper, I want to outline a possible extension of the choice function analysis presented in section 3 that deals with a problem that is known under the label 'association with focus phrase'. Building on work done by Steedman (1991) and Drubig (1994), Krifka (1996) presents a semantic argument that suggests that in an example like (69) *only* in fact associates with the complex noun phrase the *woman who introduced Bill to John* rather than with the focused constituent embedded within this island for operator movement. In Drubig (1994) complex noun phrases (and other comparable islands) in such a configuration are called 'focus phrase' (FP).

(69) Sam only talked to [the woman who introd. BILL$_F$ to John]$_{FP}$

Krifka (1996) illustrates his argument with the movement approach to association with focus. Suppose that Mary introduced Bill to John and Tim to John, and that Sam only talked to Mary. In this context (69) is intuitively true, but it is predicted to be false relative to the semantics of *only* proposed in section 2.1, cf. (70) with alt(*Bill*)={*Bill, Tim*}.

(70) $[\![only]\!]^w(\langle Bill, \lambda x.$ *Sam talked to the woman who introduced x to John*$\rangle)$
        $= 1$ iff $\forall y \in$ alt(*Bill*)((*Sam talked to the woman who introduced*
        y to John)(w) $= 1 \rightarrow y = Bill$)

Although alternative semantics uses a different semantics for *only* (a semantics that compares intensions rather than extensions, cf. section 2.2), it is nevertheless subject to exactly the same objection, cf. (71), since there is a one-to-one mapping from the set of alternatives introduced by the focus *Bill* to the restrictor C of *only*.

(71) $[\![only]\!]^w(C)$ (*Sam talked to the woman who introduced Bill to John*)=1 iff
        $\forall p \in C(p(w) = 1 \rightarrow p =$*Sam talked to the woman who introduced Bill to*
        John), where $C = \{$*Sam talked to the woman who introduced Bill to John,*
        Sam talked to the woman who introduced Tim to John$\}$

5.1. KRIFKA (1996): ASSOCIATION WITH FOCUS PHRASE

Krifka (1996) concludes from this discussion that there is something real to Drubig's distinction between 'focus phrase' and 'focus', and that the focus phrase the *woman who introduced Bill$_F$ to John* associates with *only* by island-sensitive covert operator movement, cf. (72).[18]

(72) only [the w. who intr. Bill$_F$ to John]$_{FP}$ λt_1 [Sam talked to t_1]

On the other hand, Krifka also observes that – first – the focus *Bill* can occur arbitrarily deep within FP, cf. (73),

(73) Sam only talked to [the man who mentioned [the woman who introduced
     Bill$_F$ to John]]$_{FP}$

and – second – that the set of alternatives corresponding to FP is determined
by its internal linguistic structure: In (72), for example, it is the set {*the
woman who introduced x to Bill*; $x \in$ alt(*Bill*)}. This suggests that the focus *Bill*
associates with the focus phrase FP by some island-insensitive mechanism,
and that it is linked to the focus particle *only* only in an indirect way. Krifka
(1996) assumes that the mechanism in question is alternative semantics.
Within alternative semantics the simplest way to interpret (72) is via the
condition in (74a).

(74) $[\![only[FP\phi]]\!]^w = 1$ iff
    a. $\forall y \in [\![FP]\!]_F ([\![\phi]\!]^w (y(w)) = 1 \rightarrow y = [\![FP]\!])$
    b. $\forall y \in [\![FP]\!]_F ([\![\phi]\!]^w (y(w)) = 1 \rightarrow y(w) = [\![FP]\!]^w)$
    where $[\![\phi]\!] = \lambda w \lambda x.$ *Sam talked to x in w*, $[\![FP]\!]_F = \{\lambda w.$ *the woman who
    introduced Bill to John in w*, $\lambda w.$ *the woman who introduced Tim to John
    in w*}, and $[\![FP]\!] = \lambda w.$ *the woman who introduced Bill to John in w*.

Since the meaning rule for *only* in Krifka (1996) somewhat blurs the
distinction between extensions and intensions, it is at first glance not quite
clear whether it has to be interpreted as the condition in (74a) comparing
the intensions y and $[\![FP]\!]$, or as the condition in (74b) comparing their
extensions. But if one has a closer look at it, it becomes evident that *only*
must be interpreted along the lines of (74b): Since there is a one-to-one
mapping from the set of alternatives introduced by the focus *Bill* to the
set $[\![FP]\!]_F$, only the condition in (74b) predicts the correct truth-conditions.
This is somewhat unexpected, because in general *only* quantifies over in-
tensions rather than extensions (cf. for example the discussion in Rooth,
1985), and may suggest that the observation made in Krifka (1996) is an
epiphenomenon of the semantics of determiners as proposed in von
Heusinger (1997a, 1998).

