

INFORMATION TO USERS

This reproduction was made from a copy of a document sent to us for microfilming. While the most advanced technology has been used to photograph and reproduce this document, the quality of the reproduction is heavily dependent upon the quality of the material submitted.

The following explanation of techniques is provided to help clarify markings or notations which may appear on this reproduction.

1. The sign or "target" for pages apparently lacking from the document photographed is "Missing Page(s)". If it was possible to obtain the missing page(s) or section, they are spliced into the film along with adjacent pages. This may have necessitated cutting through an image and duplicating adjacent pages to assure complete continuity.
2. When an image on the film is obliterated with a round black mark, it is an indication of either blurred copy because of movement during exposure, duplicate copy, or copyrighted materials that should not have been filmed. For blurred pages, a good image of the page can be found in the adjacent frame. If copyrighted materials were deleted, a target note will appear listing the pages in the adjacent frame.
3. When a map, drawing or chart, etc., is part of the material being photographed, a definite method of "sectioning" the material has been followed. It is customary to begin filming at the upper left hand corner of a large sheet and to continue from left to right in equal sections with small overlaps. If necessary, sectioning is continued again—beginning below the first row and continuing on until complete.
4. For illustrations that cannot be satisfactorily reproduced by xerographic means, photographic prints can be purchased at additional cost and inserted into your xerographic copy. These prints are available upon request from the Dissertations Customer Services Department.
5. Some pages in any document may have indistinct print. In all cases the best available copy has been filmed.

**University
Microfilms
International**

300 N. Zeeb Road
Ann Arbor, MI 48106

8410273

Chierchia, Gennaro

TOPICS IN THE SYNTAX AND SEMANTICS OF INFINITIVES AND GERUNDS

University of Massachusetts

PH.D. 1984

**University
Microfilms
International** 300 N. Zeeb Road, Ann Arbor, MI 48106

TOPICS IN THE SYNTAX AND SEMANTICS
OF INFINITIVES AND GERUNDS

A Dissertation Presented
By
Gennaro Chierchia

Submitted to the Graduate School of the
University of Massachusetts in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY
Linguistics
February 1984

© Gennaro Chierchia
All Rights Reserved

TOPICS IN THE SYNTAX AND SEMANTICS
OF INFINITIVES AND GERUNDS

A Dissertation Presented
By
Gennaro Chierchia

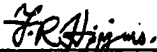
Approved as to style and content by



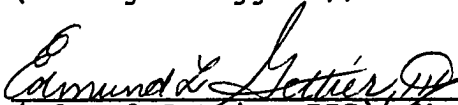
(Barbara Hall Partee), Chairperson of
Committee



(Emmon Bach), Member



(F. Roger Higgins), Member



(Edmund Gettier III), Member



(Emmon Bach), Department Head,
Linguistics

ACKNOWLEDGEMENT

Various circumstances in my life prevent me from being a student any longer. For this reason, it was necessary to finish the present dissertation. But this should not be taken to imply that I regard it in any way as a complete piece of work.

The University of Massachusetts at Amherst is a very special place for theoretical linguistics. Many things contribute to make it special: the quality of its faculty, its long peaceful winters, the fact that people working within different frameworks communicate with each other. I benefitted enormously from all this and it is impossible for me to thank individually all the regular and occasional Amherstians who directly or indirectly helped me out through these years.

Some of my teachers, like Lisa Selkirk, Alan Prince and Lyn Frazier, influenced deeply my way of thinking about language. Ed Gettier and especially Hans Kamp helped me a lot to improve the first part of the present work. Edwin Williams, through his writings and our conversations, also had an important influence. F. Roger Higgins' thorough comments on major and minor aspects of my work have been of invaluable help.

Before getting into linguistics I studied philosophy in Italy. On the plane during my first trip to the

States, I decided to say goodbye to hard-core philosophy by reading a paper by Cocchiarella which sat in the back of my mind for a while to pop out again when I began to think about G. Carlson's beautiful theory of bare plurals. Since then, a long correspondence with Cocchiarella and further acquaintance with his work had a tremendous impact on me, which is reflected throughout this work.

I can now confess that it took me a while to get used to Emmon Bach's teaching style. He (with Lyn Frazier) taught me my first syntax class, the hardest (and perhaps most fruitful) I ever took. Then, through more courses and long conversations, I really came to admire Emmon's deep, unconventional, insightful mind. My debts to his work, his teaching and his friendship are enormous.

Barbara Partee's fame among formal semanticists is one of the things that attracted me to Amherst. I was delighted to meet her and enjoyed thoroughly working close to her these years. Her continuous, varied and challenging teaching, her ideas and her friendship were absolutely crucial for me. I think that if I hadn't met her, I wouldn't be in this field today. For, let's face it, there are easier ways of making a living.

In the ideal worlds, students have a privileged access channel to their teachers' time and minds. We know that our world is not very close to the ideal ones. In this regard, however, Amherst is a fragment of an ideal

world. Moving away from it, I regret that I am bound to lose that privileged access channel.

I shared with my fellow students long discussions on the meaning of linguistics and life, which will be a nice part of my student memories. I am especially grateful to my friends Yasuaki Abe, Dan Finer, David Lebeaux, James Pustejovsky and, in particular, Mats Rooth.

I also wish to thank the Commission for Cultural Exchange between Italy and the U.S.A. for awarding me a Fulbright scholarship that partially supported me during my first year in Amherst.

My parents and my uncle Gianni Marinangeli played an important role supporting me morally and materially during all my student years. However, the person to whom I owe most is my wife Isa. Since we met, 15 years ago, she always was near me, carrying the burden of staying optimistic during various difficult moments, and keeping me up, often against myself, with her love. Now we also have little Gabriele to keep us up, especially at night.

ABSTRACT

Topics in the Syntax and Semantics of
Infinitives and Gerunds

February 1984

Gennaro Chierchia

Dottore in Filosofia, University of Rome

Ph.D., University of Massachusetts

Directed by: Professor Barbara Hall Partee

This work is intended as an inquiry into the nature of predication in natural languages. The theory of properties and predication built into Montague semantics is based on Russell's theory of types and a possible world approach to intensionality. It is argued that any theory so conceived turns out to be inadequate for the analysis of linguistic facts related to nominalization. An alternative theory inspired by work on second order logic by Cocchiarella is developed. Such a theory tries to make precise Frege's insight that properties have two modes of being: as predicable structures properties are intrinsically unsaturated (something like propositional functions); such structures are, however, systematically correlated to individuals (actions, kinds, etc.). Nominalized predicates purport to refer to individual correlates of propositional functions.

This semantic model of predication (SMP) is built into a Montague-style categorial grammar and argued to provide several results. In the first place, it embodies constraints on possible types of natural language meanings that rule out a number of a priori conceivable and yet unattested grammatical processes. It also sheds light on various distributional paradigms of English predicative expressions. In particular, the distribution of PRO (the null subject of infinitives postulated within the transformational tradition) is shown to be deducible from SMP.

Our account is crucially based on the hypothesis that infinitives and gerunds are not clausal or propositional constructions and tries to provide support for such an hypothesis, also drawing from analyses of various anaphoric phenomena.

If infinitives and gerunds do not have a propositional nature, a new account of control is called for. Control is argued to be an entailment licensed by a certain class of predicates. A general characterization of such entailments is offered in the form of a meaning postulate. It is furthermore argued, following an idea of Williams, that semantic predication is encoded in the syntax as some sort of coindexing between subjects and predicates. This device interacts with the notion of semantic function-argument structure to provide elements of a theory of binding not based on phrase structural configurations.

TABLE OF CONTENTS

ACKNOWLEDGEMENT	iv
ABSTRACT	vii
Chapter	
I. THEORIES OF PROPERTIES AND NATURAL LANGUAGE SEMANTICS	1
1. Nominalization and the Theory of Types	5
1.1. Infinitives and gerunds	6
1.2. Other nominalization phenomena	17
1.3. Conclusions	21
2. On Weak Intensionality	22
2.1. Conjunction	24
2.2. Passive	26
2.3. Predicate modifiers	31
2.4. Summary	37
3. A Second Order Theory of Properties	38
3.1. Model theoretic vs. axiomatic theories of properties	40
3.2. HST*: a second order counterpart of type theory	45
3.3. $\lambda \square$ HST*: the syntax and (weakly intensional) semantics of a modal λ -extension of HST*	57
3.4. IL*: an HST*-based logical form for English	64
4. IL* as a Restrictive Component of the Grammar	74
4.1. The three-layers hypothesis	77
4.2. The "no functor anaphora" constraint	83
4.3. Logical types vs. semantic sorts	88
5. Some Comparisons	90
5.1. Cresswell's models for λ -categorial languages	91
5.2. Parsons' "floating types"	93
5.3. Scott's semantics for the untyped λ -calculus	99
5.3.1. A sketch of Scott's semantics	100
5.3.2. Sketch of an application	103
5.4. First order theories of properties	111
5.5. Concluding remarks	119
Footnotes	121
II. LOCALLY CONFIGURATIONAL GRAMMAR	123
1. The Notion of Function-Argument Structure	124
2. The Format of Syntax	134

3.	The Syntax-Semantics Map	143
3.1.	Categories vs. types: the functional correspondence constraint	143
3.2.	Some sortal distinctions	154
3.3.	A note on semantic deviance	170
4.	Features	175
4.1.	On the nature of morphological operations	178
4.2.	Government and percolation	189
4.3.	Consequences and comparisons	196
4.3.1.	Possible governors	196
4.3.2.	Possible agreement features	197
4.3.3.	The phrasal character of inflectional processes	199
4.3.4.	The domain of government and agreement	200
4.3.5.	Long distance agreement	202
4.3.6.	Conclusions	205
4.4.	Governing functions, heads and the "case filter"	206
	Footnotes	211
III.	DISTRIBUTIONAL PROPERTIES OF PREDICATIVE EXPRESSIONS AND THE SEMANTICS OF NOMINALIZATION	215
1.	On the Constituency of Infinitives and Gerunds	218
1.1.	A non argument	218
1.2.	An argument	224
1.3.	The semantics of nominalization and the "PRO theorem"	231
2.	A Typology of Clausal Constructions and the Notion of Finiteness	239
2.1.	Finite clauses	240
2.2.	NP-to-VP and for-to clauses	245
2.3.	On gerunds	255
2.4.	Conclusions	266
3.	A Case Study: Lexical Subjects in Italian Infinitives	267
3.1.	The data	267
3.2.	Rizzi's account	271
3.2.1.	Aux-to-COMP	271
3.2.2.	Exceptional case marking	272
3.2.3.	Case assignment to traces in COMP	274
3.2.4.	Subcategorizations	275
3.3.	A categorial alternative	277
3.3.1.	The PA property	279
3.3.2.	Some consequences	280
3.3.3.	The WHE property	283
	Footnotes	286

IV. CONTROL AND SEMANTIC STRUCTURE	291
1. Relevant Data	294
1.1. Obligatory control	295
1.2. Semi-obligatory control	298
1.3. Prominence control	302
2. Control as a Semantic Entailment	305
2.1. Conditions on admissible models of English	306
2.2. On the modal nature of control predicates	311
3. Control and Thematic Structure	322
3.1. Θ -roles and complexes: an excursus into English Metaphysics	324
3.2. Further conditions on Θ -roles	329
4. Control and Locality	334
4.1. Syntactic approaches	336
4.2. Semantic accounts	339
5. Inheritance	344
6. Capturing Obligatory Control	349
7. Capturing Semi-Obligatory Control	356
8. Capturing Prominence Control	362
9. Equi vs. Raising	368
9.1. On virtual arguments	373
9.2. The virtuality-opacity correspondence	376
10. Some Comparisons	381
10.1. Syntactic approaches	382
10.2. Semantic approaches	393
10.3. Conclusive remarks	398
Footnotes	400
V. INFINITIVES, GERUNDS AND ANAPHORA	403
1. External Anaphora	404
1.1. PRO _{arb} : its truth conditional import	404
1.2. Sloppy and non sloppy anaphora	417
1.3. A few more sound arguments	427
1.4. Summary	431
2. Internal Anaphora	431
2.1. Non reflexive pronouns	432
2.2. Reflexives	442
3. Opacity and Control	450
3.1. A "propositional" account	450
3.2. Syntactic predication and control	456
3.3. Conclusions	460
Footnotes	467
REFERENCES	470

C H A P T E R I
THEORIES OF PROPERTIES
AND NATURAL LANGUAGE SEMANTICS

It is quite uncontroversial that a semantic theory for natural languages should specify what properties are and what it means to attribute a property to an individual in an act of predication. A theory of properties and predication seems in fact to be one of the most central tasks (if not the most central one) that any general semantics has to face. Montague's semantics provides a general theory of properties that relies on two major subcomponents: a possible world analysis of intensionality and the theory of simple types.

According to the first of these subcomponents a property is analyzed as a function from possible worlds into extensions (sets or characteristic functions of sets). A consequence of this view is that two properties having the same extension at all possible worlds will be identical. So any two necessarily equivalent properties (such as being autoidentical and being such that $2 = 1 + 1$) will be the same entity. While such a notion of

property allows us to model in an interesting way certain types of intensional contexts, it runs into serious problems, as is well known, in connection with propositional attitudes and related notions. We will therefore call a possible world based notion of property (as well as any other theory that identifies necessary equivalents) "weakly intensional".

The second subcomponent of Montague semantics, i.e. the theory of simple types, allows us to generalize the theory of properties sketched above from basic individuals (or urelements) to everything we might possibly want to talk about (including any higher order entities) in a paradox-free way. According to the theory of types, properties have to be ranked in a hierarchy based on the things they can be meaningfully attributed to. So let e be the type of basic entities (as in standard Montague Grammar, MG from now on). A 1-place property of basic entities will be of type $\langle e, t \rangle$.¹ A property of properties of basic entities will have to be of type $\langle \langle e, t \rangle, t \rangle$. and so on. In general, in a predicative structure of the form $\beta(\alpha)$, β (the predicate) will have to be of a higher type than α (the argument). A consequence of this is that cases of self-application of properties (i.e. predicative structures of the form $\beta(\beta)$) are ruled out once for all as meaningless.

Apart from being a theory of predication, type theory plays another role within the semantic component of a grammar. It provides a classificatory device for semantic universes. Each possible semantic object is classified as belonging to a certain type. Types are then systematically related to syntactic categories. The connection between categories and types should provide a representation of how items of different categories differ in meaning.

The original synthesis of type theory and possible world semantics developed by Montague has proven to be able to take natural language semantics out of the foggy limbo in which it used to be confined. Formal semantics seems now to be established as a reasonably explicit level of linguistic theory in which it has been shown that significant empirical generalizations can be optimally captured. However, Montague semantics left several major problems open. Consider, for instance, type theory as a classificatory device for semantic universes. There is a quite clear sense in which this role type theory is inadequate. For one thing, it makes too many distinctions. It individuates classes of entities that no natural language "uses" as semantic values for items of any category. No natural language has categories related to entities of type $\langle t, e \rangle$. In some way, though, type theory also makes too few distinctions. For instance, the

logical type of verbs like try and of adverbs like slowly turns out to be the same in Montague's system, and it seems hard to believe that the distinction between these two classes of items is purely syntactic. Further elaborations of Montague's system have not managed to do any better in this regard. So type theory provides classificatory criteria that match rather defectively those that seem to be operating in a grammar for natural language.

In addition to type theory's defects as a classificatory schema for semantic domains, I believe that Montague semantics also runs into serious problems because of its theory of predication. This is the issue that we are going to take up in some detail in the present chapter. We will point at several major difficulties that face a weakly intensional, type theoretic notion of property once we try to base on it a grammar of English predicative constructions. We will argue that a successful analysis of various linguistic phenomena is only possible if we go beyond the limits of weak intensionality and type theory. We will then sketch a theory of properties, based on work by Nino Cocchiarella, which allows for a satisfactory treatment of the phenomena that turn out to be problematic for Montague semantics. Having done this, we will consider how our novel approach does (in comparison with type theory) as a classificatory device for semantic domains. Finally, we will attempt a first

approximation comparison between the theory we adopt and other available alternatives.

1. Nominalization and the Theory of Types.²

In PTQ, Montague proposes that there exists an extremely tight relation between syntactic categories and semantic types. In fact, he claims that such a relation can be characterized as a (type assignment) function. Such a function is recursively specified. First the types associated with primitive syntactic categories are given; then the type of complex categories is specified in terms of the types of their input categories. According to this view, then, by knowing that an expression is, say, a Noun Phrase (NP), we know, by applying the type assignment function to it, what sorts of semantic objects it denotes. Let us call this the "transparency principle". What we will try to do in what follows is to argue first that on the basis of this principle it is quite difficult to deal with cases of nominalization and second that there is no way to weaken this principle in a fully satisfactory manner. We will not claim that a type theoretic approach to nominalization is altogether impossible, but that a grammar for a sufficiently rich fragment of English based on type theory does not yield the best account of nominalization we can construct.

1.1 Infinitives and gerunds.

By nominalization we refer to grammatical processes whereby expressions that are semantically associated with properties are transformed into (or better are related to) noun-like expressions (i.e. expressions that behave more like names than like predicates). Consider the following examples:

- (1) John forces Bill to be home
- (2) John accused Bill of being a communist

To present our argument we have to fill in a few preliminaries concerning the syntactic constituency of the second underlined expressions in (1)-(2). Within the tradition of MG (see e.g. the essays in Partee 1976) the underlined expressions in (1) and (2) would be analyzed as some sort of verbal constituent. For the sake of explicitness, we will assume that their category is IV (for intransitive verb (phrase)) to which we may add a feature specification such as INF or GER. There are a number of other syntactic theories which would analyze the structures in question as some sort of verbal constituent. Examples are recent developments and extensions of \bar{X} -theories, such as Generalized Phrase Structure Grammar (GPSG, see e.g. Gazdar and Pullum 1982) or lexical Functional Grammar (LFG, see e.g. Bresnan 1982). A different hypothesis on the syntactic constituency of infinitives is put forth within Chomsky's Government and Binding theory (GB, see Chomsky 1982) and

related frameworks (e.g. Williams 1980). According to these theories infinitives should be analyzed as being clausal (i.e. sentential constituents). We will follow for the moment the former tradition in assuming that infinitives are essentially verb phrases (VPs) of some sort. In chapter III we will discuss in greater detail these various hypotheses on the syntactic constituency of the constructions in question.

Let us now turn to a first approximation hypothesis concerning the semantics of the underlined constructions in (1) and (2). Clearly the infinitive in (1) and the gerund in (2) are somehow related to the underlined NPs. This relation looks a lot like the one that obtains in simple matrix sentences between a predicate and its subject. There is clearly a sense in which Bill and to be home in (1) are related to each other in much the same way in which they are related in Bill is home. So, it seems fairly natural to hypothesize that infinitives and gerunds enter some sort of predication relation with some NP in the sentences where they occur. But what sort of entities can be predicated of something? Well, properties. This would suggest that infinitives and gerunds should be semantically associated with properties. This was, in fact, the hypothesis adopted by Montague in PTQ and followed by a number of subsequent developments of the theory (such as those in Partee 1976). So according to

this hypothesis infinitives and gerunds will denote the same sort of entities as are denoted by their inflected counterparts in simple subject/argument constructions. Hence, VPs would always denote properties: when they occur in matrix sentences and when they occur in the complement structure of other VPs. Let us call this the VP = P(property) hypothesis. The question to ask, then, is: If this hypothesis is right, what consequences follow from it within Montague's theory of properties and predication?