5.2. AN ALTERNATIVE: ASSOCIATION WITH FOCUS

Krifka's (1996) analysis of (69) is apparently hybrid in that it assumes two
different concepts related to focus – 'focus' and 'focus phrase' – each of which
corresponds to a different association type: Whereas association with focus
phrase is a purely syntactic notion and involves island-sensitive FP-move-
ment, association (of focus phrase) with focus is primarily a semantic concept
and is island-insensitive.

As an alternative to Krifka's proposal, I finally want to outline a possible
extension of the choice function approach presented in section 3 that does
without the notion of 'focus phrase'. Suppose that in (69) the complex noun

phrase *the woman who introduced Bill_F to John* is focused and associates with *only*, cf. (75).[19]

(75) Sam only$_{F1}$ talked to [the woman who introd. BILL$_{F2}$ to John]$_{F1}$

In this case, the embedded focus *Bill$_{F2}$* cannot associate with *only*. But – as we saw in section 3.5 – every focus needs to be bound. Since the alternatives to the complex noun phrase depend on its internal focus structure, and since the (bound) focus F1 introduces an alternative function alt, it seems reasonable to assume that it is in fact the bound focus F1 that binds the noun phrase internal focus F2 (cf. also Rooth, 1996, for a similar proposal). Given this assumption, (75) is assigned an LF along the lines of (76).

(76) only [F1 [Sam talked to [F2 [the woman...BILL$_{F2}$ to John]]$_{F1}$]]

Since until now the alternative function alt$_c$ is not sensitive to linguistic context, it is clear that this function needs to be generalized. To give a concrete example of such a generalization, let us first interpret the internal structure of the complex noun phrase. According to (48), this results in a structured individual concept as indicated in (77).

(77) $\llbracket$ F2 [the woman that introduced BILL$_{F2}$ to John]$\rrbracket^w = \langle$f$_{\text{Bill}}$, $\lambda f_2 \lambda w$. *the woman$_w$ that introd.$_w$ f_2(alt$_c$(Bill)) to John*$\rangle$.

An alternative to this structured individual concept is, of course, every structured individual concept that is identical in 'background' but differs in 'focus'. The relevant definition thus is the following (where $\pi_2(\llbracket \alpha \rrbracket^w)$ is the 'background' of $\llbracket \alpha \rrbracket^w$):

(78) $\text{alt}(\llbracket \alpha \rrbracket^w) = \begin{cases} \text{alt}_c(\llbracket \alpha \rrbracket^w) \text{ if } \llbracket \alpha \rrbracket^w \text{ is unstructured,} \\ \{\langle f, \pi_2(\llbracket \alpha \rrbracket^w)\rangle; \text{min-ch}(f, \pi_2(\llbracket \alpha \rrbracket^w))\} \text{ else.} \end{cases}$

If the set of contextually salient alternatives to *Bill* is the set {*Bill, Tim*}, then the set of alternatives to the structured individual concept in (77) is set *A* in (79).

(79) A $= \{\langle f_{Bill}, \lambda f_2 \lambda w.$ *the woman$_w$ that ... f_2(alt(Bill)) to John*$\rangle$, $\langle f_{Tim}, \lambda f_2 \lambda w.$ *the woman$w$ that ... f_2(alt(Bill)) to John*$\rangle\}$

(76) is then interpreted as (80), where *f* is the unique choice function defined on {*A*} so that the proposition λw. *Sam talked$_w$ to f(A)* is identical to the proposition λw. *Sam talked$_w$ to* $\langle$ *f$_{\text{Bill}}$, $\lambda f_2 \lambda w$. the woman$_w$ that introduced$_w$ f_2(alt(Bill)) to John*$\rangle$.

(80) $\llbracket$only$\rrbracket^w(\langle f, \lambda f_1 \lambda w.$ *Sam talked$_w$ to $f_1(A)$*$\rangle)$

Here we encounter a small technical problem: The predicate *talk to* is defined for unstructured objects of type *e*, but not for structured individual concepts. The solution is, of course, to destroy the structure of the structured individual concept again, cf. definition (80).

(81) If $[\![Pred]\!]^w$ is a function of type $\langle e, \tau \rangle$, τ some arbitrary type, and $[\![\alpha]\!]^w$ is a structured individual concept, then $[\![\mathrm{Pred}\ \alpha]\!]^w = [\![\mathrm{Pred}]\!]^w(((\pi_2([\![\alpha]\!]^w))(\pi_1([\![\alpha]\!]^w)))(w)) =: [\![\mathrm{Pred}]\!]^w((\downarrow [\![\alpha]\!]^w)(w))$.

Because of (81) we know that the proposition $\lambda w.Sam\ talked_w\ to\ \langle f_{Bill}, \lambda f_2 \lambda w.$ the woman$_w$ that introduced$_w$ $f_2(\mathrm{alt}(Bill))$ to John$\rangle$ is identical with $\lambda w.$ Sam talked$_w$ to the woman$_w$ that introduced$_w$ Bill to John, and that f chooses $\langle f_{Bill}; \lambda f_2 \lambda w.$ the woman$_w$ that introduced$_w$ $f_2(\mathrm{alt}(Bill))$ to John$\rangle$ from A. Thus (80) is equivalent to (82):

(82) $[\![\mathrm{only}]\!]^w(\langle f_{\langle f_{Bill}, \lambda f_2 \ldots \rangle},\ \lambda f_1 \lambda w.\ Sam\ talked_w\ to\ f_1(A)\rangle)$

One question remains to be answered: How do we interpret *only?* If *only* is interpreted along the lines of (36), section 3.3, i.e. by comparing the relevant choice functions $f_{\langle f_{Bill}, \lambda f_2 \ldots \rangle}$ and $f_{\langle f_{Tim}, \lambda f_2 \ldots \rangle}$ which *only* quantifies over, then we still derive the wrong truth-conditions, for each of these functions corresponds to exactly one element in the set {*Bill, Tim*} of alternatives to Bill. We know, however, from the discussion of Krifka's (1996) proposal that, in this case, we have to compare extensional rather than intensional objects. Roughly speaking, the relevant condition is as follows: If there is a choice function f' besides f such that the proposition $(\lambda w'.\ Sam\ talked_w\ to\ f'(A))$ is true in w, then f and f' map their argument(s) to the same extension in w, cf. (83).

(83) $[\![\mathrm{only}]\!]^w(\langle f, \lambda f_i.[\![\alpha]\!]\rangle) = 1$ iff $\forall f'\ \forall A[\mathrm{Dom}(f') = \mathrm{Dom}(f) \wedge A \in \mathrm{Dom}(f) \wedge (\lambda f_i.[\![\alpha]\!])(f')(w) = 1 \rightarrow (\downarrow f'(A))(w) = (\downarrow f(A))(w)]$.

In the case of (82) this basically means the following: If the proposition $(\lambda w'.\ Sam\ talked_{w'}\ to\ f_{\langle f_{Tim},\ \lambda f2 \ldots \rangle}(A))$ – which is identical with the proposition $(\lambda w'.\ Sam\ talked_{w'}\ to\ the\ woman_{w'}\ that\ introduced_{w'}\ Bill\ to\ John)$ – is true in w, then the woman that introduced Tim to John in w is the same as the woman that introduced Bill to John in w – and this is exactly what the relevant scenario was all about. This shows that in this way the correct truth-conditions are derived.

If this analysis is compared to the one proposed in Krifka (1996), one has to concede that it looks a bit more complicated, since it makes use of choice functions and structural information.[20] But it is in fact a very simple generalization of the analysis proposed in section 3: Apart from some minor technicalities, the only additional assumptions are, first, that the set of alternatives can also depend on linguistic structure, and, second, that in some cases the focus particle *only* only compares extensional objects rather than intensional ones. The latter assumption may cast some doubt on both the analysis proposed here and the one in Krifka (1996), but at the moment I see no straightforward way to derive the correct truth-conditions without having reference to the extensions of the relevant complex noun phrases. The way