Disregarding for the moment intensionality (which is irrelevant to the point we want to make), the semantic type of simple properties such as to be home in MG would be $\langle e, t \rangle$, i.e. properties of simple individuals would be modelled as functions from entities into truth values. But consider now the following sentences:

- (3) a. to be home
 b. being home is nice
 c. John

On the VP = P hypothesis, to be home and being home in (3a-b) denote entities of type $\langle e, t \rangle$ i.e. first order properties. However, in (3a-b) they are not attributed to a subject, but rather they are themselves the subject of predication. Notice that inflected VPs cannot occur as subjects:

- (4) *are home is nice

So, in this regard infinitives and gerunds seem to have both a verbal and a nominal nature: they can be used to predicate something of a subject and to be themselves subjects of predicates. One can put forth, therefore, the hypothesis that infinitive and gerund formation is a process that transforms predicative expressions into names or quasi-names. In this sense such a process might be regarded, from a semantic point of view, as a case of nominalization. We will leave this at an intuitive level for the moment; in chapter III we will try to develop this hypothesis more precisely. Now, given a type theoretical approach, the predicate is nice in (3a-b) could not be of type $\langle e, t \rangle$. It would have to be of type $\langle \langle e, t \rangle, t \rangle$. But in (3c) is nice is attributed to an ordinary individual, and so it would have to be of type $\langle e, t \rangle$. Hence, the VP = P hypothesis forces is nice to be associated with two different functions in (3a-b) and (3c). Furthermore, if the relation between syntactic categories and semantic types is a function, then is nice in (3a-b) cannot be of the same category as in (3c). According to this view it is a mere accident that the predicate in (3a-c) looks just the same, since its different occurrences would have to differ both in meaning and syntactic category. But this is quite implausible. So the principle of transparency, the VP = P hypothesis and the theory of types lead us to

counterintuitive results on the nature of certain predicates.

The situation is worse than that, however. Consider, for instance the following sentence:

- (5) being home is extremely nice
- (6) John is extremely nice

If the adverb in (5)-(6) is what it looks like, i.e. a VP-adverb, it will correspond semantically to a function from properties into properties. The function associated with the adverb extremely, then, would have to be of different types in (5) and (6) since it modifies properties that look alike but, as we have just seen, are very different. In (5), the adverb would be a function of type $\langle\langle e, t \rangle, t \rangle$, $\langle\langle e, t \rangle, t \rangle$, and in (6) of type $\langle\langle e, t \rangle, \langle e, t \rangle\rangle$. And again granted a functional dependency of categories-types, the two adverbs in question would have to be of different syntactic categories. The same argument could be iterated for adverbial prepositional phrases (PPs), for adverb-forming prepositions (Ps), and for quite a number of other categories. So the principle of transparency, the VP = P hypothesis and type theory not only force us to an otherwise unmotivated duplication of meanings and syntactic categories for a certain class of predicates; this duplication, in fact, spreads throughout the entire categorial system of the grammar. Most verbs, common nouns, prepositions, etc. have to be split both in

their syntactic and semantic representation into at least an individual (e-type) level and a predicate (<e, t>-type) level.

There are further problems related to the ones just considered. Take is nice in (3a-b). Presumably, it will have a gerundive counterpart, i.e. being nice, for there is no reason to suppose that gerund formation can apply only to e-level predicates. But gerunds can be themselves subjects of other predicates, as in e.g.:

(7) being nice is a quality

So the whole affair starts over again; is a quality in (7) would have to be a VP of level <<<e, t>, t>, t>; and by the reasoning developed above, we can show that this further increase in type-level spreads throughout the categorial system of the grammar. And the argument can be applied over and over again: we can form gerunds like being a quality, which can then be the subject of other predicates, etc. There is no clear upper bound to this process. So, we not only have a duplication in the syntactic and semantic representation of most items. We have in fact a multiplication of them potentially ad infinitum. And each instance of a given item at a given type theoretic level looks just like every other instance of that item (just like the is nice's in (3)). Furthermore, nothing prevents a property from applying to its own gerundive counterpart, as in:

(8) being nice is nice

It is certainly possible to imagine situations where (8) would even be true. But doesn't type theory ban self-predication as meaningless? It does. So, some further explaining is in order if we are to maintain the VP = P hypothesis and a type theoretic notion of property. Now, the arguments given above show that there is probably a potential infinity of is nice's: one of level $\langle e, t \rangle$ one of level $\langle \langle e, t \rangle, t \rangle$, etc. So, it may be that when we utter (8) we really mean something like:

(9) being nice_n is nice_{n+1}, for all n

where n is the number of t's in the type of being nice-functions. (9) has to be taken as something like a schema for an infinite number of type theoretically acceptable predications. Notice, though, that it appears *prima facie* not so easy to connect systematically all the being nice-functions with a plausible system of syntactic categories. But let us grant that a solution to this problem can be found. One should ask, then, the following question: is (9) a plausible representation of the meaning of (8)? First we should check whether there is any empirical evidence suggesting that this is the case. Certainly, there does not seem to be any syntactic evidence. There isn't any overt and systematic correlate of a ranking of predicates according to a type theoretic

hierarchy in the syntax of any natural language I know of. And isn't it funny that such a central semantic distinction would go unmarked in language after language?

Semantic evidence doesn't abound either. For instance, as pointed out in Parsons (1979), many people would consider the following a valid inference:

- (10) Every entity has the property of being
autoidentical
 every entity has a property

This argument is inexpressible in a logic based on type theory. In such a theory there is no way to speak about everything. We can only speak about everything of a given type. The closest we can get to formulating the conclusion of (10) is by doing something like (9) (i.e. everything $thing_n$ has a property $n+1$). But (9) would not be a wff in a typed logic such as Montague's IL; at best it could be a schema in an infinite set of wffs.

What about semantic intuitions? Could we claim that the semantic representation of (8) provided by (9) matches out intuitions? If I have at all intuitions about what sentences like (8) mean, they go in the direction of something like self-instanciation of properties. As far as (9) is concerned, I have trouble in understanding what it should mean even being acquainted with the theory of types.

It appears, then, that representing (8) as (9) might just be a regimentation of natural languages forced upon

us by the principle of transparency, the VP = P hypothesis and the theory of types. Such a regimentation, if we are right, lacks any independent support. It should be noted that on the VP = P hypothesis, sentences like (8) are problematic not only for the theory of types. Suppose we wanted to provide the semantics of English by translating it not into a typed language, such as Montague's IL, but into an untyped one as, say, Zermelo-Fraenkel set theory. Suppose, furthermore, since we are ignoring intensions, that we represent properties in terms of the set of things that bear them (in a specified situation). Then (8) would turn out to mean something like "the set of nice things belongs to itself". But no set can belong to itself in ZF. So there is no simple account of something like (8) even outside of the theory of types. And it should be evident that turning to a weakly intensional representation of properties in terms of possible worlds would not help at all.

To sum up, an analysis of infinitives and gerunds based on the transparency principle, the VP = P hypothesis and the theory of types runs into a host of pretty serious difficulties. Now, it would seem that the problem is independent from Montague's transparency principle. For assume we give it up. Still the meaning (if no longer the syntactic category) of most expressions of a natural language would have to be multiplied ad infinitum. The

central mechanism of type-theoretic predication forces properties that apply to urelements and properties that apply to higher order entities to be distinct. Given the recursive character of gerund formation and the inter-connection among meanings of expressions, this effect cannot be confined to a limited set of cases but invests the entire semantic system. Virtually all the meanings of natural language expressions end up being split and scattered throughout an infinite hierarchy of types. And there is no empirical evidence that this should be the case. Most relevant semantic distinctions have reflexes in the syntax of some natural language or other; the ranking of predicates characteristic of type theory doesn't. Furthermore, even if we weaken Montague's transparency principle, the multiplication of meanings induced by the theory of types is bound to have reflexes in the syntax, if syntax and semantics are related to each other. In fact, in the case of Montague's proposal of a functional dependency between categories and types, this influence appears to be devastating. But on any way of casting the category-type relation, we will have to encode in the syntax a mechanism to keep track of the semantic typing of meanings. We will consider later an instance of such an attempt due to T. Parsons. However, it seems legitimate to suspect that any such mechanism would build into the syntax a device that lacks any independent

syntactic support. Hence, everything else being equal, a theory that can handle gerunds and infinitives without forcing us to do so would appear to be preferable. We can conclude, then, that the principle of transparency is not responsible for the kind of problems we are considering.

It may be, however, that the VP = P hypothesis is the culprit, rather than type theory. For instance, if infinitives and gerunds are associated semantically with propositions the problem, in this devastating form, would not arise. Now, there are various reasons why this proposal has been put forth. In the rest of this work we will consider several (hopefully, most) such reasons and try to argue that they are inconclusive. In fact, we will try to argue that the best possible account of the properties of infinitives and gerunds must be given in terms of the VP = P hypothesis. But then the thing responsible for the problems considered here is just type theory.

But even if the VP = P hypothesis should turn out to be wrong, still I believe the problems pointed out above in connection with nominalization would arise. There are in fact other phenomena that can (and, I think should) be analyzed as processes whereby predicates are somehow turned into (or looked at as) individuals. We will now consider some such cases.

1.2 Other nominalization phenomena.

Plural NPs often give rise to group readings, as in:

(11) Tom and Harry lifted the piano

How are groups to be analyzed? A natural thing to do would be to regard them as sets. In fact, this proposal has been made, among others, by Cresswell (1973) and Bennett (1976), which still constitute the most complete treatment of plurality within MG.³ Now, sets correspond to properties in extension, whose type in Montague's IL is $\langle e, t \rangle$. But if a group is analyzed as an entity of type $\langle e, t \rangle$, a property of a group on a type-theoretical approach must be of type $\langle \langle e, t \rangle, t \rangle$. So, lift the piano will have to belong to two different types when it applies to ordinary individuals and when it applies to groups. We face, therefore, the very same problems as in section 1.1. It should also be noted that in the same way as we speak about groups, we can speak about groups of groups, groups of groups of groups, etc. It might also be that there are groups that belong to themselves. So, again, there is no evident upper limit to this process.

A related problem arises also in connection with bare plurals, i.e. constructions like the following:

(12)		extinct
		mammals
	whales are	numerous
		in short supply
		loved by John

Carlson (1977) has argued that bare plurals are syntactically NPs derived from plural common nouns and behave pretty much like proper names (with respect to anaphora and scope phenomena). To account for this, he proposes to analyze bare plurals as names of kinds. So the formation of bare plurals out of plural common nouns (see Bennett 1976 for an analysis of the latter) would amount to turning a predicate into something like a proper name of a kind. Now, in the philosophical literature it has been proposed, most prominently in Cocciarella (1976), to regard kinds as (nominalized) properties.⁴ If this proposal is correct, then we would have the following situation. Consider:

(13) Moby Dick and Moby Dick Jr. are whales

In (13) whales is used as a predicative expression in order to attribute to the Moby Dick's the property of belonging to a certain kind. In (12) whales is used as a proper name of the kind and is the subject of a predication act. Now, if we stick to a type theoretic notion of property, the CN whales in (13) will have to denote something of type $\langle e, t \rangle$; if such a CN is then turned into an NP with singular reference, the properties which are predicated of it (as in (12)) will have to be of a higher type. This will get us into the by now familiar kind of trouble.

Constructions that behave much like bare plurals are those involving mass nouns; consider:

- (14) a. This ring is gold
 b. Gold is an element

(14a) attributes to a ring the property of being made of a certain substance. So in (14a) the CN gold is used as a property-denoting item; in (14b) on the other hand gold seems to be used as a referring NP (maybe the name of a substance) of which we say that it is an element. We run again into the same pattern: a predicative expression can also act somewhat like a name. If gold in (14a) is a genuine predicate, its type will be $\langle e, t \rangle$. But what is the entity that the NP gold purports to refer to in (14b)? Ter Meulen (1980) proposes to analyze mass NPs as names of the property associated with the corresponding CN (i.e. as nominalized properties). On a type theoretic approach, this will entail that predicates that apply to mass NPs will have to be of a higher type than predicates that apply to ordinary individuals. But this again leaves open the puzzle of how to deal with predicates that apply indifferently to ordinary individuals and to substances, as e.g.:

- (15) I love gold
 touched John

By the usual argument, the verb love would have to be of a different type when it applies to John or to gold. And

this would have the usual effects on the overall organization of the grammar.

Similar problems might be argued to arise in connection with nominalization of propositions. Consider pairs like the following:

- (16) a. John is a little strange
 b. That John is here is a little strange
 c. New York is a little strange

On the assumption that ordinary individuals like John and propositions are of a different type, then to be a little strange must have different meanings in (16a) and (16b).

Note, however, that if cities are basic entities (like humans and unlike propositions) then to be a little strange would be the same function in (16a) and (16c).

What motivates drawing the line exactly where type theory does?

Other cases that look like cases of transformation of predicative expressions into names involve adjectives.

Consider the following examples:

- (17) a. This dress is red
 b. Red is my favourite colour

- (18) a. John is good
 b. Goodness is disappearing from this world

In (17a) the adjective red is used as a predicate; in (17b) as a name of an entity (a colour). In (18a) the property of being good is attributed to John; in (18b) the adjective good is turned into an NP, presumably denoting something like a quality. It is not implausible that

these processes might be analyzed in a similar way to the one sketched above in connection with mass and plural nominals. If this hypothesis is correct, their analysis would turn out to be problematic for a type theoretical approach in just the same way the cases considered above do.

Even in this very sketchy form, what we have said should be sufficient to show that problems analogous to those pointed out in connection with infinitives and gerunds arise (or are likely to arise) also in connection with other grammatical phenomena, given some reasonable assumptions. This gives us further grounds for believing that the difficulties in question are not an artifact of the VP = P hypothesis. Rather they seem to arise as a consequence of the idea that properties should not be treated on a par with their arguments, which is the central tenet of a type theoretic approach to predication.

1.3 Conclusions.

There are a number of grammatical phenomena whereby predicative expressions of English (VPs, CNs, ADJs, etc.) are turned into noun-like items, i.e. items that purport to have singular references like proper nouns do. In a type theoretic semantics properties have to be ranked on the basis of the "level" of their arguments (i.e. according to whether they apply to urelements, properties of

urelements, properties of properties of urelements, etc.). So, properties of basic elements and properties of (nominalized) properties will have to be different entities. This seems counterintuitive. Furthermore it has undesirable effects on the overall organization of a grammar (both in the semantics and in the syntax), which are totally unsupported from an empirical point of view. We are led to conclude that type theory provides an unsatisfactory basis for a linguistically adequate notion of property and predication. If we can provide a notion of property that allows us to do anything we can do in standard MG and can also deal with nominalizations without running into difficulties such as those we have considered, it would seem to provide a better basis for natural language semantics.

2. On Weak Intensionality.

One of the most serious problems for Montague semantics concerns the treatment of propositional attitudes and related phenomena. If believing is a relation between individuals and propositions and if logically equivalent propositions are identical (as they are bound to be on a possible world approach), then we are committed to the validity of many arguments that do not seem to be valid. So, such a semantics fails in providing an adequate account of the notion of logical consequence, once propo-

sitional attitudes are brought into consideration. Since it is possible to regard propositions as 0-place properties, the problem of propositional attitudes can be viewed as a consequence of the identification of logically equivalent properties, i.e. the adoption of weak intensionality in general. There seems to be, in fact, growing consensus on the idea that one of the sources of trouble (maybe the source of trouble) is the attempt to handle propositional attitudes by employing an insufficiently intensional notion of property (e.g. see Bealer 1982, Barwise and Perry 1981, Thomason 1980).

Now, suppose we had a theory equivalent in empirical coverage to Montague semantics, but not committed, unlike the latter, to weak intensionality. Such a theory would be a priori more appealing, since even though it might not solve automatically the propositional attitudes problem, it would certainly provide a framework within which such a problem might be tackled with greater hope of success. The theory we are going to propose in the next section is not committed to weak intensionality, though, as we will see, it is consistent with it. So, within such a theory, you don't have to identify properties with functions from possible worlds into sets, but you can if you want to.

Within the limits of the present work it is impossible to go in any depth into the propositional attitudes problem. What I would like to do, instead, is to point

out some more specific problems for linguistic theory that arise if one adopts a weakly intensional notion of property. We would like to individuate areas in the grammar where choosing between a weakly and a strongly intensional theory of properties has clearly noticeable and interesting empirical consequences. The structure of the arguments we will consider is the following. Suppose we have two competing linguistic analyses of the same range of facts. It can turn out to be the case that while strong intensionality is compatible with both, weak intensionality is compatible only with one. Through the lens of a weakly intensional semantics, we just cannot see one of the two analyses. If the analysis compatible with weak intensionality is the "best one" on purely internal linguistic grounds (i.e. according to the evaluation metric set up by universal grammar), this would strongly support weak intensionality. If the contrary is the case, we would have instead disconfirming evidence for such a thesis. The cases we will consider all turn out to be problematic in this sense for weak intensionality, at least according to some current proposals.

2.1. Conjunction.

In a number of recent works (Gazdar 1980, Keenan and Faltz 1978, Cooper 1983, Rooth and Partee 1983) it has been argued that virtually all uniform syntactic consti-

tuents can be directly conjoined. So, in particular, gerunds can. On the VP = P hypothesis, all the analyses of generalized conjunction mentioned above would assign to conjoined gerunds such as, say walking and talking, (at least on one reading) the property of being an x such that x walks and x talks. Let us denote such a property by walking' \cap talking'. Take now a verb phrase denoting a property that everything has necessarily, such as, say, the property of being self identical. On a weakly intentional notion of property the result of conjoining any property P with the universal property would have to be P itself. So, for instance, the result of conjoining the property of, say, jogging with the property of being selfidentical is the same as the property of jogging; i.e. jogging' \cap being selfidentical' = jogging'. But on the analysis of conjunction sketched above, this predicts that there is a reading of (19) which is synonymous with (20):

- (19) John likes jogging
 (20) John likes jogging and being selfidentical

The reading of (19) which is equivalent to (20) would be the one which is not gotten by (an analogon of) a "conjunction reduction" rule (see Rooth and Partee 1983, for discussion). Such a reading would be represented as:

- (21) like' (jogging' \cap being self identical')(John')

However there does not seem to be any sense in which (20) can be taken to be even equivalent to (21). Examples of

this kind (possibly also more convincing ones than the one I presented) can be generated ad libitum. So something has gone wrong either with the VP = P hypothesis, or with the analysis of conjunction, or else with weak intensionality. To the extent that the former two appear to be sufficiently motivated from a linguistic point of view, the source of the problem would appear to be located in the third hypothesis.

2.2 Passive.

In the tradition of generative grammar, discussions concerning the analysis of passive have played a crucial role in the development of various frameworks. Within MG, the analysis of passive in general most adopted is in two stages; one passive rule takes care of agentless passive and one of passive with agent. Briefly, here is how the analysis goes. The agentless passive rule can be viewed as a process that maps transitive verb phrases into intransitive ones (by applying to them passive morphology); the semantic side of the rule would amount to an existential quantification over subject position. A purely illustrative version of the agentless passive rule, together with an example is given in what follows:

(22) Agentless passive (AP): if $\alpha \in TVP$, then $PAS(\alpha) \in IV[+PAS]$, where $PAS(\alpha) = be\alpha ed$
 $PAS(\alpha)' = \lambda x \exists y \alpha'(x)(y)$.

(23) kick \rightarrow be kicked
 $be\ kicked' = \lambda x \exists y [kick'(x)(y)]$

For passive with agent there is a rule that syntactically combines transitive verbs with by-phrases, yielding a passive VP; semantically such a rule inserts the NP meaning of the argument of by in the 'subject' slot. This is illustrated in what follows:

(24) Passive with agent (PWA): if $\alpha \in \text{TVP}$ and $\beta \in \text{NP}$,
 then $\text{PAS}(\alpha)$ by $\beta \in \text{IV}[+\text{PAS}]$
 If $\alpha \in \text{TVP}$ and $\beta \in \text{NP}$, then $\text{PAS}(\alpha)$ by β
 translates as $\lambda x[\alpha'(x)(\beta')]$.