this information is made available is, however, quite different from the analysis in Krifka (1996): Whereas in Krifka's approach the existence of a 'focus phrase' is assumed which is moved to the focus phrase sensitive particle *only*, the choice function approach starts from the assumption that the complex noun phrase is focused and associates with the focus particle *only* in the same way as the embedded focus associates with the focus on the complex noun phrase. This raises of course the question of why this complex noun phrase is focused to begin with, and why this focus – in contrast to the embedded focus – seems to be sensitive to restrictions on operator movement. In Reich (2003) I argued that, first, *wh*-phrases (in German) behave island-sensitively, and that, second, complex *wh*-phrases induce corresponding foci in answers. Thus, in a model like the one in Roberts (1996) that relates every utterance to an explicit or implicit *wh*-question, the focus on the complex noun phrase as well as its apparent island-sensitivity can be considered to be an epiphenomenon of the properties of complex *wh*-phrases. This, of course, is an empirical question that needs thorough investigation.

6. Summary

In this paper, I have proposed a selective binding approach to association with focus based on choice functions. Since it is a binding approach, focus is interpreted in situ and association with focus is not subject to any island constraints imposed on LF-movement; since it is selective, cases of selective association with focus can be accounted for, too. Moreover, introducing choice functions into the analysis of association with focus allows for an explicit treatment of alternative sets and, more importantly, to consider indefinites, (in situ) *wh*-phrases and association with focus as a natural class of 'weak' phenomena. Having compared this analysis to a similar proposal made by Wold (1996), I have proposed a possible extension of the choice function approach that copes with the so-called problem of association with 'focus phrase' without assuming the existence of such an additional concept.

Notes

[1] Cf. von Stechow (1991) and Rooth (1996) for a more comprehensive comparison.

[2] I also refrain from a discussion of more indirect approaches like 'association with presupposition' (but cf. Rooth, 1999, and Fox et al., 2001) or Schwarzschild's (1997) approach based on 'felicity conditions' for utterances. If Beaver and Clark (2000) are right in claiming that a 'mixed' approach to association with focus phenomena is called for, this presentational

decision is, to a certain extent, justified by restricting the following discussion to focus particles like *only* or *also*.

[3] *Bill*, being a second occurrence of the focused constituent *Bill* in (20a), apparently lacks a pitch accent. However, as is argued in Krifka (1997), Beaver and Clark (2000) and references therein, there is good evidence from weak pronouns that nevertheless second occurrence expressions (SOEs) are syntactically focused and that this syntactic focus is realized by features other than pitch, e.g. duration or amplitude. Here, and in the following, SOEs are indicated by small caps.

[4] For a thorough investigation of the syntax and semantics of German focused and unfocused *auch* (also) as well as a proposal for a unified treatment of both uses, cf. Reis and Rosengren (1997).

[5] The introduction of choice functions into the analysis of indefinites and in situ *wh*-phrases raises a lot of important questions concerning their exact definition in case the complement's extension is empty (cf. e.g. Winter, 1997, von Stechow, 2000) as well as in intensional contexts (the so-called 'Donald Duck' problem; cf. i.a. Reinhart, 1994, von Stechow, 2000, Dekker, this volume), questions concerning the applicability to overtly moved *wh*-phrases (cf. von Stechow, 2000) or definites (cf. von Heusinger, 1997b), and finally questions concerning the binding process (cf. e.g. Reinhart, 1997 vs. Kratzer, 1998). Although, of course, some of these problems carry over to the choice function approach to association with focus proposed in this paper, I can't enter into a discussion of these matters here for obvious reasons (but cf. e.g. Farkas, 2002, and references therein, for recent discussion).

[6] This, of course, raises the question why (in situ) *wh*-phrases in German behave island-sensitively, while indefinites and association with focus do not. In Reich (2002b, 2003) it is argued (i) that in German *wh*-phrases should be analysed as functional expressions with an indefinite core, and (ii) that the functional part of *wh*-phrases is subject to movement operations and thus triggers island effects. In a nutshell, the answer is that *wh*-phrases contain an additional island-sensitive component that the semantics of indefinites and association with focus lack.

[7] For a somewhat different position cf. e.g. Schwarzschild (1997).

[8] Here it is absolutely crucial that the domain of f is in fact minimal in the sense that it only contains the set of alternatives to Bill. If, for example, the domain of f contains another set, say X, then any choice function f' with $f'(\text{alt}(Bill)) = $ Bill, but $f'(X) \neq f(X)$ will falsify the condition in (36). The minimality requirement thus captures the fact that other arguments than alt(*Bill*) are simply not relevant for calculating the truth-conditions of (32).