(25) kick, John \Rightarrow kicked by John
 kicked by John' = $\lambda x \text{kick}'(x)(\text{John}')$

A detailed discussion and motivation for this analysis can be found in Bach (1980a). Against it there are a number of arguments that might be raised. For instance, on the basis of (22) and (24) a grammar of English would not look more complicated if the passive morphology brought about by AP and PWA looked totally different. The evident identity between the two would appear to be accidental. Furthermore, by-phrases can attach also to NPs (as in a book by John, etc.), to which they seem to add an agentive argument, just like they do with passive VPs. It is not clear how this evident relatedness between the two constructions might be captured on an analysis based on (22) and (24). These simple minded linguistic objections to a two-rules analysis of passive, however, are not conclusive and can be countered in various ways (see the discussion in Bach (1980)). Nevertheless several authors have pointed out undesirable empirical consequences of a

two-rule analysis (Bresnan (1982), Halvorsen (1982), Abe (forthcoming)) and argued in favour of a one rule analysis. Now, within the framework we are considering a one rule analysis could be developed along the following lines. It seems plausible to assume that by-PPs are just predicate modifiers (which on independent grounds we know can attach to both CNs and VPs) that in some sense add an agent to both VPs and NPs. The relation between passive VPs and their active counterpart might be captured by resorting simply to a meaning postulate like the following:

$$(26) \text{by}'(x)(P)(y) \leftrightarrow \exists R [P = \hat{\lambda}x \exists y [\check{R}(x)(y)] \& \check{R}(y)(x)]$$

where P and R are variables of type <s, <e, t>> and <s, <e, <e, t>>> respectively

It should be noted that the meaning postulate in (26) is limited to VPs. However, the present approach could be generalized, in principle, to an analysis of the by-PP modifiers of CNs, roughly as follows:

$$(27) \text{by}^+(x)(\text{CN}') (y) \leftrightarrow [\text{CN}'(y) \& \text{by}'(x)(\text{brought about}') (y)]$$

where by' is as in (26)

So agentive modification of CNs could be defined in terms of agentive modification of VPs, which is a plausible first approximation towards capturing the relatedness between the two.

An analysis along the preceding lines was in fact considered in the early times of MG (see e.g. the remarks on this in Dowty 1978). It had, though, to be abandoned,

for the following reasons (pointed out originally by R. Thomason). There are in natural languages items related as follows:

- (28) a. $\Box [\text{buy}'(x)(y)(z) \leftrightarrow \text{sell}'(z)(y)(x)]$
 b. $\Box [\text{lend}'(x)(y)(z) \leftrightarrow \text{borrow}'(z)(y)(x)]$

I am assuming that the PPs subcategorized for by these verbs are in fact internal arguments; i.e. we analyze the verbs in question as 3-place predicates (ditransitives). To assume that a PP is an internal argument of a verb amounts to claiming that such a PP has an NP meaning and its preposition acts somewhat as an overt case marker. The present assumption is made only for sake of simplicity, and the following argument would go through even if from- and to- PPs in (28) were predicate modifiers. The oblique argument of buy and sell can be deleted by a very general process ('argument drop') which in the case at hand would map ditransitive verbs into transitive ones and existentially quantify over the position of the argument that is deleted.⁵ The result of applying argument drop to buy and sell will give:

- (29) $\text{buy}'_{TV} = \lambda x \lambda y \exists z [\text{buy}'_{TTV}(z)(x)(y)]$
 $\text{sell}'_{TV} = \lambda x \exists y \exists z [\text{sell}'_{TTV}(z)(x)(y)]$

One can then apply passive (i.e. AP) to buy_{TV} and sell_{TV}, thereby obtaining:

- (30) a. be bought' = $\lambda x \exists y \exists z [\text{buy}'_{TTV}(z)(x)(y)]$
 b. be sold' = $\lambda x \exists y \exists z [\text{sell}'_{TTV}(z)(x)(y)]$

An agentive by-phrase (say, by John) could then apply to (30a) and (30b). However, in virtue of (28), be sold' and be bought' turn out to have the same intension; hence by weak intensionality, they have to be the very same entity (the very same function from worlds into sets). This gives us the following results:

- (31) a. $\hat{\text{be sold}}' = \hat{\text{be bought}}'$
 b. $\text{by}' (\text{John}') (\hat{\text{be bought}}') = \text{by}' (\text{John}') (\hat{\text{be sold}}')$

So the property of being bought by John and being sold by John should be the very same property. But this is absurd. Something has gone wrong either with the one-rule analysis of passive, or with weak intensionality. If we want to stick to weak intensionality, we are forced to adopt a two rule-analysis of passive. This seems to provide us with a rather clear case where a semantic theory constrains syntactic analysis, and heavily so. Does it constrain syntax in the right direction?

By contrast, if we adopt a strongly intensional property-theory, it is easy to see that we could maintain a one rule-analysis of passive, for according to such a theory two properties may be distinct even if they have the same extension at all worlds. So be bought' and be sold' could be equintensional but distinct properties, and thus by-PPs can yield different outputs when applied to them. To the extent to which a one-rule approach would appear to be better motivated on purely internal linguis-

tic grounds, the thesis of weak intensionality would turn out to be disconfirmed.

2.2 Predicate modifiers.

It might be argued that the case made for passive can be extended to the analysis of predicate modifiers in general. Predicate modifiers in standard MG are grouped into two main categories, that of verb modifiers (IV/IV), which contains mainly adverbs and adverbial PPs, and that of common nouns modifiers (CN/CN), which contains adjectives and PPs. Semantically, predicate modifiers are analyzed uniformly as functions that map properties into new properties. It follows, then, that an adverbial applied to two different VPs that denote the same property should yield the same output. But take again pairs like be bought and be sold or be lent and be borrowed, and consider the following examples:

- (32) a. This book was bought in a hurry
 b. This book was sold in a hurry
- (33) a. This book was borrowed with interest
 b. This book was lent with interest

If the PPs in (32)-(33) are IV/IVs, these examples are problematic for weak intensionality, for the a-variant of each pair would have to be synonymous with the b-variant, while they don't even have the same truth conditions.

However, the little that is known about the semantics of adverbs is not sufficient to grant the preceding

conclusions. Even though the issues involved are extremely complicated, it is worth considering them in some more detail, for I believe that the area of adverbial modification might be a serious source of problems for weak intensionality. What we can do within the limits of the present work is to sketch a couple of current approaches to the semantics of adverbs and to point out how they would inherit the be sold-be bought problem.

One possibility put forth by Keenan (1980), among others, would be to analyze the adverbs in (32)-(33) as TVP-modifiers; passive would have then to apply to the already modified TVP (e.g. buy in a hurry). This move could be motivated on the basis of the fact that we need to find a systematic way of relating adverbs with the arguments of the verbs they modify. Consider the following valid entailments:

(34) John kicked Bill at the party \Rightarrow Bill was at the party
 killed
 washed $\not\Rightarrow$ John was at the party

(35) John saw Bill from the roof \Rightarrow John was on the roof
 signalled
 attacked $\not\Rightarrow$ Bill was on the roof

So, locative and source adverbials seem to differ in their entailment properties; locatives seem to apply to objects and source-adverbials to subjects. On the basis of these data, Keenan seems to suggest that adverbials that "apply to" objects should, in general, be regarded as TVP-modi-

fiers and adverbials that "apply to" subjects should be regarded as IV-modifiers.⁶ But then by this criterion the adverbials in (32) and (33) should be regarded as IV-modifiers (cf. John bought this book in a hurry) and the be sold-be bought problem would arise again.

Halvorsen (1982) has made in this regard an interesting proposal. His idea is that adverbs might be regarded uniformly in the semantics as being property-modifiers. The difference between subject oriented and object oriented adverbs could then be accounted for by assuming that adverbs are somehow sensitive to thematic relations (cf. for a related proposal, Jackendoff (1972), among others). For instance, one might try to say, in first approximation, that locative adverbials (such as at the party) require themes to be at the relevant location, while source-adverbials would require sources to be at the relevant location. This would account for the data in (34)-(35). In general, then, the format of entailments licensed by adverbials would look like the following:

$$(36) \gamma [(\beta)(\alpha_1) \dots (\theta_i) \dots] (\alpha_n) \rightarrow \gamma' (\alpha_i)$$

where γ' is the 'predicative' meaning associated with γ
 β is a VP-meaning
 and θ ranges over thematic relations.

What (36) states is that an adverbial γ "applies" to the argument of β which bears to it the relation θ . Of course, this analysis relies crucially on θ -relations,

notions regarded as murky by many. We will discuss at some length the status of Θ -relations within model theoretic semantics in chapter IV. At any rate, it is quite evident that if Halvorsen's proposal can be worked out, then the uniform analysis of adverbials as VP-modifiers would give rise to some version of the be bought-be sold-problem on a weakly intensional theory of properties.

One of the most interesting recent discussions of the logic of adverbial modification can be found in McConnell-Ginet (1982). She argues that a bipartition of adverbs between IVP and TVP modifiers (aside from S-adverbs) is insufficient and argues for a different way of classifying predicate modifiers. According to her, the latter group divides up into what she proposes to call Ad-verbs and IV/IVs. Ad-verbs are basic entities which should not be regarded as functions on predicates but rather as internal arguments of predicates. To accommodate this view, she formulates the notion of 'admissible augmentation' of an n-place predicate. An admissible augmentation of an n-place predicate R is an n + 1 predicate $R^+ = R \cup S$, where $S \subset R \times Y$ and Y is the type of the Ad-verb (i.e. a manner, a rate, a place, etc.). For example, a given Ad-verb such as quickly can modify a verb, say run, only if there is an admissible augmentation of run which has a rate as one of its arguments. IV/IV-

modifiers are then derived from Ad-verbs through the following schema:

- (37) If ξ is an Ad-verb, ξ' is an IV/IV, and for any β in IV, $\xi'(\beta) = \text{act } \xi \text{ to } \beta$

On the basis of this hypothesis, McConnell-Ginet is able to provide a plausible explanation for a quite intricate set of facts in the behavior of adverbs, including, e.g., the ambiguity in:

- (38) a. Lisa rudely departed
b. Lisa departed rudely

(38a) can mean that given the circumstances it was rude that Lisa departed (independently of the manner of Lisa's departure); in (38b), however, the 'locus' of the rudeness must be in the way Lisa's departure took place. Rudely, it should be noted, according to standard tests could not be regarded as an S-adverb (see on this Thomason and Stalnaker 1973); given that depart is intransitive one cannot blame the ambiguity in (38) on a IVP- TVP-modifiers distinction either. McConnell-Ginet argues that in (38b), rudely is an Ad-verb in her sense, while in (38a) it is an IV/IV-modifier. She also applies her theory to passive sensitive adverbs, such as:

- (39) Fido was sold reluctantly (by John)

In this example, the pleasure in the selling might belong either to the patient or to the agent. She accounts for this by claiming that the agent sensitive reading within her system must be the one gotten through the Ad-verb

reluctantly, while the patient sensitive would be obtained through the corresponding IV/IV-modifier (McConnell-Ginet 1982, pp. 172-4). It is evident, then, that on a weakly intensional notion of property and on the IV/IV variant of reluctantly, (39) should be synonymous with:

(40) Fido was bought reluctantly

Are they synonymous? It is not clear. The suggested paraphrase (Fido acted reluctantly to be sold/be bought) is dubiously grammatical which makes it hard to decide. It would seem to me that a more plausible interpretation of (39) and (40), on the subject oriented reading of the adverb, would be something like "Fido was sold/bought but he didn't want to". Now it seems to me that one might want to be sold without wanting to be bought (just like one might have 'contradictory' intentions). But then (39) and (40) would not be synonymous on the intended reading, which, on the basis of weak intensionality would be a puzzle. Thus, also on the otherwise quite convincing semantic theory of adverbs developed by McConnell-Ginet, the be sold-be bought problem might arise.

It is also likely that some version of the problem under consideration will arise in the case of CN-modification. For instance, Partee (1983) points out that on a weakly intensional theory of properties, a uniform analysis of of-PPs as CN-modifiers is impossible. One is

forced to assume that there are "transitive" CNs (of category, say, CN/+of PP). Whether this move is desirable (or necessary) for the analysis of genitive constructions in English is an open question. But having a weakly intensional theory of properties prejudices the question independently of empirical considerations of generality and the like.

Thus analyzing predicate modifiers as functions from properties into properties, coupled with a weakly intensional notion of property, leads to a series of (at least potentially) undesirable empirical consequences.

2.4 Summary.

In the present section we have illustrated some problems for the thesis of weak intensionality that might arise in a theory of grammar independently of the general philosophical problem of propositional attitudes. We considered briefly various proposals that have been made in connection with conjunction, passive, and predicate modification. In the analysis of these phenomena, empirical considerations of generality, simplicity, etc. might lead us to select hypotheses that potentially clash with a general semantic theory based on weak intensionality. Another way to put it is that a weakly intensional notion of property clearly makes a big difference in the empirical analysis of the data in question. In many of the

above examples we seem to be faced with the alternative of either tinkering somehow with the analysis or rejecting weak intensionality. To the extent that such an analysis-tinkering would appear to be otherwise unmotivated, it might be more plausible to abandon weak intensionality. This is so, especially in consideration of the likely eventuality that such a move will be necessary anyhow to deal properly with propositional attitudes, as several philosophers have argued.

3. A Second Order Theory of Properties.

We will now present a theory of properties which (1) is not based on the theory of types, (2) allows for a very general treatment of nominalization, and (3) is not committed to weak intensionality. Such a theory is due to Nino B. Cocchiarella.⁷ Cocchiarella's theory takes the notion of property as basic and provides an axiomatic characterization of how properties behave and what structural features predication has. Property and predication are thus taken as primitive notions in pretty much the same way in which set and set-membership are within set-theory. Linguists in the Montague tradition are accustomed to taking individuals and possible worlds as basic and to defining properties in terms of those. *Prima facie*, though, it does not seem to be the case that the notions of individual and possible world have in any sense

a "clearer" status than the notion of property. Our intuitive understanding of what an individual is does not provide us with a procedure for deciding in any given circumstance whether two individuals (or rather two individual-manifestations) are the same. Identifying individuals in different circumstances is a notoriously problematic question. The situation with possible worlds is, if anything, worse. Similarly, we certainly have an intuitive understanding of what a property is, even though we do not have ready made identification criteria for properties. So, adding properties to our stock of primitives does not seem to introduce in our ontology creatures of greater darkness than those we already have, especially since we will carry out this move by means of a theory provably consistent and complete (if ZF is). At any rate, it is not our job to provide a philosophical justification for the notion of property. What we want to do is show that by resorting to a theory of properties such as Cocchiarella's we can solve the problems (concerning nominalization and intensionality) considered above in one fell swoop.

Even if we cannot discuss in any detail the philosophical assumptions regarding the background of a theory of properties, there are a few general consequence that adopting such a theory has on linguistic semantics that are worth discussing right away. Given the limits of the

present work, we will not be able to give to the following issues the consideration they deserve. We wish, nevertheless, to point them out at least in a preliminary fashion, for they are directly relevant for the present enterprise.

3.1 Model theoretic vs axiomatic theories of properties.

Montague's theory of properties is cast in model theoretic terms. The existence of 3 basic sets is assumed: the set E of individuals, the set t of truth values and the set W of possible worlds. Then models of natural languages are defined using set-theoretic devices. Formally, such models are characterized by means of definitional extensions of some standard version of (Von Neumann-Gödel) set theory. Properties, in particular, are identified as certain objects in the model (roughly, functions from W into sets). Questions such as whether a property P exists, or whether P can be, say, conjoined with Q can be answered in terms of the definition of the property in question out of W , t and E and, ultimately, in terms of the logic of set-theory (i.e. the logic of ϵ). The resulting notion of property is typed in two senses. First, Montague defines the relevant set of models in order to provide an interpretation for IL, i.e. a typed language into which natural languages can be translated. Second, the underlying set theory is a

standard one, where \in -chains (i.e. chains of the form $x \in \dots \in x$) are disallowed (cf. the axiom of foundation).

Given that properties are sets built out of E, t and W, and given Montague's definition of function application, it follows that no property can truly apply to itself.

Let us now consider some general metatheoretical consequences that the adoption of Montague's theory has for linguistics and compare them with those that an axiomatic theory would have.

First, the models defined by Montague are standard in the sense of Henkin (1950). What this means, intuitively, is that, e.g., all the functions in, say, t^E are taken as possible property-extensions. If we use standard models to characterize a notion of logical consequence with respect to some higher order language L, it is well known that no finitely specifiable axiomatization for such notions exists (see the discussion in Gallin 1975). So in Montague semantics, not all arguments (of the form ψ , hence ϕ) that are valid could be possibly recognized as such by finite machines. This state of affairs might be argued to be not ideal, for the following reasons. Linguistics is not concerned with arguments that are valid simpliciter. This may be the concern of logic. What is in question is the validity of arguments on the basis of our implicit knowledge of English grammar. Let us call such a notion of validity E-validity (E for English).

Most linguists think that a grammar should represent in an idealized form the implicit knowledge of a native speaker of a language. So, syntax should characterize our ability to recognize a sentence as wellformed, semantics our ability to recognize a sentence as E-valid. It is commonly held, on the line of Chomsky (1957), that "human abilities to recognize" should be represented in the form of some algorithmic device (let us say, for explicitness, a Turing machine). This is not a very strong requirement. Among logically conceivable algorithmic devices there are many that could not possibly be (part of) natural language grammars. But still it seems to be a plausible requirement. I will not dwell on that. Now, the that Montague semantics is meant to be a characterization of E-validity, but there is no algorithmic device that corresponds exactly to such a semantics, goes against Chomsky's claim that an algorithm is what in principle is needed to represent human capabilities to make yes-no decisions. To the extent that such a claim is well motivated, a non axiomatizable characterization of E-validity might be regarded as a problem. In other words, the same reasoning which induces us to want syntax to be algorithmic should induce us to want our semantics to be axiomatizable (i.e. in some sense algorithmic too). To move towards such a goal one thing that we need is an axiomatic characterization of what a property is.

Montague semantics as it is cannot be provided with such a characterization.

A second feature of Montague's theory of properties seems to be its commitment to realism. If properties are sets of ordered pairs of the form $\langle w, e \rangle$ (where w is in W and e is a property-extension), then they cannot be mental constructs because, intuitively speaking, they are too big.⁸ It may be, however, that properties really are not mental constructs, and that anyway to decide whether they are or are not is not the linguist's job. But this seems to be precisely an argument in favour of a theory that, unlike Montague semantics, is not committed to the position that properties could not in principle be, say, mental constructs or logico-grammatical roles.

If, on the other hand, we have an axiomatic characterization of what a property is, then it would seem that we are not necessarily forced to a realistic position. This will depend on the theory itself. For example, we might still maintain that there is an infinite number of properties. But such an infinity might be only potential: i.e. an extrapolation from our capacity to apply iteratively basic logical operations. We know that we can form new properties out of given ones by using \exists , \rightarrow and \neg in formulae of indefinite complexity, and so the set of properties must be infinite. But we don't have to assume that all properties are sitting in an independently given

infinite set (say, t^E) which our linguistic capacity is able to scan in its entirety. In other words, not all members of t^E need to be possible property-extensions, but only some subset of it satisfying certain conditions of closure with respect to basic logico-linguistic operations. Shifting from t^E to some subset of it amounts to going from standard models to general ones (again in the sense of Henkin, 1950), which will allow us to establish completeness for our theory of properties.

A further point which is worth touching upon briefly concerns the ontological commitments of our theory. If we cast a theory of properties solely in model theoretic terms, and if we furthermore assume that properties should be related to (or maybe defined in terms of) constructions out of possible worlds (and, say instants), then we would seem to be committed to the actual existence of possible worlds (and durationless temporal atoms). Otherwise, what are we talking about when we say, for instance, that a property P is true of u in the possible world w ? Doesn't the variable w of our theory refer, just as u does? And if so, to what? Suppose, instead, we have an axiomatic theory of properties. Then we can say that properties are whatever behaves in the way characterized by our theory. Constructions in terms of possible worlds may be useful model theoretic devices to test certain features of our theory, i.e. to characterize in an independent way reason-

ing patterns that the theory characterizes axiomatically. But as far as our theory is concerned, possible worlds would not have any independent reality, besides that of a useful artifact of model theory. So an axiomatic theory of properties and in general an axiomatic ("proof theoretic") semantics can help us to isolate a level of ontological commitment which makes all and only the assumptions to which the grammar of English forces us. We might want to claim for philosophical purposes that there actually are possible worlds. But what is the evidence for claiming that such an assumption is necessary to build a grammar of English? Couldn't we build an adequate and explanatory grammar without committing ourselves to the existence of possible worlds? If we have an axiomatic characterization of what properties are it would seem clear that we can answer the latter question positively.

These considerations are certainly preliminary and superficial given the complexity of the topics they concern. But even in this form, they seem to show that there are a number of metatheoretical reasons for having an axiomatic characterization of what properties are, rather than having solely a model theoretic one.

3.2 HST*: a second order counterpart of type theory.

In what follows we are going to present the syntax and semantics of a formal system (Cocchiarella's HST*);

such a system can be regarded as a non standard second order logic, which provides a representation of the theory of (homogeneous) simple types. Later we will present extensions of HST* containing λ 's and modal operators. On the basis of these extensions we will present a new semantic theory for (a fragment of) English.

The following introduction is essentially intended as an informal presentation to put linguists in a condition to evaluate the system as a candidate for doing natural language semantics. We will not be able to consider all the technical aspects and the philosophical issues that HST* as a general theory of properties raises; for this the reader is referred to the literature mentioned in fn 7.

The syntax of HST* will have individual variables $\alpha_1, \dots, \alpha_n, \dots$ and n-place predicate variables, β^n, γ^n, \dots . The primitive logical symbols of the language are taken to be $\neg, \rightarrow, \forall$. Other connectives and the existential quantifier will be defined as usual. The major difference between HST* and standard second order logic consists in the fact that the formation rules of HST* allow for predicate variables to occur in argument position (i.e. allow for formulae like $\beta(\gamma)$ or even $\beta(\beta)$ to be well-formed). So, if we let μ_0, \dots, μ_n range ambiguously over predicate and individual expressions, then the formation rule for atomic well-formed formulae (wffs) of

standard second order logic is replaced in HST* by the following clause:

- (41) If β^n is an n-place predicate expression, then $\beta(\mu_0, \dots, \mu_n)$ is a wff.

Argument position occurrences of predicative expressions are said to be nominalized predicative expressions.

Alternatively, as suggested in Cocchiarella (1974) we can add to the apparatus of second order logic a name building device, say ' $\hat{}$ ', that transforms predicative expressions into singular terms. Accordingly instead of (41), we can have:

- (42) a. If β is an n-place predicative expression, $\hat{\beta}$ is a singular term.
 b. If β is an n-place predicative expression and $\alpha_1, \dots, \alpha_n$ are singular terms, then $\beta(\alpha_1, \dots, \alpha_n)$ is a wff.

The other formation rules of HST* are standard. So, this logic distinguishes two roles of predicative expressions. Their primary role is the usual one: attributing properties to subjects in predication acts. Their secondary role is that of being subjects of acts of predication. To allow for the latter, the logical grammar of HST* accommodates a nominalization process (in the form of (41) or (42)). Nominalized predicative expressions are singular terms derived from predicative expressions. This is intuitively close to what seems to take place at the surface of English syntax, where we have, for example,

fully inflected VPs (like runs) that can only occur as predicates, and gerunds and infinitives which can only occur as arguments (i.e. there are no matrix sentences of the form *John to run). Predicative expressions of HST* might be intuitively understood as inflected VPs, nominalized predicative expressions as infinitives or gerunds. The role of 'λ' in HST* would therefore be analogous to that of infinitive and gerund formation. Another example would be the relation between adjectives such as good and their nominal derivative goodness. Good is a predicate, goodness a nominalized counterpart of the predicate. In this case the role of 'λ' would be similar to that of the derivational affix '-ness'. To take a further example, one might consider the relation between sentences (such as John eats a sandwich) and their nominal or quasi nominal counterparts that John eats a sandwich, John's eating the sandwich, and for John to eat the sandwich. In HST* propositions are regarded as 0-place properties and wffs can be, accordingly, regarded as 0-place predicative expressions. Hence they can be nominalized by prefixing 'λ'.⁹ In the case at hand, 'λ' would play a role similar to that of the complementizers that and for or the POSS-ing construction. These examples are only meant to be suggestive of what appears prima facie to be a natural interpretation of the name building device 'λ' of HST*. But it is precisely this very general nominalization

schema built into the logic of HST* that we will try to exploit in providing an analysis for the phenomena just mentioned.

So, the main feature of HST* is that property denoting expressions can be nominalized, and hence properties can be turned into legitimate value of individual variables. But since Russell's paradox of predication, we know that it is impossible to let any old property we can imagine be the value of individual variables, at the risk of inconsistency. A type theoretic regimentation of logical grammar was Russell's own way out of the paradox of predication, a way that Montague built into his theory of properties. What will be our way out? The characteristic axiom of second order logic is the comprehension principle, i.e.:

$$(43) \text{ (CP) } \exists P \forall x_1, \dots, \forall x_n [P(x_1, \dots, x_n) \leftrightarrow \varphi]$$

where P does not occur free in φ

What (43) does is to sanction formally the intuitive requirement that any open wff should "correspond to" a property. By doing this we automatically guarantee that properties may be conjoined, disjoined, etc. (for any operator in the language). Suppose now that in HST* we extend (43) to the following generalized CP:

$$(44) \exists \beta \forall \mu_1, \dots, \forall \mu_n (\beta(\mu_1, \dots, \mu_n) \leftrightarrow \varphi)$$

The following would then be an instance of (44) and therefore provable:

$$(45) \exists R \forall P \forall x [R(\wedge P, x) \leftrightarrow P(x)]$$

(45) posits the existence of a relation which correspond to predication; let us call it Pred. Pred can then be used in further instances of CP such as the following:

$$(46) \exists P \forall x [P(x) \leftrightarrow \neg \text{Pred}(x, x)]$$

According to (46) there is a property that holds of any property that does not hold of itself. By Russell's argument then, such a property will both hold and not hold of itself. Clearly, thus, once we switch from standard second order logic to a system which allows for nominalized predicates we must adopt a restricted version of CP. There are various ways of restricting it, which give rise to various logics of nominalization (Cocchiarella's *-systems). We will be interested in a technique for restricting CP which exploits the same idea that underlies type theory. Such a technique is called stratification.¹⁰ We say that a wff φ of HST* is homogeneously stratified (or h-stratified) if there is a function g from the predicative and individual expressions in φ into natural numbers such that (i) for any predicative expression β , $g(\beta) = g(\wedge \beta)$, and for any atomic sub-wff $\beta(\mu_1, \dots, \mu_n)$ of φ (ii) $g(\mu_i) = g(\mu_j)$, where $1 \leq i, j \leq n$, and (iii) $g(\beta) = g(\mu_i) + 1$. For instance, no formula which contains subwffs of the form $P(\wedge P)$ or $R(\wedge R, x)$ will be h-stratified, because there is no function g from the set $\{P, \wedge P, R, \wedge R, x\}$ into the set of natural

numbers that satisfies (i)-(iii). Also a wff that contains two atomic subwffs like $P(x)$ and $R(\wedge P, x)$ will not be h-stratified, as the reader can easily verify. Intuitively, it is clear that an h-stratified formula is in some sense type-theoretically "safe", for predicates in an h-stratified formula must be exactly one stratum above their arguments. We can then use stratification to restrict CP. Only properties that can be specified in terms of h-stratified wffs are guaranteed to exist by CP. The following will be the CP of HST*:

$$(47) \exists \beta \forall \mu_1, \dots, \forall \mu_n \beta(\mu_1, \dots, \mu_n) \leftrightarrow \varphi \quad (\text{HSCP*})$$

where (i) β is not free in φ and (ii) the entire biconditional is h-stratified

It can easily be seen that we will not be able to posit a predication relation via HSCP*, since (45) above is not h-stratified, and so the existence of Russell's property will be disprovable in HST*. It is in view of the stratification requirement in HSCP* that HST* can be regarded as a second order counterpart of the theory of (homogeneous) simple types.

In this way we will have a theory whose syntax is untyped (apart from the distinction between predicative and individual terms), but an analogue of a typing requirement is used in the metalanguage to restrict CP. The reader at this point will wonder what we are gaining by relocating types from the object language into the metalanguage. Consider the following formula:

$$(48) \exists P \forall x [P(x) \leftrightarrow x = x]$$

(48) is an instance of HSCP* and hence is a theorem of HST*. But (48) posits the existence of a property, call it U, that holds of everything, including therefore its own nominalization. Hence the following will be a theorem of HST*:

$$(49) \exists P [P(\hat{P})]$$

So not only self-instantiation of properties is meaningful in HST*, but it can actually be proven that there are properties which hold of themselves. This shows that the restricted comprehension principle allows us to have (via stratification) the advantages of type theory (i.e. no Russell's property) and at the same time to make sense of self-predication.

To sum up, HST* is a formal system that arises by changing second order logic in two ways: we allow predicative expressions to be nominalized and switch to an h-stratified CP. The resulting system is provably consistent relatively to ZF, and complete. An important thing to notice is that HST* does not necessarily require that properties with the same instances have to be identical. In other words the system leaves open whether properties are to be thought of in extensional or intensional terms. If we want to think of them as sets, all we need to do is to add an axiom saying that two properties with the same instances are identical. The question is especially

relevant for modal extensions of HST* where properties will have varying extensions in different worlds. We might require then that necessarily equiextensional properties must be identical by assuming:

$$(50) \Box \text{Ext}^* \forall R \forall Q [\Box \forall x_1, \dots, \forall x_n [R(x_1, \dots, x_n) \leftrightarrow Q(x_1, \dots, x_n)] \rightarrow R = Q]$$

Assuming $\Box \text{Ext}^*$ amounts to adopting essentially Montague's weakly intensional view of properties, which would allow us to model them as functions from possible worlds into truth values. So it is possible to adopt the general treatment of nominalization that HST* provides without having to abandon weak intensionality.¹² However, HST* as such is not committed to $\Box \text{Ext}^*$, a fact that is particularly interesting in view of the problems that weak intensionality creates. In fact, the "intended" model of HST* that we shall present in sec. 3.4. below is one where $\Box \text{Ext}^*$ fails.

So far we have presented HST* as an uninterpreted formal system, without saying anything about the semantics that comes with it. Now, as already noted, properties are projected in two roles in the logical grammar of HST*: as predicative expressions and derivatively as singular terms. This is the trade mark of Cocchiarella's *-systems. If, as many philosophers seem to think, the role of a logical grammar should be to display somehow the structure of the world, then one of the central theses of

HST* seems to be that properties may be looked at from two different points of view. Qua predicable entities, they appear to be essentially incomplete or "unsaturated" structures. Something like runs cannot stand by itself: it is a structure with a gap in it. If we fill that gap we get a sentence. So properties as predicables should be conceived as intrinsically incomplete, "gapped", or "functional" structures. These structures are whatever they are: platonic concepts, mental constructs, logico-grammatical roles. One might think of them as functions from individuals into propositions, ways of displaying individuals in a state of affairs, etc. In any case on what seems to be the most natural interpretation of HST*, properties as predicable structures should not be regarded on a par with individuals (i.e. saturated particulars).

Properties, however, can be "nominalized" and nominalized predicative expressions can be subjects in predication acts. This strongly suggests that the unsaturated structures associated with predicative expressions (in their primary linguistic role) can somehow be "projected" as individuals, or have individual counterparts, which is what nominalized predicative expressions refer to. So, on this view, properties have two modes of being: one as "intrinsically functional" entities, the other as individuals systematically correlated to those entities. So, for instance, men used as a predicative

expression (as in John and Bill are men) would be associated with something like a propositional function. Used as a subject (as in men are mammals) it purports to refer to an independently existing individual, the kind of men. Similarly, the verb run used as a predicate is something that maps individuals into propositions (whatever the latter may be); used as a subject (in its infinitival or gerundive transform) it may refer to some abstract individual, say the activity of running. This idea of properties having a double nature seems to be close in spirit to the Fregean distinction between concepts and objects (where concepts have to be kept distinct from but do have as correlates objects, conceived possibly as Wertverläufe). In fact, Cocchiarella claims that HST* can be used to provide a fair representation of Frege's view of nominalization processes.

For the purpose of providing a model for HST*, properties will have to be modelled as some sort of set theoretic construct, viz. as saturated entities. For instance, if we adopt weak intensionality, we can represent properties (qua predicables) as functions from worlds into sets. We can introduce, then, a function f that maps properties (as propositional functions) into the domain of individuals. Intuitively, what f does is to associate, in Frege's terminology, each "concept" with its individual projection or objectual correlate. The exact nature of

this individual counterpart of properties may be left for specializations of the general theory to determine. So a model theoretic frame for HST* will look roughly as a triplet of the form $\langle U, P, f \rangle$ where U is the set of individuals, P the set of properties and f a map from P into U . That the set P , whatever structure it might have, is bound to be a set of "saturated" entities only shows that model theoretic constructs are useful tools in studying formal theories. Models, however, are not to be confused with the actual objects that theories are about, though they might be good (or bad) guides to the intended interpretation of a certain theory. A further point to be noted concerning the frames in question is that nothing prevents P from being a subset of U and f from being the identity function. In other words, HST* allows for an interpretation where predicative expressions and their nominalizations have the same semantic value. But, in the general case, this is not to be assumed since nothing in the theory enforces it as a necessary requirement. It might also be pointed out that Cocchiarella's Fregean semantics for HST* and related systems bears a certain resemblance to a semantic technique that should already be familiar to linguists, namely Thomason's (1979) "levelling function". Such a function is supposed to be a map from propositions into the domain of individuals. Thomason mentions briefly that extending the levelling technique to

properties would require major adjustments to semantic theory. Cocchiarella's HST* provides a general framework to carry out such necessary adjustments.

This concludes our informal presentation of the basic features of HST* as a general theory of properties and predication. We now turn to a more formal description of two versions of the system.

3.3 $\lambda\Box$ HST*: the syntax and (weakly intensional) semantics of a modal λ -extension of HST*.¹³

The system we are going to consider is an extension of HST* containing λ 's and modal operators. Such a system will be provided with a standard possible world semantics that validates the thesis of weak intensionality (i.e. \Box Ext*). The reasons for presenting this system are twofold. On the one hand it may help linguists who, like myself, are accustomed to thinking in terms of Montague's IL to work their way smoothly towards more radical alternatives to Montague semantics that can be built on the basis of $\lambda\Box$ HST*. On the other hand $\lambda\Box$ HST* (+ \Box Ext*) provides a very general semantics for nominalization processes that might be appealing for the weak intensionalist, i.e. to those who believe that properties should be thought of as functions from possible worlds to extensions.

The λ -operator in $\lambda\Box$ HST* serves as a device for explicitly defining new properties in terms of open wffs

of the language. Now, properties in $\lambda\Box\text{HST}^*$ are not ranked according to the kind of arguments that they take. So the λ -operator is allowed to bind only individual variables (i.e. only λ -abstracts of the form $\lambda x_1, \dots, \lambda x_n \varphi$ are taken to be well-formed). If we allowed the λ -operator to bind also predicative variables, it would suggest that certain properties can be meaningfully applied only to properties, and this is not so. Properties can be applied meaningfully to any individual, though some properties might be true (and furthermore necessarily so) only of (nominalized) properties. A further preliminary remark concerns the way "improper" λ -abstracts are going to be treated. As we know already, only properties specifiable in terms of h-stratified wffs exist as far as HST^* is concerned. So we have got to do something about λ -abstracts $\lambda x_1, \dots, \lambda x_n \varphi$ where φ is not h-stratified. One way to go would be to assign to those some sort of "null" relation, say the empty n-place relation. It turns out to be simpler, though, for technical purposes, to assume that improper λ -abstracts simply are not well formed terms of $\lambda\Box\text{HST}^*$. The definition of well-formedness for an expression of $\lambda\Box\text{HST}^*$ can be easily adjusted to his end, which will save us a lot of work. Perhaps it should be noted that there are other ways of extending HST^* via the λ -operator which do not have this feature. Cocchiarella develops, for instance a system, called

HST^*_λ , where all λ -abstracts (including non h-stratified ones) denote a property when used as predicative expressions. However, only some properties turn out to have individual projections, i.e. only some predicative expressions refer when nominalized.

In describing $\lambda \square HST^*$ it is convenient to identify the set of types with natural numbers. Intuitively, 0 is taken to be the type of individual expressions, 1 the type of sentences (wffs), 2 the type of 1-place predicative expressions, n the type of n-1-place predicative expressions. It should also be noted that expressions of the language L within which $\lambda \square HST^*$ is going to be described are ranked only according to the adicity of predicates and not also according to the rank of arguments, as in standard type theory. Let Var_n and $Cons_n$ be the variable and constants of type n of a language L. The set of meaningful expressions of L of type n ($ME_n(L)$) is recursively defined as follows:

- (51) (i) $Var_n, Cons_n \subseteq ME_n$, for all n
 (ii) if $\beta \in ME_n$ (where $n \geq 2$)¹⁴ and $\alpha \in ME_0$,
 then $\beta(\alpha) \in ME_{n-1}$
 (iii) if $\varphi, \psi \in ME_1$ and $\alpha \in Var_n$, then: $\neg \varphi$,
 $\varphi \rightarrow \psi$, $\square \varphi$, $\forall \alpha \varphi \in ME_1$
 (iv) if $\beta \in ME_n$ (where $n > 0$), then $\neg \beta \in ME_0$
 (v) if $\varphi \in ME_1$ and $\alpha_1, \dots, \alpha_n \in Var_0$ (where
 $n > 1$), $\lambda \alpha_1, \dots, \lambda \alpha_n \varphi \in ME_{n+1}$

The definition of h-stratification goes as before, except that we now require that for any wff ψ such that $\lambda x_1, \dots, \lambda x_n \varphi$ occurs in ψ all the x_i must receive the same numerical index and the λ -abstract is assigned the successor of this index for ψ to be h-stratified. We then say that a wff φ of L is a wff of $\lambda \square \text{HST}^*$ iff every λ -abstract occurring in φ is h-stratified. We can also define ' \equiv ' (indiscernibility) as usual, i.e.:

$$(52) \text{ } \alpha \equiv \beta \leftrightarrow \forall \xi [\xi(\alpha) \leftrightarrow \xi(\beta)]$$

The axioms of $\lambda \square \text{HST}^*$ are taken to be all the tautologies, plus a standard axiomatic basis for (second order) predicate logic with identity, the axioms of S5 propositional modal calculus, the Carnap-Barcan exchange law for \forall and \square (see Cocchiarella, 1983 for details) and the following:

$$(53) \text{ (Id}^*) \quad \lambda x_1, \dots, \lambda x_n [P(x_1) \dots (x_n)] \equiv P$$

$$\text{(\lambda-Conv}^*) \quad [\lambda x_1, \dots, \lambda x_n \varphi](\alpha_1) \dots (\alpha_n) \leftrightarrow \\ \varphi(\alpha_1/x_1, \dots, \alpha_n/x_n),$$

where α_i is free for x_i in φ

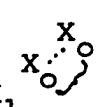
$$\text{(HSCP}^*) \quad \exists F \lambda x_1, \dots, \lambda x_n \varphi \equiv F$$

$$\text{(LL}^*) \quad a \equiv b \rightarrow [\varphi \leftrightarrow \psi]$$

Given LL*, we can safely use '=' for ' \equiv '. Furthermore we should note the simple form that the comprehension principle takes here. $\square \text{Ext}^*$ (see above sec. 3.2.) can be consistently added to $\lambda \square \text{HST}^*$. It is worth pointing out that (1) $\lambda \square \text{HST}^*$ is consistent, if Zermelo-Fraenkel set

theory is, and (2) for any wff φ of $\lambda \Box \text{HST}^*$ there is λ -free wff ψ of $\lambda \Box \text{HST}^*$ such that $\varphi \leftrightarrow \psi$ is a theorem of $\lambda \Box \text{HST}^*$ (i.e. $\lambda \Box \text{HST}^*$ is a conservative extension of (modal) HST^* -- Cocchiarella, 1983).