[9] For reasons of readability, I omit in the following the variable assignment g.

[10] To keep things simple, I will also assume, here and in the following, that the subject *John* is syntactically reconstructed to its base position within VP.

[11] Without further assumptions, this analysis apparently cannot cope with examples involving hyperintensionality. Since this is a general problem of the model this analysis is couched in, I won't enter into a discussion of this issue here.

[12] It should be noted here that the structured proposition constructed via definition (48) is always immediately destroyed by the focus-sensitive particle to which the binder index originally belongs. In this way, complicating projection principles for structured objects, as stipulated in Krifka (1991), are avoided.

[13] In fact this is only possible under the VP-internal subject hypothesis.

[14] But cf. also the discussion in Jacobs (1983), Büring and Hartmann (2001).

[15] In Reich (2002b, 2003) it is argued that this notion needs to be generalized to other rhetorical relations like, e.g., contrast and answer.

[16] Cf. also Schwarzschild (1997) for another interesting case.

[17] The presupposition in (66a) ensures that there is a way to interpret the focused constituent as being unfocused.

[18] The LF in (72) is, again, simplified by assuming that the subject is syntactically reconstructed to a position within VP.

[19] Structures of this kind are easily interpreted by phonology, cf. Schwarzschild (1999) and Reich (2003).

[20] In Reich (2002a, 2003) it is argued that the use of choice functions and structural information is essential for an explanation of the so-called pied piping problem in the related case of *wh*-questions.

References

Anderson S. R. (1972) How to get *even*. *Language*, 48, pp. 893–905.

Beaver D. I., Clark B. (2000) "Always" and "only": Why not all Focus Sensitive Operators are Alike. First Draft.

Büring D., Hartmann K (2001) The Syntax and Semantics of Focus-Sensitive Particles in German. *Natural Language and Linguistic Theory,* 19, pp. 229–281.

Dekker P. (this volume) A Pragmatic View upon Indefinites. *Research on Language and Computation.*

Drubig H. B. (1994) Island Constraints and the Syntactic Nature of Focus and Association with Focus, Technical Report 51, Arbeitspapiere des Sonderforschungsbereichs 340, Stuttgart/Tübingen.

Farkas D. (2002) On Some Indefinites in English. In: von Heusinger K., Kempson R., Meyer-Viol W. (eds.), *Proceedings of the Workshop "Choice Functions and Natural Language Semantics",* Arbeitspapier 110, pp. 35–62.

Fox D., Nissenbaum J., Sauerland U. (2001) Does Focus Introduce a Presupposition?, Ms., MIT, Harvard, and Tübingen University.

Heim I., Kratzer A. (1998) *Semantics in Generative Grammar*. Blackwell, Oxford.

Jackendoff R. (1972) *Semantic Interpretation in Generative Grammar*. MIT Press, Cambridge, MA, USA.

Jacobs J. (1983) *Fokus und Skalen. Zur Syntax und Semantik von Gradpartikeln im Deutschen*. Niemeyer, Tübingen.

Jacobs J. (1984) Funktionale Satzperspektive und Illokutionssemantik. *Linguistische Berichte* 91, pp. 25–58.

Kratzer A. (1991) 'Modality'. In von Stechow and Wunderlich (1991), pp. 639–650.

Kratzer A. (1998) Scope or Pseudoscope? Are there Wide-Scope Indefinites?. In: Rothstein S. (ed.), *Events and Grammar*. Kluwer, Dordrecht pp. 163–196.

Krifka M. (1991) A Compositional Semantics for Multiple Focus Constructions. In: Moore S., Wyner A. (eds.), *Proceedings of SALT I. Ithaca*, N.Y., Cornell University. pp. 127–158.

Krifka M. (1996) Frameworks for the Representation of Focus. In: Krujiff G.-J., Morrill G., Oehrle D. (eds.), *Formal Grammar*. Prague: Eighth European Summer School for Logic, Language, and Information, pp. 99–112.

Krifka M. (1997) Focus and/or Context: A Second Look at Second Occurrence Expressions. In:. Kamp H., Partee B. H. (eds.), *Context Dependence in the Analysis of Linguistic Meaning*, University of Stuttgart Working Papers.University of Stuttgart, Proceedings of the Workshops in Prague and Bad Teinach, 1995, pp. 253–275.