We turn now to presenting a weakly intensional semantics for $\text{HST}^* + \text{Ext}^*$ in terms of which the completeness results are obtained. This will serve the purpose of illustrating how the Fregean semantics of Cocchiarella's works, but it should not be taken as the intended model of $\lambda \Box \text{HST}^*$. By an intensional Fregean frame we refer to an n -indexed triplet $\mathcal{Q} = \langle X_n, W, f \rangle_{n \in \omega}$ such that:

- (54) (i) X_0, W are non empty sets
(ii) $X_1 \subseteq 2^W$
(iii) $X_{n+1} \subseteq X_1$  n -times
(iv) f is a function from $\bigcup_{n > 0} X_n$ (i.e. the set of n -place properties) into X_0 .

X_0 is taken to be the domain of individuals, X_{n+1} the domain of n -place properties (in particular, X_1 is a set of characteristic functions on the set W). f associates properties to their individual projections; we take the function f to be rigid, i.e. a given property will have the same individual counterpart at all worlds. But this is not forced upon us. By an assignment in an intensional Fregean frame we mean a function g with domain $\bigcup_{n \in \omega} \text{Var}_n$ such that if $\alpha \in \text{Var}_n$, $g(\alpha) \in X_n$. An intensional interpretation for a language, given a Fregean frame \mathcal{Q} ,

will be an ordered pair $M = \langle A, F \rangle$, where F is a function with domain $\bigcup_{n \in \omega} \text{Cons}_n$, such that for any $\alpha \in \text{Cons}_n$, $F(\alpha) \in X_n$. We now define recursively the value of a meaningful expression α of L with respect to an assignment g and an interpretation M (in symbols $V_{g,M}(\alpha)$) as follows:

(55)

- (i) if $\alpha \in \text{Var}_n$, $V_{g,M}(\alpha) = g(\alpha)$; if $\alpha \in \text{Cons}_n$, $V_{g,M}(\alpha) = F(\alpha)$
- (ii) if $\alpha \in \text{ME}_0$ and $\beta \in \text{ME}_n$, $V_{g,M}(\beta(\alpha)) = V_{g,M}(\beta)(V_{g,M}(\alpha))$
- (iii) if $\varphi \in \text{ME}_1$, for any $w \in W$, $V_{g,M}(\neg\varphi)(w) = 1$ iff $V_{g,M}(\varphi)(w) = 0$ similarly for
- (iv) if $\varphi \in \text{ME}_1$ and $\alpha \in \text{Var}_n$, for any $w \in W$, $V_{g,M}(\forall\alpha\varphi)(w) = 1$ iff for all $u \in X_n$, $V_{g(\alpha/u),M}(\varphi)(w) = 1$
- (v) if $\varphi \in \text{ME}_1$, then $V_{g,M}(\Box\varphi)(w) = 1$ iff for all w' , $V_{g,M}(\varphi)(w') = 1$
- (vi) if $\varphi \in \text{ME}_1$, $V_{g,M}(\lambda\alpha_1, \dots, \lambda\alpha_n\varphi)$ is that \underline{u} in ${}^W D$ such that for any a_1, \dots, a_n in X_0 , $u(a_1) \dots (a_n)(w) = 1$ iff $V_{g(\alpha_1/a_1 \dots \alpha_n/a_n), M}(\varphi)(w) = 1$
15
- (vii) if $\beta \in \text{ME}_n$, $V_{g,M}(\wedge\beta) = f(V_{g,M}(\beta))$

An intensional interpretation M for L is a general interpretation iff for all $n \geq 0$, all $\alpha \in \text{ME}_n$ and all

assignments g in M , $V_{g,M}(\alpha) \in X_n$. Satisfaction and truth can then be defined as usual. The completeness results for $\lambda \square \text{HST}^*$ are obtained with respect to all the general interpretations that satisfy all instances of HSCP^* .

The model we have considered should be understood as a mere device for testing certain formal properties of $\lambda \square \text{HST}^*$. This done, we can adopt HST^* as an autonomous framework for natural language semantics. A semantics of English can be provided by means of a translation procedure into the language of $\lambda \square \text{HST}^*$. The David Lewis argument that translating is not doing semantics does not apply here, just as it does not apply to Montague's translations into IL. We know that $\lambda \square \text{HST}^*$ has models (if ZF is consistent), i.e. we know that $\lambda \square \text{HST}^*$ can be interpreted in terms of extralinguistic entities. Natural language semantics (as intended here) is not concerned, in general, with the actual meaning of single words or with the actual truth conditions of particular sentences, but with the structure of meaning. To display this a translation procedure into a formal language (known to be interpretable) suffices. Furthermore, if one task of natural language semantics is to characterize a notion of logical consequence, then this can be done in terms of the notion of "provable in $\lambda \square \text{HST}^*$ ". Completeness guarantees that the set of pairs of wffs picked out by this notion is

the same as the one characterizable in semantic terms. Thus, it would seem that there is nothing that can be accomplished by a direct map from English into models but could not be accomplished by a map from English into the language of $\lambda \square \text{HST}^*$. This allows us to adopt $\lambda \square \text{HST}^*$ as a basis for semantics independently of any particular model that one might want to choose as the "intended interpretation". Still a semantics by translation along the lines just sketched would be model theoretic and truth conditional, since the possibility of putting particular models, as it were, in parenthesis is granted only by the fact that model theoretic interpretation is a known quantity.

3.4 IL_{*}: an HST*-based logical form for English.

We are going to present now the syntax and semantics of a logical language (IL_{*}) that will be used as a language of the translation in providing an explicit semantics for a fragment of English. The logic that we adopt is that of $\lambda \square \text{HST}^*$ without $\square \text{Ext}^*$. The semantics of IL_{*} serves the purpose of illustrating what a model for a strongly intensional logic like $\lambda \square \text{HST}^*$ without $\square \text{Ext}^*$ might look. Such a semantics is meant to be simply a heuristic into possible "natural interpretations" of our new semantic framework.

We will add a new category to our logical grammar: the category of n -place functors. Functors are taken to be expressions of IL_* that so to speak modify the type of other expressions. So, for instance, there can be functors of IL_* that map 1-place properties into other 1-place properties (like adverbs). We regard n -place functors in the way we regard n -place relations: an n -place functor applied to an argument yields an $n-1$ place functor. The addition of n -place functors is completely straightforward from a technical point of view; it amounts essentially to the addition of (third order) functional constants (i.e. constants that apply to properties of basic elements) to our language. This will not take us beyond the second order logic of $\lambda \square HST^*$, because we do not assume the possibility of quantifying over second order functors. IL_* has thus expressions of three logical types: individual expressions, predicative expressions, functor expressions. Its logic, however, is second order, for we only have variables of the former two logical types.

The motivation for this implementation is the intuition that natural language items such as adverbs, prepositions, determiners, complementizers should not be treated on a par with predicates (common nouns, verbs, adjective, predicate nominals).¹⁶ What motivates this claim is the fact that $\lambda \square HST^*$ is a logic for nominalized

predicates: any predicate expression can be nominalized. So if we treat adverbs, prepositions, etc. on a par with predicates they too would be nominalizable (i.e. they would have individual counterparts). But in natural languages they never are.¹⁷ A simple way of accounting for this is to say that things like adverbs or prepositions are out of the domain of the Fregean correlation f between predicates and individuals, which is what we do by assuming that they are higher order functors. Another possibility would be to treat all non predicates as syncategorematic symbols at the level of our logical language. This course strikes me as rather implausible for several reasons. For instance, it would force us to have a probably infinite number of primitive syncategorematic symbols, at least in the case of adverbs (given that adverbs are closed under conjunction and maybe disjunction). This would raise serious learnability problems. Furthermore, syntactic categories that play a role analogous to that of adverbs, prepositions, determiners, etc. are pretty widespread crosslinguistically, yet the meaning of single adverbs or prepositions or determiners varies greatly. If we treat these items syncategorematically how could we express the fact that their type is attested, at least as a tendency, universally? In other words there are a number of linguistic generalizations that seem to get lost on this hypothesis.

There are various other ways in which HST* could be extended to accommodate adverbs, prepositions, etc. Instead of considering all of them, we will pursue our proposal (adding third order functional constants to HST*) and then try to show some nice consequences that this move seems to have.

For the time being, we will adopt a very strict hypothesis on the nature of n-place functors. We will assume that (a) there are no variables of such type and (b) that there are no 4th or higher order functors (i.e. functors that apply to functors). This hypothesis will turn out to have interesting consequences for natural language semantics which we will try to defend in sec. 4. We now turn to a formal characterization of the syntax of IL_* .

- (56) 1. Type is the smallest set such that:
 $e, p \in \text{Type}$; if $b \in \text{Type}$ and $a = e$ or $a = \langle e, {}_n, p \rangle$ (where $\langle e, {}_n, p \rangle = \langle e, \dots \langle e, p \rangle \dots \rangle$), then $\langle a, b \rangle \in \text{Type}$.
 n times

(Remark: e is the type of individuals, p the type of prepositions. For any type we will have denumerably many constants of that type. If $a = e$ or $a = \langle e, {}_n, p \rangle$ we will also have denumerably many variables of type a)¹⁸

2. Recursive definition of ME_a ('meaningful expression of type a') for any $a \in \text{Type}$:
- (i) $\text{Var}_a, \text{Cons}_a \subseteq ME_a$
 - (ii) If $\beta \in ME_{\langle a, b \rangle}$ and $\alpha \in ME_a$, $\beta(\alpha) \in ME_b$
 - (iii) If $\alpha \in \text{Var}_a$ and $\beta \in ME_b$, $\lambda \alpha \beta \in ME_{\langle a, b \rangle}$
 - (iv) If $\varphi, \psi \in ME_p$ and $\alpha \in \text{Var}_a$, then the following are in ME_p as well: $\neg \varphi, \varphi \rightarrow \psi, \Box \varphi, \text{H}\varphi, \text{W}\varphi, \forall \alpha \varphi$
 - (v) If $\beta \in ME_{\langle e_n, p \rangle}$, $\hat{\beta} \in ME_e$

In providing a semantics for IL_* we will treat individuals and propositions as basic. Properties (as predictable entities) will be represented as functions from individuals into propositions.¹⁹ Since we leave open the exact nature of propositions, it will not in general be the case that two logically equivalent properties turn out to be identical entities. A Fregean embedding f will associate each property with its individual projection. Worlds will then be treated essentially as (time-indexed) sets of propositions. Constraints on worlds will guarantee that truth conditions come out right (i.e. if p is true in a world w , then not p is false at w , etc.). In order to guarantee that each formula φ is assigned a proposition p such that p is a function of (the semantic counterparts of) the syntactic operations in terms of

which \mathcal{Q} is build up, we will use algebraic techniques. In particular for each primitive logical symbol of IL_* there will be an operator in the model mapping propositions into new propositions. The universal quantifier will be treated as customarily in algebraic semantics as a function $g_{\forall, n}$ that maps n-place properties into n-1 place properties (e.g. $g_{\forall, 1}$ will map the 1-place property run into the 0-place property everybody runs'). It should again be emphasized that none of these moves are forced upon us by IL_* as such. There are many conceivable alternatives. For instance we might take n-place relations as primitive and define propositions in terms of them and the notion of an individual. Cocchiarella (1983, ch. 6) develops a semantics for $\lambda \Box HST^*$ (and other * -systems) that does not validate $\Box Ext^*$ and does not require having propositions as primitives. Such a semantics is based on a (primitive) predication-relation. Thus, the model we are going to provide represents only one of the many available options and is only meant as a heuristics.

(57) 1. A Fregean predication frame (FP-frame) is an ω -indexed tuple $\mathcal{Q} = \langle U, P, X_n, W, J_{\leq}, f, g_{\forall, m}, g_{\neg}, g_{\rightarrow}, g_{\Box}, g_W, g_H \rangle_{m, n \in \omega}$ where:

- (i) U, P are non empty sets (individuals and propositions respectively)

- (ii) for each $n \in \omega$, if $n = 0$, $X_0 = P$
 if $n \geq 1$, $X_n \in P^{nU}$,
 where for any $Y, Z, Z^Y =$
 Z^Y n -times
- (iii) J_g is a set partially ordered by \leq
- (iv) f is a function from P^{nU} into U
- (v) (a) $g_\square, g_\neg, g_W, g_H$ are all distinct
 members of P^P
 (b) g_\rightarrow is a member of $P^P \times P^P$
 (c) for all $n \in \omega$, $g_{\forall, n}$ is a function from
 $P^{D_n^{D_2} D_1}$ into $P^{D_n^{D_2}}$,
 where for each i , $n > i > 1$, $D_i = U$ or $D_i = X_n$ (for some $n \in \omega$)
- (vi) W is some set of functions from $P \times J$ into $\{0, 1\}$; each $w \in W$ must satisfy the following conditions, for each $p, q, r \in P$, $k \in P^{D_n^{D_1}}$ (where each D_i is as in (vc)) and for each $j \in J$:
- (a) if $p = g_\neg(q)$, then $w(p, j) = 1$ iff $w(q, j) = 0$
- (b) if $p = g_\rightarrow(r, q)$, then $w(p, j) = 1$ iff either $w(r, j) = 0$ or $w(q, j) = 1$
- (c) if $p = g_\square(q)$, then $w(p, j) = 1$ iff for all $j' \in J$ and $w' \in W$, $w'(q, j') = 1$

- (d) if $p = g_H(q)$, then $w(p, j) = 1$ iff there is a $j' \in J$ such that $j' \prec j$ and $w(q, j') = 1$
- (e) if $p = g_W(q)$, then $w(p, j) = 1$ iff there is a $j' \in J$ such that $j \prec j'$ and $w(q, j') = 1$
- (f) for any $e_1 \in D_2, \dots, e_{n-1} \in D_n$ (where each D_i is as in (vc)), $w(g_{\forall, n}(k)(e_1) \dots (e_{n-1}), j) = 1$ iff for all $e \in D_1$, $w(k(e)(e_1) \dots (e_{n-1}), j) = 1$.
2. A value assignment h in a FP-frame \mathcal{CQ} is a function such that if $\alpha \in \text{Var}_e$, $h(\alpha) \in U$, and if $\alpha \in \text{Var}_{\langle e_n, p \rangle}$, $h(\alpha) \in X_n$
3. The set D_a of possible denotations of type a is defined as follows:
- (i) if $a = e$, $D_a = U$
- (ii) if $a = p$, $D_p = P$
- (iii) if $a = \langle e_n, p \rangle$, $D_a = X_n$
- (iv) if none of (i)-(iii) obtains and $a = \langle b, c \rangle$,
 $D_a = D_c^{D_b}$
4. An FP-interpretation is a pair $\langle \mathcal{CQ}, F \rangle$, where \mathcal{CQ} is a FP-frame and F is a function such that for any $\alpha \in \text{Cons}_a$, $F(\alpha) \in D_a$
5. For each $\alpha \in \text{ME}_a$, we define the value of α with respect to an FP-interpretation M and

a value assignment h (in symbols $\llbracket \alpha \rrbracket^{M,h}$) recursively as follows:

- (i) (a) if $\alpha \in \text{Var}_a$, $\llbracket \alpha \rrbracket^{M,h} = h(\alpha)$
 (b) if $\alpha \in \text{Cons}_a$, $\llbracket \alpha \rrbracket^{M,h} = F(\alpha)$
 (ii) If $\beta \in \text{ME}_{\langle a,b \rangle}$ and $\alpha \in \text{ME}_a$, $\llbracket \beta(\alpha) \rrbracket^{M,h} = \llbracket \beta \rrbracket^{M,h}(\llbracket \alpha \rrbracket^{M,h})$
 (iii) if $\alpha \in \text{Var}_a$ and $\beta \in \text{ME}_b$, where $b = \langle a_1, \langle a_2, \langle \dots, a_n \rangle \rangle \dots \rangle$, $\llbracket \lambda \alpha \beta \rrbracket^{M,h} = d$,

where d is that member of $D_a^{D_{a_1} \dots D_{a_n}}$ such that for all $u \in D_a$, $d(u) = \llbracket \beta \rrbracket^{M,h}(\alpha/u)$

- (iv) if $\varphi, \psi \in \text{ME}_p$ and $\alpha \in \text{Var}_a$, then:
 (a) $\llbracket \neg \varphi \rrbracket^{M,h} = g_{\neg}(\llbracket \varphi \rrbracket^{M,h})$
 (b) $\llbracket \varphi \rightarrow \psi \rrbracket^{M,h} = g_{\rightarrow}(\llbracket \varphi \rrbracket^{M,h}, \llbracket \psi \rrbracket^{M,h})$
 (c) $\llbracket \forall \alpha \varphi \rrbracket^{M,h} = g_{\forall, n}(r)(h(\alpha_1)) \dots (h(\alpha_{n-1}))$ where $\alpha, \alpha_1, \dots, \alpha_{n-1}$ are all the (pairwise disjoint) free variables occurring in φ and r is that member of

$D_p^{D_1 \dots D_n}$ (where for $n \geq i \geq 1$, $D_i = D_a$

and a is the type of α_i) such that

for any $e_1 \in D_1, \dots, e_n \in D_n$, $r(e_1) \dots (e_n) = \llbracket \varphi \rrbracket^{M,h}(\alpha/e_1, \alpha_1/e_2, \dots, \alpha_{n-1}/e_n)$

- (d) $\llbracket \Box \varphi \rrbracket^{M,h} = g_{\Box}(\llbracket \varphi \rrbracket^{M,h})$
 (e) $\llbracket W \varphi \rrbracket^{M,h} = g_W(\llbracket \varphi \rrbracket^{M,h})$
 (f) $\llbracket H \varphi \rrbracket^{M,h} = g_H(\llbracket \varphi \rrbracket^{M,h})$

- (v) if $\beta \in ME_{\langle e_n, p \rangle}$, then $\llbracket \wedge \beta \rrbracket^{M,h} = f(\llbracket \beta \rrbracket^{M,h})$
4. An FP-interpretation M is a general FP-interpretation iff for each $\alpha \in ME_a$ and each value assignment $h, \llbracket \alpha \rrbracket^{M,h} \in D_a$
5. A wff φ is true with respect to some $w \in W$, $j \in J$ and some general FP-interpretation M iff for all value assignments $h, w(\llbracket \varphi \rrbracket^{M,h}, j) = 1$
6. If φ is a wff of $\lambda \square HST^*$, then φ is IL_* -valid iff for each general interpretation M , each $w \in W$ and $j \in J$, if all instances of LL^* and $HSCP^*$ are true in M with respect to w and j , then φ is true in M with respect to w and j .

It is clear that all the axioms of $\lambda \square HST^*$ (except for LL^* and $HSCP^*$) are true in general FP-interpretations at any w and j . Now, let the axioms of IL_* be those of $\lambda \square HST^*$ (without $\square Ext^*$) complemented by some proper set of axioms for H and W plus a proper extension of λ -conversion to cover also functors. We may then conjecture that Cocchiarella's completeness results extends to IL_* with respect to the relevant class of general FP-interpretations (i.e. all the interpretations in which all instances of LL^* and $HSCP^*$ are satisfied):

(58) Conjecture: the set of IL_* -valid wffs is equal to the set of theorems of $\lambda \square HST^*$.

It is also convenient, as we will see, to have in the language of IL_* something which can "undo" what the

nominalization operator ' \wedge ' does. We will denote this as ' \smile ' and call it "the predicator". Intuitively, the predicator ' \smile ' is to be interpreted as the inverse of the embedding map f . Syntactically, ' \smile ' maps individual expressions into predicative ones. We may simply define it contextually as follows:

$$\smile x(y) =_{df} \exists P[\wedge P = x \ \& \ P(y)]$$

Notice that the following is IL_* -valid:

$$(58) \ \smile \wedge P(x) \leftrightarrow P(x)$$

(58) is reminiscent of Montague's "down-up" cancellation.

4. IL_* as a Restrictive Semantic Component of the Grammar.

The motivation for implementing IL_* as a logical form for a grammar of English is mainly its general (type-free) treatment of nominalization and its non commitment to weak intensionality. Of course we have not built yet a grammar of English predicative expressions. Probably, though, the Montague grammarian can already imagine the lines along which such a grammar can be constructed. We will devote subsequent chapters of the present work to this task. In a sense, the kinds of arguments in favour of adopting IL_* we have considered thus far are of a "negative" nature. There are phenomena that can be treated only in a very complicated way given

Montague's theory of properties. IL_* allows for a simpler treatment of those phenomena. Essentially, IL_* as a much more simple and flexible type-theoretic structure than IL: only three logical types instead of infinitely many, and furthermore a device for "lowering" arguments (instead of "raising" functions). This gives us the flexibility we need in handling certain facts and this is why IL_* is preferable to IL.

Now, for well over a decade the main concern of (generative) linguists has been the search for constraints on grammars. Given all the grammars one could imagine, only some of them turn out to be possible as grammars of natural languages. The choice of a formal theory (such as say, in syntax, \bar{X} -theory) sets up the logical space of what counts as a possible grammar. Then, the task of a substantive linguistic theory, as usually conceived, is to narrow down this logical space up to a set that (apart from possible "accidental gaps") contains all and only the grammars of natural languages. The focus of research for constraints has been mainly in the theory of syntax. Research in semantics has rather been devoted to specifying in a systematic fashion truth conditions for sentences involving a variety of semantic constructions, in the hope of finding systematic semantic explanations of phenomena for which a satisfactory purely syntactic explanation was lacking. We would like to argue that IL_* represents a

step towards a more restrictive semantic component than the one provided by IL. IL_* can be regarded as a way of constraining the semantic component of a grammar in much the same way in which the recent work on generalized quantification provides a principled account of why only some of the logically conceivable quantifiers are attested in natural languages.²⁰ If IL_* gives us more freedom in handling certain phenomena than IL, it also gives us less freedom in handling other phenomena. There are, specifically, two ways in which IL_* can be viewed as imposing severe restrictions on the structure of meaning. One has to do with constraints that IL_* imposes on the translation mapping and through them on the syntactic constituency of certain constructions (particularly, infinitives and gerunds). We will consider them in detail in chapter III. The second has to do with the types of entities that on the basis of IL_* we should expect to find as possible meanings of natural language expressions. We mentioned in the introduction to this chapter that one of the roles of type theory was to provide a set of classificatory principles for semantic domains and that in this role too there were reasons for being dissatisfied with it. On totally independent grounds we have been led to formulate a theory which contains only a threefold type theoretic distinction. We wish now to suggest that such a theory represents also a better approximation to the

individuation of classificatory devices on semantic domains built into grammars.

To put it in a nutshell, adopting IL_* as logical form in a grammar amounts to making the following two claims: (a) the meanings of natural language expressions have to fit the mold of three and only three logical types, namely (i) individual (or argument) expressions, (typically, NPs), (ii) predicates (or propositional functions: verbs, CNs, adjectives, sentences) and functors (everything else); (b) there are no grammatical processes that involve quantification over functors.

4.1 The three-layers hypothesis.

Let us discuss claim (a) first. Jespersen (1 24 ch. VII) puts forth the idea that there are essentially three "levels" of expressions in a grammar: (1) things that are always arguments (which he calls "primary"), (2) things that take primary elements as arguments (which he calls "secondary") and (3) things that modify other things (tertiary). Jespersen seems to be somewhat uncertain as to whether to recognize further distinctions within this third "level" (as we will be too -- see below). Now, what is this a classification of? What is its motivation? The hypothesis I wish to suggest is that it is a claim concerning the logical types associated with syntactic categories. As far as natural languages are concerned,

the structure of semantic universes is articulated in three layers, which induces a tripartition on the set of syntactic categories. In other words, syntactic categories group up in three natural classes determined by their logical types. Claim (a) above amounts to saying that there is no purely syntactic way to account for why syntactic categories should form just this pattern: the explanation lies in the way categories are related to types.

Some preliminary evidence for this claim is provided by nominalization phenomena: they can affect only predicates (verbs, CNs, adjectives, etc.) and not functors (determiners, adverbs, prepositions, etc.). However, since this argument was used to motivate the system in the first place, we would need some independent evidence to really back it up. I think that such evidence can indeed be found.

Consider first the fact that in natural languages there are operations that manipulate the structure of propositional functions. Arguments of a verb can, for instance, be deleted (by argument drop), swapped around (by passive) or added (by causativization). These argument manipulating operations are usually morphologically marked: something happens to the syntax to signal which alteration of argument structure has taken place. A priori one could conceive similar operations taking place

on functors. For instance, prepositions can be regarded as 2-place functors: they take an NP (e.g. in + the park) to give something which applied to a predicate (e.g. in the park + run) yields another predicate (to run in the park). An operation like, say, passive that swaps around arguments is certainly no less conceivable, from a logical point of view, for prepositions than for verbs. However, operations of this sort are not attested, I would think, universally. In general, there is no analogous argument manipulating operations for things like adverbs, prepositions or determiners. Our system might provide the beginning of an explanation for this fact. Let us illustrate it with an example. Consider causativization. When we form the causative transform of a verb, we alter its argument structure by adding an extra argument (e.g.: a dog walks \leadsto John walked his dog, John keeps laughing \leadsto Mary keeps John laughing, etc.). So causative morphology (which in English can often be zero morphology, but in most languages is overtly realized) applies to, say, a 1-place propositional function (an intransitive verb) and gives us back a two-place propositional function. The semantic counterpart of this operation will therefore be a functor (call it CAUSE) that maps predicates into predicates. Let us imagine that we had an analogon of causativization for, say, adverbs. We would then need a functor CAUSE* mapping adverbs (i.e. 1-place functors from

properties into properties) into "causative adverbs" (possibly prepositions, if they are regarded as suggested by Jackendoff (1972) as "transitive" adverbs) as to have alternations like $\text{slowly}(X) \sim \text{CAUSE}^*(\text{slowly})(X)(Y)$. But now since adverbs are functors, CAUSE^* would have to be a higher order functor that takes functors as inputs. But this cannot be, since in IL_* we have only third-order functors. Our semantics cannot accomodate anything like CAUSE^* : it would not fit its three-layered type-theoretic mold. Similar arguments might be constructed for passive, argument drop, etc. So, it would seem that the lack of argument manipulating operations for prepositions, determiners, etc. may receive at least the beginning of a principled explanation on the hypothesis that the semantic component of a grammar looks like a second order logic with third order functional constants. Obviously such an explanation would not be available on the basis of Montague's IL , for there it is equally easy to define a passive-like operation for any function (no matter which type) as for verbs.

Perhaps, I should restate all this more cautiously. IL_* is not in principle incapable of expressing a higher order functor (such as CAUSE^* above). Any such functor might always be added by some definitional extension. The point is that as it is IL_* makes no provisions for items of this sort. Something special has to be done to the

system in order to accomodate them. So, what IL_* really does is not to legislate once for all what might and what might not exist; but it certainly provides at least a markedness scale concerning what sort of items one might expect to find in natural languages. Things requiring a special definition are marked.

Let me try to forestall right away some of the possible objections that might be raised to our hypothesis. Consider the conjunction and. It doesn't fit straightforwardly any of the logical types we are assuming: the meaning of and is clearly neither an individual nor a property, nor is it a simple functor (given that and does not merely serve to conjoin sentences but items of all major categories). So, something special has to be done about it: our semantic must devise some special operator to represent properly the meaning of natural language conjunction. But the point is that a move of this sort seems to be needed anyhow, on anybody's account.²¹ The cross-categorial character of natural language conjunction, would seem to force us to treat it differently from other non cross-categorial functors. Furthermore, and has a universal character: every language has some device for expressing conjunction and its meaning tends to stay roughly constant across languages. Thus, treating conjunction as a special semantic operator is not costly, because we can regard it as part

of universal grammar, i.e. as part of a (possibly innate) conceptual schema.

In natural languages there are functors (like almost, even, only, very, etc.) that seem to be higher order modifiers (i.e. seem able to modify other functors, e.g. very slowly). On the basis of the preceding discussion of how and is to be accommodated within our theory, we are led to predict that these items should behave similarly to conjunction. So, they should operate cross-categorially and have a tendency to have the same universal semantic content across languages.²² This seems to be roughly right. Consider items like very or almost. They tend to combine in English with expressions of most major categories (e.g. almost combines with adjectives (John is almost right), verbs (John was almost elected), determiners (almost every man), etc. Also it seems quite plausible to expect that some item with the same function as almost is attested in any language. But this is exactly what would justify treating them as distinguished operators, outside of the three-layers mold. Their cross-categorial and tendentially universal character does not make this move costly, for they can be provided for by universal grammar. What would be costly would be a 4th order functor (a functor that modifies other functors) which is specific to certain categories (say, it applies only to prepositions) and has a language particular

semantic content. Special provisions would have to be made in our semantics to accommodate such an item.

We started off saying that higher order functors cannot exist. We see now that this of course is not quite right. What our semantics actually says is that any such "functor" is expected to have certain characteristics (cross-categoricity, universal semantic content) or else is funny (i.e. a marked construction). IL_* as a logical form in a theory of grammar provides us with a markedness scale concerning what sorts of things can be expected to be found among natural language meanings. Such a scale leads us to make empirically testable predictions on the structure of meaning and our preliminary considerations seem to show that such predictions might be borne out.

4.2 The "no functor anaphora" constraint.

Let us now turn to a consideration of claim (b) made above. Saying that there are no variables of a certain logical type amounts to saying that we cannot refer to arbitrary entities of that type. On the assumption that anaphoric processes involve a capacity to refer to arbitrary entities in a given domain (formally represented by a notation with variables ranging over that domain) our system predicts that functors do not enter anaphoric processes in natural languages. Hence, we should expect that processes such as pronominalization, VP-deletion,

wh-movement etc. never involve determiners, prepositions, adverbials, etc. It seems to me that there is something basically right about this generalization. For instance, in general determiners, prepositions, complementizers, etc. do not undergo wh-movement and have no pro-forms. Similarly there is no analogue of VP-deletion or one-anaphora for these items. All such facts would receive a rather simple and uniform semantic explanation within our system.

Items that might constitute a serious problem for the present hypothesis are obviously adverbs. Adverbs seem to enter various anaphoric processes such as wh-movement or comparative formation, and to have proforms (thus, so). Maybe one simply has to allow for a theory that has variables ranging over adverb-meanings and that our claim (b) has to be weakened accordingly. But the simplicity of the "no functor anaphora" constraint makes it rather appealing, and I think it would be nice if one could maintain it in its strongest possible form. Here is a line of defense for it.

There are roughly two major classes of adverbials that should be considered from the present perspective. Adverbials of the first class have a "predicative counterpart", i.e. a predicate to which they are semantically and syntactically related. This is for instance the case for most prepositional phrases. In the park is an adverbial

in John runs in the park and a predicate in John is in the park. One way of accounting for this might be to say that in the park is primarily a predicate that is mapped into a predicate modifier by applying to it some function (say, the adverb-forming function $\lambda P \lambda Q \lambda x [P(x) \ \& \ Q(x)]$). Similarly, manner adverbials such as slowly have a corresponding adjective, slow, from which they are derived. Adverbials of the second class do not have predicative counterparts. Examples are again, too, a lot, almost, etc. For adverbs of the first class it would be possible to treat anaphoric processes that involve them in terms of their predicative counterparts (for which we do have variables) by exploiting the independently needed adverb-forming functions that the grammar makes available (e.g. ly' or $\lambda P \lambda Q \lambda x [P(x) \ \& \ Q(x)]$). Let me try to clarify this by means of an example. Consider wh-movement of adverbs. Let us assume, just for sake of discussion, Karttunen's analysis of questions, according to which the meaning of questions should be represented (roughly) as a set of propositions (the set of "true answers"). On the basis of this analysis we might represent something like (59a-b) as in (60a-b):

- (59) a. Where did John play?
 b. How did John play?

- (60) a. $\lambda p \exists P [p \ \& \ \wedge p = \wedge H [\lambda Z \lambda Q \lambda x [Z(x) \ \& \ Q(x)] (P) (\text{play}') (j)]]$
 b. $\lambda p \exists P [p \ \& \ \wedge p = \wedge H [\text{ly}' (P) (\text{play}') (j)]]$

As you can see in (60) we quantify just over properties and we use then an appropriate adverb-forming function to build the corresponding adverb. So no direct quantification on adverbs would be necessary. This example is meant to be purely illustrative of a plausible strategy that we might adopt to deal with anaphoric processes that seem to involve adverbs.

Now, an approach along the preceding lines would make an interesting, clear cut prediction. Adverbs that do not have a predicative counterpart (i.e. adverbials of the second class above) cannot be involved at all in anaphoric processes, because there is no corresponding predicate to quantify over. This means, e.g., that almost, again, too etc. should not represent a possible answer to how-questions (i.e. should not be able to undergo wh-movement). Similarly, they should not enter comparative formation or act as the antecedent of proforms like thus or so. This seems to be correct.

So there seems to be some evidence that the "no functor anaphora" constraint might be right. By disallowing direct quantificational reference to functors, we are forced to cast what looks like functor anaphora as predicate anaphora in disguise. This apparently strange move, however, seems to shed some light on why just a certain, special class of functors (i.e. adverbials with

predicative counterparts) rather than other conceivable classes display anaphoric tendencies.

A possible problem for the above line of explanation might be represented by time and degree modifiers (for two days, at a faster speed, etc.). They can be wh-moved (when, how fast, etc.) and can be antecedents of proforms (that, so), but do not seem to always have natural predicative counterparts, e.g.:

- (61) a. ?? the trip was at a fast speed
 b. ? the trip was for two days

The point in this case is, however, that independent considerations seem to show that we have got to have times and degrees among our individuals. Consider for instance:

- (62) a. the time of his arrival is June 15
 b. John's height is amazing

What do the referring NPs the time of his arrival and John's height refer to? Whatever analysis we wish to put forward for times and degrees they will have to be, it seems, in our domain of individuals. Hence we can simply use individual level quantification to deal with anaphoric processes which involve them.

In conclusion, the no functor anaphora constraint seem able to account not only for the impossibility of certain logically conceivable anaphoric relations (e.g. the lack of wh-phenomena for prepositions, determiners, complementizers and the like) but also for why other

anaphoric relations involving functors (some adverbials, degree modifiers) are instead possible. It is clear that no comparable explanation would be available within Montague's IL, where we have variables for (and can quantify over) any logical type whatsoever.

4.3. Logical types vs. semantic sorts.

The preceding considerations are preliminary and speculative to a large extent. Actually testing the adequacy of the three-layers hypothesis and the no functor anaphora constraint will require a lot of work. But what is interesting is that we are led to these claims on the basis of a semantic theory of properties which has been motivated on totally independent grounds (nominalization and weak intensionality). Such a theory directs us towards a restrictive and a priori unlikely hypothesis on the structure of natural language semantics which, strikingly, seems to make a number of correct predictions (and thereby provides an interesting explanatory model). In the next chapters we will discuss a number of other predictions that derive from the present semantic theory, which we hope to convince the reader are less speculative.

We have achieved this by drastically cutting down the number of logical types to be employed in a semantics for natural languages. In Montague's system there was a fairly unstructured set of basic entities (the set D_e);

all the "structural" aspects of meanings where represented in terms of logical types. We have tried to argue that this approach runs into serious problems.

A logically conceivable alternative would be to make a move similar to the one made in shifting from type theory to set theory in mathematical logic. We could get rid of types altogether by having a richly structured domain, similar to the domain of set theory. We could also explicitly sort out the domain of individuals (by using a many sorted logic) to make sure that we have all the entities we may need. Along such lines it would be possible to do natural language semantics in terms of a first order language. Below we will consider this possibility in more detail, and point out some problems it runs into. A view similar to this is for example adopted in Chierchia (1982, 1983). Even though the theory adopted there is also an application of Cocchiarella's HST*, the "intended" interpretation of the nominalization function which was explicitly assumed was that such a function is just the identity map. This interpretation, however can arise only in very special marked circumstances (see the discussion in Chierchia 1982, pp. 321-323).

Between these two extremes (all the possible types - no types) there are many intermediate solutions. One of them is the one we are exploring here. According to it certain aspects of meaning, those that pertain to predica-

tion and predicate modification, should be handled as differences in logical types. Other relevant semantic distinctions should instead be handled by sorting out the domain of individuals, as we will see in subsequent discussion. We are defending here the idea that natural languages use exactly three logical types. Whether or not we succeed in making a convincing defense, the present work should be able to show that constraining Montague's type theory is a necessary and rewarding enterprise, and that there are interesting empirical issues that can be seen in a new light from this perspective.

5. Some Comparisons.

We will now consider some of the many alternatives to Montague's theories of properties and try to compare them with the one we adopt. A detailed consideration of all of them would take us too far afield. So the following considerations will have to be to a large extent preliminary and informal. We will first consider approaches that essentially maintain type theory as a framework for modelling predication and then we will consider non type theoretic approaches. Throughout the following discussion, it might be worth to bear in mind what aspects of our theory we take to be fundamental, namely:

- a) the possibility of turning properties into arguments of other properties of the same level

- b) the distinction between properties qua propositional functions vs. properties qua individuals
- c) having a sufficiently general theory of predication

We do not consider the particular axiomatization of IL_* proposed above as fundamental as properties (a)-(c). The general line of our argumentation concerning alternative approaches to nominalization will be that none of them has all of the characteristics (a)-(c). Almost all of the approaches we will consider could be extended as to satisfy (a)-(c). It is not clear, though, whether the result of such extensions will turn out to be essentially different from the framework we already have developed.

5.1 Cresswell's models for λ -categorical languages.

The system of Cresswell (1973), further extended in Bigelow (1978), allows for a very general treatment of nominalization. Cresswell defines a typed (categorical) language containing λ 's and provides a system of semantic domains in terms of which such language is interpreted. The construction of the models for Cresswell's λ -categorical language can be informally described as follows. One starts off with two basic domains D_0 (intuitively the set of possible denotations of the singular terms of the language) and D_1 (the set of possible proposition values). Then one builds recursively sets of partial functions out of D_0 and D_1 by means of standard set

theoretic techniques (i.e.: $D_{\langle \sigma, \tau_1, \dots, \tau_n \rangle} = D_{\sigma}^{D_{\tau_1} \times \dots \times D_{\tau_n}}$).