Lenerz J. (1977) *Zur Abfolge nominaler Satzglieder im Deutschen*. Narr, Tübingen.

Reich I. (2002a) Pied Piping and the Syntax and Semantics of Complex wh-Phrases. In: Mauck S., Mittelstaedt J. (eds.), *Georgetown University Working Papers in Theoretical Linguistics, Volume 2 (2002)*. Georgetown University, pp. 263–286.

Reich I. (2002b) Question/Answer Congruence and the Semantics of wh-Phrases. *Theoretical Linguistics* 28, 73–94.

Reich I. (2003) *Frage, Antwort und Fokus*, No. 55 in studia grammatica. Akademie Verlag, Berlin.

Reinhart T. (1994) Wh-in-situ in the Framework of the Minimalist Program, Technical Report OTS-WP-TL-94-003, OTS Working Papers, Utrecht University.

Reinhart T. (1997) Quantifier Scope: How Labor is Divided between QR and Choice Functions. *Linguistics and Philosophy* 20, pp. 335–397.

Reis M, Rosengren I. (1997) A Modular Approach to the Grammar of Additive Particles: the Case of German Auch. *Journal of Semantics* 14, pp. 237–309.

Roberts C. (1996) Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics. In: Hak Yoon J., Kathol A. (eds.), *Ohio State University Working Papers in Linguistics, Volume 49*. Ohio State University.

Rooth M. (1985) Association with Focus. Ph.D. thesis. University of Massachusetts, MA, USA

Rooth M. (1992) A Theory of Focus Interpretation. *Natural Language Semantics* 1, pp. 75–116.

Rooth M. (1996) Focus. In: Lappin S. (ed.), *The Handbook of Contemporary Semantic Theory*. B. Blackwell, Oxford pp. 271–298.

Rooth M. (1999), Association with Focus or Association with Presupposition. In: Bosch P., van der Sandt R. (eds.), *Focus*. Cambridge University Press, Cambridge. pp. 232–244.

Ross J. R. (1967) Constraints on Variables in Syntax, Ph.D. thesis. MIT.

Sag I. A. (1976) Deletion and Logical Form, Ph.D. thesis. MIT.

Schwarzschild R. (1997) Why Some Foci Must Associate. Ms, Rutgers University.

Schwarzschild R. (1999) GIVENness, AVOIDF and Other Constraints on the Placement of Accent. *Natural Language Semantics* 7, pp. 141–177.

Steedman M. (1991) 'Surface Structure, Intonation, and Focus'. In: Klein E., Veltman F. (eds.), *Natural Language and Speech*. Dordrecht Kluwer, pp. 260–296.

von Heusinger K. (1997a) Focus in Complex Noun Phrases. In: Dekker P., Stokhof M., Venema Y. (eds), *Proceedings of the 11th Amsterdam Colloquium*. University of Amsterdam. pp. 49–54.

von Heusinger K. (1997b) *Salienz und Referenz. Der Epsilonoperator in der Semantik der Nominalphrase und anaphorischer Pronomen*, No. 43 in studia grammatica. Akademie Verlag, Berlin.

von Heusinger K. (1998) Association with Focus in Definite DPs, Technical Report LRC-98-01, University of California, Santa Cruz.

von Stechow A. (1981), Topic, Focus and Local Relevance. In: Klein W., Levelt W. (eds.), *Crossing the Boundaries in Linguistics*. D. Reidel, Dordrecht: pp. 95–130.

von Stechow A. (1991) Current Issues in the Theory of Focus. In von Stechow and Wunderlich (1991), pp. 804–825.

von Stechow A. (2000) Some Remarks on Choice Functions and LF-Movement. In: von Heusinger K., Egli U. (eds.), *Reference and Anaphoric Relations*. Kluwer, Dordrecht: pp. 193–228.

von Stechow A., Wunderlich D. (eds.), 1991, *Semantik – Ein internationals Handbuch zeitgenössischer Forschung. Semantics – An International Handbook of Contemporary Research*. de Gruyter, Berlin.

Winter Y. (1997) Choice Functions and the Scopal Semantics of Indefinites. *Linguistics and Philosophy* 20, pp. 399–467.

Wold D. (1996) Long Distance Selective Binding: The Case of Focus. In: Galloway T., Spence J. (eds.), *Proceedings of* SALT VI. Ithaca, N.Y., pp. 311–328.