Finally one requires that all the sets of partial functions so constructed be themselves subsets of D_0 (i.e. potential values of individual variables). Since the functional domains contain only partial functions, we can let them be members of their own ranges by assuming that they will always be undefined when applied to themselves.

There are several objections that can be raised against Cresswell's models viewed as a theory of natural language nominalization. First it should be pointed out that Cresswell never actually shows that those models exist. He gives, however, a plausibility argument by pointing out that the set D_0 ought to look more or less like the domain of a model of set theory. Also, Cresswell analyzes propositions as sets of worlds, which commits him to a weakly intensional theory of property. I don't know what would be the effect of dropping this assumption within his theory. More compelling seems to be the fact that Cresswell's models are still built upon the type theoretic tenet that self-attribution of properties is meaningless (functions in his system have to be undefined when applied to themselves). So if one adopts it as the basis for a theory of nominalization in natural languages, one is bound to do a considerable amount of explaining away for those cases where self attribution seems to arise (e.g. being crazy is crazy, etc.). Finally, there is a

sense in which Cresswell's construction appears to be too powerful for natural language semantics. All the functional domains are pushed back down into D_0 . If we take this as a theory of nominalization in natural languages, then we ought to expect everything to be nominalizable (including determiners, complementizers, etc.). If we don't take Cresswell's construction directly as a way of nominalization, then such a theory has to be built on top of it as an extra layer.

5.2 Parson's "floating types".

T. Parsons (1979) has also developed a semantic technique to handle the problems that arise in Montague semantics in connection with nominalization. On the basis of this technique, Parsons defines a PTQ-like fragment of English that can accommodate items such as fun or even property that on a type theoretic approach have to "float" endlessly in a sea of types. According to Parsons, in order to accommodate nominalization phenomena one has to abandon the idea that the relation between syntactic categories and semantic types can be specified as a simple function like the one proposed by Montague. A more complex indexing system seems to be called for. In Montague's semantics, types are a system of indices that indexes (a) the family of semantic domains in the model and (b) the set of expressions of IL, the intermediate

language that in some sense overtly reflects the structure of the semantic system. What Parsons does is to suggest using types also to index sets of expressions of English. English expressions will then be indexed by two systems of indices: the set CAT of categories and the set Type of types. The set CAT is designed to capture syntactic generalizations about the distribution of expressions. The set Type is primarily designed to encode a theory of semantic function-argument structure; but now it must also be used as part of a feature system built into the syntax of English. According to Parsons' proposal we will now have a system of categories supplemented by type-indices, e.g. in the following way:

(63) Category	Type
S	t
CN ^τ	<τ, t>
IV ^τ	<τ, t>
NP ^τ	<<τ, t>, t>
ADV ^τ	<<τ, t>, <τ, t>>
etc.	

τs are variables ranging over types. Parsons calls the syntactic categories indexed by types "fixed types". A consequence of Parsons move is that we will not be able to tell what sort of semantic function is associated with a word α by knowing the syntactic category of α. By knowing that, e.g., α ∈ IV^τ we know that α is associated with some characteristic function (<τ, t>), but we no longer know what the domain of that characteristic

function is. Words like be fun would be of category IV^e when applied to individuals (i.e. we would have a word which might be denoted as befun^e or be fun [+ e-level]); when applied to, say, gerunds, be fun would be in $IV^{<e,t>}$ (i.e. be fun^{<e,t>}), and so on. Parsons claims that in this way items of various categories would no longer be associated with types but with type-templates which should now characterize the structural aspect of meanings previously characterized by types. However, it should be pointed out that the discriminatory power of type templates is very weak. For instance, members of $IV^{<e,t>}$ and of NP^e turn out to be associated with the same logical type, even though they certainly do not form a natural class in any linguistically significant way.

Furthermore, as Parsons points out, we will need an infinity of type levels for each category. Consider for instance the VP have a property. Clearly we can say, possibly truthfully, something like everything has a property. The only way of making sense of this sentence, given a type theoretic approach is by assuming that what we are really saying is that every entity of type τ has a property of type $\langle \tau, t \rangle$. This seems to show the need for a VP having a property ^{τ} for any type τ .

It should be pointed out that while on Parsons' approach we have somehow a way of accomodating items such as be fun in a grammar of English, this is accomplished at

the price of (a) giving up Montague's strong requirement that categories and types are functionally related (i.e. the transparency principle), (b) weakening the classificatory power of classical type theory and (c) building an extra indexing device into the syntax. This last piece of machinery is used to rule out things like:

(64) John^e is fun^{<<e,t>, t>}

In (64) we are applying an <<e,t>, t>-level be fun to an e-level NP. (64) can be ruled out in either of two ways. Either we say that (64) is syntactically ill-formed or that it is semantically deviant.

Strictly speaking, however, be fun^e and be fun^{<e,t>} are two different words with the same phonological realization. In other terms, an item like be fun is treated on Parson's approach as being indefinitely ambiguous. It is unclear what disambiguation strategy such a theory would have to ascribe to the speakers of the language in order to avoid the undesirable result that each be fun would have to be learned separately. Maybe, though, be fun really has different meanings when applied to an individual and when applied to a property. It is quite clear intuitively that different acceptations are involved in say John is fun and jogging is fun. But so it is with New York is fun, that cloud is fun, syntax is fun (!?), etc. All the acceptations in which be fun is used in the preceding sentences are very different from each

other. Why should our semantics attribute such a special role to the distinction between, say, an e-level and an $\langle e, t \rangle$ -level predicate? Aren't other sortal distinctions at least as clear cut? One is lead to suspect that the special role attributed to e- versus $\langle e, t \rangle$ -level predicates is merely an artifact of a type theoretic approach to predication. Moreover, all the different acceptations of be fun are clearly related. Hence an account of this relatedness is called for.

What Parsons suggest in this regard is that there is a basic sense of be fun out of which all others are somehow derived. So one has to learn only this basic sense, all the more specialized meanings being derivable by general principles. The way Parsons proposes to account for the relatedness of floating items is by a further elaboration of the (semantically motivated) indexing of syntactic categories. One can define a set of "word-forming" functions and assign them to categories on the basis of the category of their outputs. Let π be a function from types into types, and let us indicate by $\pi(\zeta)$ the fact that ζ is an argument slot of π (i.e. π is a function of ζ). A word-forming function g will be a function from types into words of fixed types (which, let us recall, are type-indexed categories); we will assign g to a category $A^{\pi(\zeta)}$ iff g applied to a type τ gives a word of fixed type $A^{\pi(\tau)}$.²³ So, for instance we

might have a function g_1 in $VP^{\Pi(\zeta)}$, where $\Pi : \tau \rightarrow \langle \tau, t \rangle$, such that for any τ in Type, $g_1(\tau) = \text{be fun}^{\langle \tau, t \rangle}$. So, g_1 is a function that generates all the be funs and thereby in some sense "ties together" their meanings. We might call g_1 "be fun ζ ". In a similar way we might treat all floating items. Appropriately, Parsons calls the categories of these word-forming functions "floating types". Floating types are probably not to be taken as categories of English syntax, but maybe as categories of the word-formation component of a grammar.

Parsons offers, thus, a way of somehow systematizing the ambiguity of the various be fun's by generating those words in terms of a function sensitive to types. The solution is highly ingenious and probably the best one can give within type theory. It should be noted, however, that what we were after according to Parsons himself was a way of capturing the relatedness in meaning of different be fun's. What we get, instead, are word-forming functions. It would seem that the intuitive insight we are after is not really reflected by the formalism proposed by Parsons.

What about everything has a property? Here Parsons follows essentially Russell's appeal to pragmatics. When we utter sentences like the preceding one, to the extent to which we really intend to speak about everything, we use them as schemata for the infinite number of sentences

that our grammar can generate in terms of the word-forming functions thing² and property² .

As a summary it is worth singling out in schematic form the main drawbacks of Parsons' approach to nominalization phenomena.

- (65) a. Parsons theory has to introduce into the grammar complex indexing procedures to keep track of the type theoretic structure of function argument application. These procedures have no independent syntactic support. Semantically the only reason for introducing them seems to be to get the type theoretic function-application right.
- b. Parsons theory weakens the relation between syntactic categories and semantic types to a complex relation. The classificatory power of type theory (for natural language semantic universes) is accordingly weakened.
- c. Given Parsons approach, the fact that natural languages exploit some set of fixed types rather than others is left in need of further explanation. (Why for instance are there no words of fixed type $CN\langle\langle e,t\rangle, \langle e,t\rangle\rangle$ -- i.e. "adverb-level" CNs?).
- d. Self instantiation of properties has to be treated as a schema whose instances are infinite type theoretically good predications.

5.3 Scott's semantics for the untyped λ -calculus.

Recent work by Turner (1982, 1983) has drawn linguists' attention to the potential that Scott's semantics for the λ -calculus might have for dealing with phenomena related to nominalization. Scott's semantics provides a way of constructing domains that allow embeddings of their own function spaces. By using Scott's technique, we can

build domains that satisfy conditions such as the following: $D \cong [D \rightarrow D]$, where ' \cong ' is to be read as 'is isomorphic to', and $[D \rightarrow D]$ is some specified subset of D^D .

Each member of $[D \rightarrow D]$ will have an isomorphic copy in D , i.e. there is a function ψ from $[D \rightarrow D]$ into D such that for any $k, h \in [D \rightarrow D]$, if $h \neq k$, then $\psi(h) \neq \psi(k)$. If we assume that $[D \rightarrow D] = D^D$, then by Cantor's argument we know that an isomorphism such as ψ cannot exist. *

It is convenient in what follows to sketch briefly those aspects of Scott's semantics that might be relevant in constructing models for nominalization phenomena. We will not be able to present here any proof for the claims we will make. For this the reader is referred to the aforementioned papers by Turner, whose exposition we will follow closely. This done, we will try to indicate points of contact and differences between an approach based on Scott's semantics, as we understand it, and the one we are adopting.

5.3.1 A sketch of Scott's semantics. Scott's construction can be carried out in the domain of complete partially ordered sets. A complete partially ordered set is an ordered pair $\langle D, \leq \rangle$ such that (i) \leq is a partial order on D , (ii) for any ω -chain $d_0 \leq d_1 \leq d_2 \dots$ of elements of D there is a unique upper bound

$(\bigcup_{n \in \omega} d_n)$ and (iii) there is an element \perp of D such

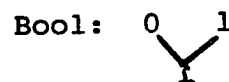
that for each $d \in D, \perp \leq d$. So, instead of working with simple sets as in Montague semantics, we will use sets that have the additional structure of complete partially ordered sets. Following Turner, we will reserve in the present section the term "semantic domain" to any complete partially ordered set. As pointed out above, in building a domain D such that $D \cong [D \rightarrow D]$ we will have to limit ourselves to some selected subset of D^D . In particular we can take $[D \rightarrow D]$ to be the set of continuous functions in D^D . A function f in D^D is said to be continuous iff for each ω -chain $d_1 \leq d_2 \leq d_3 \dots$ in D , f

$$\left(\bigcup_{n \in \omega} d_n \right) = \bigcup_{n \in \omega} f(d_n).$$

A continuous function is, then, an order preserving map defined on semantic domains. It can be shown that given two semantic domains A and B , the set of continuous functions $[A \rightarrow B]$ also forms a semantic domain.

We will now illustrate Scott's construction by defining from two basic domains A and Bool a domain E_∞ isomorphic to IV_∞ , where $IV_\infty = [E_\infty \rightarrow \text{Bool}]$. Following Turner, then, we will try to indicate how a domain so constructed could be used in defining models for English.

A and Bool are to be thought of as semantic domains pictorially represented as follows:



Intuitively, A might be regarded as the set of basic entities and $Bool$ as the set of truth values (augmented by ' \perp '). The construction of E_∞ proceeds as follows. We first define a sequence of domains E_0, E_1, \dots where each E_n contains functions constructed from E_{n-1} according to the following definition:

$$(67) \quad \begin{array}{ll} E_0 = \{\perp\} & E_{n+1} = A + IV_n^{25} \\ IV_0 = [E_0 \rightarrow Bool] & IV_{n+1} = [E_n \rightarrow Bool] \end{array}$$

We also define inductively two sequence of functions

$\psi_n: E_{n+1} \rightarrow E_n$ and $\phi_n: E_n \rightarrow E_{n+1}$. Intuitively, ψ_n provides a counterpart of each member of E_{n+1} in the "lower" set E_n and ϕ_n provides a counterpart of each E_n in the "higher" set E_{n+1} .

ψ_n and ϕ_n are defined recursively as follows:

$$(68) \quad \begin{array}{l} \text{if } a \in E_0, \phi_0(a) = \perp_{E_1} \\ \text{if } a \in E_1, \psi_0(a) = \perp \\ \\ \text{for } a \in E_n \quad \phi_n(a) \quad \left\{ \begin{array}{l} = a \text{ if } a \in A \\ = a \circ \psi_{n-1}, \text{ otherwise} \end{array} \right. \\ \\ \psi_n(a) \quad \left\{ \begin{array}{l} = a \text{ if } a \in A \\ = a \circ \phi_{n-1}, \text{ otherwise} \end{array} \right. \end{array}$$

These functions satisfy $\psi_n(\phi_n(f)) = f$ and

$\phi_n(\psi_n(g)) \leq g$. In terms of them we will be able to

define E_∞ along the following lines:

$$(69) \quad E_\infty = \left\{ \begin{array}{l} \langle e_n \rangle_n \quad : \quad e_n \in E_n \text{ and} \\ \quad \forall n (\psi_n(e_{n+1})) = e_n \} \\ \text{and where for } \langle e_n \rangle_n \in \omega, \langle d_n \rangle_n \in \omega \\ \text{in } E_\infty \\ \langle d_n \rangle_n \in \omega \leq \langle e_n \rangle_n \in \omega \text{ iff } \forall n (d_n \leq e_n) \end{array} \right.$$

Now consider the set $IV_{\infty} = [E_{\infty} \rightarrow \text{Bool}]$. We can define function application in the following way: for each $e \in E$ and $f \in [E_{\infty} \rightarrow \text{Bool}]$,

$$(70) f(e) = \bigcup_{n \in \omega} f_{n+1}(e_n)$$

What is particularly interesting is that E_{∞} and $[IV_{\infty} + A]$ are isomorphic. It is possible to define an isomorphism $\rho : [[IV_{\infty} + A] \rightarrow E]$ as follows:

$$(71) \text{ for each } a \in [IV_{\infty} + A],$$

$$\rho(a) \begin{cases} = a & \text{if } a \in A \\ = \bigcup_{n \in \omega} (\lambda y \in E_n. (a(y))_n), & \text{otherwise} \end{cases}$$

It can be shown that ρ is well defined and one-to-one.

5.3.2 Sketch of an application. Scott's construction allows, thus, to have sets IV_{∞} and E_{∞} such that each member of IV_{∞} will have exactly one isomorphic copy $\rho(a)$ in E_{∞} . How could all this be put to work in dealing with the problems considered in section 1? The simplest move, suggested by the notation adopted would be to let IVs take their values in IV_{∞} , to let the denotations of NPs take their values in E_{∞} and to let ρ play a role similar to that of our Fregean correlation function. So, consider a case of self-application of properties such as:

$$(72) \text{ being crazy is crazy}$$

Let $\llbracket \cdot \rrbracket$ be the value assignment function. The meaning of (72) might be represented as something like:

$$(73) \llbracket \text{crazy} \rrbracket (\rho(\llbracket \text{crazy} \rrbracket))$$

By (70), (73) would be equivalent to:

$$(74) \bigcup_{n \in \omega} \llbracket \text{crazy} \rrbracket_{n+1} (\llbracket \text{crazy} \rrbracket_n)$$

In very rough terms, (74) is an ω -sequence of function applications where each $\llbracket \text{crazy} \rrbracket_n$ applies to its own initial segment $\llbracket \text{crazy} \rrbracket_{n-1}$ all the way up to ω . Furthermore, each $\llbracket \text{crazy} \rrbracket_n$ will be the "best approximation" to $\llbracket \text{crazy} \rrbracket_{n+1}$ we can get in E_n , i.e. when we restrict $\llbracket \text{crazy} \rrbracket_{n+1}$ to E_n what we get is just $\llbracket \text{crazy} \rrbracket_n$. So this gives us a way of modelling self-application of functions.

All this is extremely sketchy. However it should suffice to convince us that Scott's semantics is certainly of great potential interest for building domains representing structural properties of natural language universes of discourse. Various questions now arise. Can Scott's semantics be turned from a theory of functionality (which is what the λ -calculus is) into a theory of properties and predication as general as the one needed by natural language semantics? And, more specifically, can we construct on the basis of Scott's approach a substantive linguistic theory of predication (i.e. a theory with linguistically relevant empirical content)? The answer to these questions appears to be far from obvious.

We certainly won't be able to provide them within the limits of the present work.

What we can do, however, is to point out a few questions that might be relevant in addressing the issue. We will also attempt a very preliminary consideration of some differences between Scott's semantics (or rather Turner's version of it) and our IL_* .

A first question that I believe deserves serious consideration concerns the expressive limits of Scott's semantics. One of the crucial aspects of semantic domains is the drastic limitation imposed on the construction of function spaces: we are allowed to have only continuous functions. If we represent all natural language meanings as continuous functions are we going to have an adequate representation of truth conditions for natural languages? It is not obvious that we do. The intuitive idea underlying the notion of continuity adopted in Turner's formulation of Scott's semantics is that the value of a function g for an argument \underline{a} should be determinable in terms of all the values of g for all the "finite segments" of \underline{a} . From a linguistic point of view this would be a very appealing condition, for it would make meanings finitely representable entities (which, therefore, one could hope to be graspable by finite brains). But, always speaking at an intuitive level, meanings in natural languages might not always work that way. For example,

all of arithmetic and real number analysis can be developed in HST* (supplemented with an axiom of infinity -- see Cocchiarella (1983) for discussion). It seems plausible to maintain that arithmetic and analysis can also be developed within natural languages. However, the λ -calculus is only a theory of recursivity and, unless extended in a substantial way, lacks the expressive power of HST*. Hence, it would seem, it also lacks the expressive power of natural languages, with respect to the possibility of doing arithmetic and real number analysis.

These considerations are of course informal and inconclusive. One will have to wait for more detailed applications of Scott's semantics to natural languages, to really see whether they provide adequate representations for natural language meanings. At any rate, if Scott's semantics for the untyped λ -calculus is to be incorporated into a semantic theory for natural languages, some sort of foundational question concerning expressive power has to be addressed, since such a semantics works by drastically limiting the standard set-theoretic devices for building functions. These limitations may turn out to be all for the best. But we cannot just take this for granted.

With respect to the differences between a Scott-like approach and one based on IL_* , two obvious ones are that the former, but not the latter, is necessarily committed to truth value gaps, and that the latter is not committed

to weak intensionality, while I don't know how the former stands with respect to this. Also, as Turner underscores, Scott's technique allows us to provide semantic domains that satisfy any condition specified in terms of function space construction, cartesian products and disjoint union. So it is just as easy to write conditions that require adverbs, prepositions, etc. to be nominalizable as to write down conditions to the contrary. So in terms of Scott's semantics we can state the fact certain items in natural languages do not nominalize, but we seem to lack any principled basis for making such statements. This is not so for IL_* (as we tried to show in section 4) which is precisely what makes it an interesting basis for natural language semantics. In relation with this, we will also try to argue in chapter III that IL_* , given some plausible assumptions, really forces us to regard tensed VPs as propositional functions and infinitives and gerunds as their nominalized counterparts. This in turn will lead us to make interesting claims concerning the constituency of various constructions involving infinitives and gerunds. On the other hand, on Turner's approach it seems to be equally possible to regard tensed and untensed VPs as being either members of, say, E_∞ or of IV_∞ (because of the isomorphism between the two). In this way, as the reader will be able to verify, we would lose an interesting insight in the nature of infinitives

and gerunds. What these considerations seem to show is that some work seems to be needed before applying Scott's technique to natural languages, if we want to achieve the results that we are able to achieve on the basis of IL_* . But this does not entail that it can't be done.

A further point worth discussing is the way self-application of properties is modelled in Scott's semantics. If you look back at (74), it is evident that strictly speaking we never apply a function to itself. What looks like self application really is an infinite sequence of applications where a function of level n is applied to something of level $n-1$. So if you believe that self-predication is what the name suggests, i.e. attribution of a property to itself, then probably a semantic analysis of (72) along the lines of (74) will not be fully satisfactory. Of course, one might reply that Scott's construction is only a way of modelling self-predication and that any structure that matches it will do. But in this form, such a reply would be rather vague. Either we take Scott's semantics as a theory of properties which can accommodate self-instanciation, and then self-instanciation is what that theory says it is. Or we take Scott's technique as a model theoretic heuristics for building a theory of properties, but then it cannot be taken as such to be such a theory.

Let me emphasize once more that I am not trying to argue that applications of Scott's semantics to natural languages are uninteresting. On the contrary, I think that it constitutes a very promising field of research. What I am suggesting, however, is that such applications are probably not as direct and straightforward as e.g. Turner (1982) seems to suggest.

Of particular interest might be a detailed comparison between HST* (and in general Cocchiarella *-systems) and Scott's semantics, for something very similar seems to be going on in the two approaches from an intuitive point of view. It might turn out to be the case that Scott's semantics for self-application of functions provides specialized models for HST* (regarded as a general theory of properties) models that might be taken as the "intended ones" if the basic theory is channelled in certain ways by integrating it with further assumptions.

Let me try to give more plausibility to the preceding conjecture by sketching a very tentative procedure to turn Scott's domains as those employed by Turner into models of HST*. For the sake of simplicity, let us limit ourselves to a non modal version of monadic HST* with an axiom of extensionality for properties (call the resulting system NFU*).²⁶ Cocchiarella provides models of NFU* in terms of Fregean frames of the following form $\langle U, X, f \rangle$, where U is the domain of individuals, $X \subseteq 2^U$ and $f: X \rightarrow U$.²⁷

The models of NFU* are represented by the class of Fregean frames that verify every instance of the h-stratified version of the comprehension principle. Now, let E_∞ , IV_∞ and ρ be defined as in sec. 5.3.1. Let, furthermore G_{IV_∞} be a family of functions defined as follows:

$$(75) G_{IV_\infty} = \{ g \in 2^{E_\infty} : \exists h \in IV_\infty \text{ such that for each } e \in E_\infty, g(e) = 1 \text{ iff } h(e) = 1 \text{ and } g(e) = 0, \text{ otherwise.} \}$$

Clearly for each h in IV_∞ there is a g in G_{IV_∞} such that for any $e \in E_\infty$, $h(e)$ yields 1 iff $g(e)$ yields 1. So if IV_∞ is used to model predicates, for each predicate in IV_∞ there will be a predicate in G_{IV_∞} that holds true of just the same individuals. We can now further define a map γ from G_{IV_∞} into E_∞ in the following way:

$$(76) \text{ for each } g \in G_{IV_\infty}, \gamma(g) = \rho(h), \text{ where } h \text{ is some member of } IV_\infty \text{ such that for any } e \in E_\infty, h(e) = 1 \text{ iff } g(e) = 1$$

For any set E_∞ , IV_∞ and any map γ defined as above, let $\langle E_\infty, G_{IV_\infty}, \gamma \rangle$ be a Scott-frame. Clearly any Scott-frame is a Fregean frame (since $G_{IV_\infty} \subseteq 2^{E_\infty}$, and $\gamma: G_{IV_\infty} \rightarrow E_\infty$). Now the questions to ask seem to be the following. Would any Scott-frame satisfy HSCP*, i.e. the (stratified) comprehension principle of NFU*? If this is so, then such frames would yield models of NFU*. Or would rather only some Scott-frames satisfy HSCP*? If so it would be interesting to ask which version (if any) of the comprehension principle would all such frames satisfy. It might turn out to be an unstratified comprehension

principle such as the one of Cocchiarella's T* (see Cocchiarella, 1974).

If any of the preceding conjecture should be borne out, then by adopting Cocchiarella's *-systems one could use Scott's domains as specialized (possibly "intended") models.

5.4 First order theories of properties.

In the philosophical literature, a number of first order theories of properties have been recently developed (e.g. Jubien 1981, Bealer 1982). Such theories provide a type-free notion of predication and are not committed to weak intensionality. Hence they provide potentially interesting alternatives to Montague semantics. It is not our task to evaluate the respective merits and faults of first vs higher order theories of properties from a philosophical point of view in general. What we will do in what follows is to point out the main reasons that led us to choose a second order theory (such as HST*) over first order alternatives from the perspective of a theory of grammar. For the sake of explicitness we will focus the discussion on the theory developed in Bealer (1982). We believe, though, that the remarks made below, are of a general nature and apply to any first order theory of properties.

Bealer's theory is cast in a language which contains individual variables and constants (t_1, \dots, t_n, \dots) and predicative constants (F_1, \dots, F_n, \dots). A formation rule of the language guarantees that if φ is a wff and v_1, \dots, v_n are distinct individual variables occurring in φ , then $[\varphi] v_1, \dots, v_n$ is a singular term. A singular term of the form $[\varphi] v_1, \dots, v_n$ is taken to denote a proposition if $n = 0$, a 1-place property if $n = 1$, a 2-place relation if $n = 2$, etc. So in Bealer's logical grammar, the operator '[']' plays a role similar to that of the name-building device '^' in IL_* , viz. turning predicative expressions into individual terms. However in Bealer's theory, a predicative constant, say F , and its "nominalization", $[F(x)]_x$, have the very same semantic value. So the distinction between properties qua "predicables" and properties qua "individual projections" does not arise. Predicative expressions and their nominalizations are distinct only syntactically. Bealer explicitly criticizes Frege precisely because he regards the distinction between predicates and their nominalizations to be also semantic in nature. As we have partly tried to argue in preceding sections, from a linguistic point of view it seems extremely hard to account for the distribution of VPs and their nominal counterparts (i.e. infinitive and gerunds) in purely syntactic terms. The most extensive attempt to provide an autonomous syntactic

explanation of the differences between tensed and untensed VPs is the one put forward within GB (which we will discuss in chapter III). Such an account treats infinitives and gerunds as being syntactically clausal with a phonologically null subject. On this view it would seem that nothing other than sentences could play the role of singular terms in a logical grammar that one might employ as semantic theory. In other terms, the only syntactic attempt to provide a principled explanation of different grammatical properties of tensed and untensed VPs would make the most interesting part of a theory of properties unnecessary for a semantic analysis of natural languages. On the other hand, if one rejects the GB analysis of infinitives and gerunds, and one tries to regard them as "noun-like" counterparts of predicative expressions, then a semantic explanation of their different behavior is called for. And this is what HST*, by making precise Frege's insight, provides. The semantic value of runs is something like a propositional function; the semantic value of to run or running is an individual systematically related to that propositional function.

In Bealer's theory a special logical 2-place constant ' Δ ' is singled out to play the role of a predication of instantiation relation. The logical behavior of Δ is determined by a set of axioms (basically, the axioms of some set theory minus extensionality). The theory is

provided with an algebraic semantics with respect to which the main metatheoretical results (consistency and completeness) are provided.

Besides the syntacticization of the difference between predicates and their nominalizations, another aspect of a first order property theory that appears to be problematic from a linguistic point of view is the way in which adverbs, prepositions, determiners, etc. are going to be represented. Consider for instance the following sentences:

(77) every man runs

(78) John runs slowly

If we adopt a Montague style analysis of determiners and adverbs and we try to provide a representation of (77) and (78) in a theory such as Bealer's, then there would appear to be two ways to go, namely:

(79) a. $\text{every}'([\text{man}'(x)]_x)([\text{run}'(x)]_x)$
 b. $\langle [\text{man}'(x)]_x, [\text{run}'(x)]_x \rangle \Delta$
 $[\text{every}'(x,y)]_{x,y}$

(80) a. $\text{slowly}'([\text{run}(x)]_x)(j)$
 b. $\langle j, [\text{run}(x)]_x \rangle \Delta$ $[\text{slowly}'(x,y)]_{x,y}$

In (79) and (80) every and slowly are treated as 2-place relations, on a par with other two place relations (such as, say, transitive verbs). But they are relations of a very different sort from, say, kick. In Chierchia (1982), it has been argued that a representation analogous to (79) and (80) is really all that is needed to handle deter-

miners, adverbs, and the like. However, such a line, as we have seen in section 3.4., puts a very heavy burden on syntax. For example in logics like Bealer's determiners, adverbs, etc. can be values of individual variables. Then to account for the fact that they are never nominalizable in natural language, either we have to look for some syntactic account, or we have to build some extra apparatus in the logic. In the setting of HST* the change that we need is minimal (the addition of 3rd order functional constants). It would appear that to get similar results in a first order setting we would have to make analogous changes (e.g. adding second order functional constants), which, however, would look much less natural within a first order theory.

A further problem has to do with the lack of predicate variables and quantification in Bealer's theory. Predicates can be values of variables only in their nominal form. So, according to Bealer, an argument like the one in (81) ought to be represented as in (82):

(81) Whatever John wants, Bill wants
John wants to jog
 Bill wants to jog

(82) $\forall x$ [want'(j,x) \rightarrow want'(b,x)]
want'(j,[jog(x)]_x)
 want (b,[jog(x)]_x)

On the other hand, Bealer recognizes that ultimately a simple sentence like John loves it should be represented

as $\text{love}'(j,x)$. By the same token we might expect that a sentence like John does it might be represented as $\text{do}'(j,x)$. Even if to do is a verb whose properties are very different from those of, say, to love, its verbal character is, nevertheless, beyond question. But consider the following argument:

(83) John does anything that Bill does
 John jogs
 Bill jogs

If we have predicate variables, (83) could be represented as follows:

(84) $\forall P[P(j) \rightarrow P(b)]$
 $\text{jog}'(j)$
 $\text{jog}'(b)$

However, in Bealer's system we have to resort to something like:

(85) $\forall x[j \Delta x \rightarrow b \Delta x]$
 $\text{jog}'(j)$
 $\text{jog}'(b)$

The inference in (85) is sanctioned by the axioms regulating the logical behavior of Δ , which make equivalences like the following valid:

(86) $\text{jog}'(x) \equiv j \Delta [\text{jog}'(x)]_x$

Now, the verb to do in English can act as a verbal pro-form in processes such as VP-deletion. Such a process is illustrated by the following sentence:

(87) John jogs and Bill does too

Among the various proposals that have been made on how to deal with VP-deletion, the best motivated appear to be those of Williams (1977) and Sag (1976). Simplifying a lot, the idea they put forward is that to do translates as a predicate variable, whose value is determined, roughly, by the context. So, in the same spirit of (84), (87) would be translated as follows:

(88) $jog(j) \ \& \ P(b)$

An interpretive procedure would ensure that the predicate variable P gets interpreted in the right way, i.e. in the case at hand as the predicate jogs. Not having predicate variables, (87) would have to be represented as:

(89) $jog'(j) \ \& \ b \ \Delta \ x$

Some interpretive procedure has to ensure, then that x in (89) is interpreted as $[jog(x)]_x$. However, (89) looks quite unnatural. The two conjoined sentences in (87) have exactly the same syntactic structure; but their logical forms in (89) look very different. Furthermore, Bealer underscores that $jog'(j)$ and $j \Delta [jog'(x)]_x$, though logically equivalent are different propositions. So, for instance one might believe one and disbelieve the other. This would lead us to expect that it is possible to believe (90a) and disbelieve (90b), for on the analysis of VP-deletion just sketched they would end up having very different logical forms (illustrated in (91a-b)).

- (90) a. John jogs and Bill jogs
 b. John jogs and Bill does too
- (91) a. jog'(j) & jog'(b)
 b. jog'(j) & b Δ [jog'(x)]_x

But I think that such an expectation is unwarranted. It may be possible for somebody to believe that London is pretty and disbelieve that Londres is pretty. But nobody (not even Pierre) might believe (90a) and disbelieve (90b) (or vice versa). So a theory with only individual variables seems to give counterintuitive results on the basis of the best grammatical analysis currently available of anaphoric phenomena involving VPs (such as VP-deletion). This argument is certainly far from conclusive but it does show that a theory that does not allow for variables over predicates as such (i.e. in predicate position) makes it harder to deal with cases of VP anaphora, which is an extremely widespread phenomenon cross-linguistically. I am sure that one could amend a first order theory of properties to cope with all the problems we have pointed out in the present section. What I am not sure about is whether what we will end up having, after all the necessary amendments, will look substantially different from the framework we are advocating.

5.5 Concluding remarks.

All the semantic frameworks considered in the present section are viable in principle as a basis for developing a semantic component in a theory of grammar. Each of them, I think, might constitute an interesting improvement of Montague's original theory. Especially interesting are recent developments of various techniques for building reflexive domains (i.e. domains containing properties that can be predicated of themselves) and strongly intensional approaches to properties. The beginning of a systematic consideration of some such theories (for metamathematical purposes) can be found in Feferman (1982). Here we have barely scratched the surface of the problem.

Investigations in type-free theories of properties can constitute a major breakthrough in formal semantics. However, as far as I can see, the second order framework of IL_* appears to be better designed as a semantic theory for natural languages than the alternatives we have been able to consider here. Such a framework allows for a very general treatment of nominalization that bypasses type theory, is not committed to weak intensionality, and does not force us to adopt a preconstituted model of what predication should be. This flexibility is certainly quite appealing, or at least so I would think, but our framework, though flexible in certain directions, is also

rather rigid in others. It makes strong claims on what sorts of things might be expected to be found as natural language meanings. Now all the claims embodied in the structure of IL_* might ultimately be falsified. But testing the data against the most restrictive grammatical theory available has proven to be a successful research strategy. We might hope that this will be the case also for the present attempt.

Of course, the arguments presented in this chapter are all quite general and abstract. As such, they provide only a very preliminary motivation for adopting IL_* as a semantic theory for natural languages. We have not built yet a grammar of English. The following chapters will be devoted to the actual construction of a formal grammar for a rather rich fragment of English. Such a formal grammar will be primarily intended to provide a linguistic analysis of certain kind of constructions involving infinitives and gerunds. It is on the basis of our success or failure in providing a convincing account of those constructions that the semantic framework adopted here should be ultimately evaluated.

FOOTNOTES

¹We are disregarding here intensions. As the reader will be able to see later on, considering intensions will not effect our main point in any way.

²This section contains some material from Chierchia (1982).

³Link (1983) has proposed to introduce groups (or "plural individuals") as primitives in the domain of individuals. He then imposes on groups an algebraic structure isomorphic to the algebra of sets. Under such a proposal, the problem in question wouldn't arise.

⁴Carlson (1977) considers this possibility himself, but rejects it for the very problem we are pointing at. He introduces kinds as primitives and relates them to the corresponding CN-intension by means of meaning postulates. This move, however, appears to be rather ad hoc. See Chierchia (1982, 1983) for some discussion.

⁵See Dowty (1978) for a detailed analysis of argument drop and related phenomena within MG.

⁶A similar point was made by Bach (1982) in connection with the analysis of purpose clauses.

⁷See Cocchiarella (1974, 1978, 1979, 1983).

⁸See Partee (1979b) for discussion of the issues involved.

⁹This not so for the version of HST* developed in Cocchiarella (1979). It is however the case for the version of the system in Cocchiarella (1983).

¹⁰Such a technique was originally developed by Quine (1958).

¹¹The identity sign in HST* can be defined in terms of indiscernibility, i.e. $\mu_i = \mu_j \leftrightarrow_{df} \forall \beta [\beta(\mu_i) \leftrightarrow \beta(\mu_j)]$.

¹²This was the line followed in Chierchia (1982).

¹³ $\lambda \alpha$ HST* has been developed in Cocchiarella (1983).

¹⁴This restriction excludes things like $\varphi(x)$, where φ is a wff.

15We take $ND = D \dots D$ n-times

16In Chierchia (1982) adverbs, prepositions, etc. were regarded as being of the same logical types as predicative expressions such as CNs, IVs, etc.

17F.R. Higgins points out to me the existence of nominalizations such as 'downness' or 'inness'. Nominalizations of this kind seem to be more frequent in languages such as German, Greek or Latin. However, they are probably to be analyzed, on independent grounds, as nominalizations of prepositional phrases (rather than of prepositions). "Downness" means something like "being down something".

18We assume that $\langle e_0, p \rangle = p$.

19This is essentially what Thomason (1980) proposes. It should be emphasized, however, that nothing in IL^* as such forces us to this particular choice of primitives.

20See e.g. Barwise and Cooper (1981), Van Benthem (forthcoming).

21See on this Keenan and Faltz (1978), Gazdar (1980), Partee and Rooth (1983), Cooper (1983).

22This was originally suggested to me by Barbara Partee.

23We are ignoring intensions here; to accomodate them within Parsons' system is unproblematic, and does not change the substance of the points we will raise.

24Notice that π 's are functions from types into types and g 's are functions from types into words.

25 '+' is taken to be an order preserving analogous of set-theoretical union. See Turner (1983) for details.

26NFU* stands for 'New Foundations with Urelements' a system developed by Jensen (1968) of which NFU* is a representation (see on this Cocchiarella, 1979).

27Actually, the Fregean embedding f in Cocchiarella's system is a function from $U \cup X$ into U which is the identity map when restricted to U . But this is done only for technical convenience and nothing hinges upon it.

C H A P T E R I I
L O C A L L Y C O N F I G U R A T I O N A L G R A M M A R

We are now going to consider some of the general assumptions concerning the organization of a grammar that we will be making. This will provide us with an implementation of the semantic framework developed in Ch. I within an (empirical) theory of language. What we are going to sketch is the result of extensions of (and constraining principles on) the framework originally developed by Montague (especially in UG and PTQ).

In particular, we will outline the basic format of syntactic theory, of the syntax-semantics map and of morphological theory as they will be needed for more specific investigations of English predicative expressions in subsequent chapters. We will also try to point out at least some of the general consequences of the way of doing things that we want to endorse. Occasionally we will draw some comparisons between our approach and other current theories.

1. The Notion of Function-Argument Structure.

Following the tradition originated by Montague's work, we take a fundamental notion of grammatical theory to be that of function-argument structure. A speaker exposed to a set of linguistic facts (such as those in (1a)-(4a)) analyzes them into functions and arguments nested in a certain way (let us say, as in (1b)-(4b)).

- (1) a. John runs
b. run'(John')
- (2) a. John kisses Mary
b. Kiss'(Mary')(John')
- (3) a. John runs slowly
b. (slowly'(run'))(John')
- (4) a. John tries to run
b. tries'(run')(John')

It should be noted that from a mathematical point of view there are an indefinite number of different function-argument structures that could be associated with (1)-(4). To decide how a given sentence is organized in a particular nesting of functions and arguments is an empirical issue. In the following discussion we will provide a number of criteria for the analysis of sentences into functions and arguments. The assumption that function-argument structure is fundamental is not peculiar to MG (or to closely related frameworks, such as GPSG) but is clearly a feature that other frameworks share with it, such as LFG or, in some way, GB. What is peculiar to MG is the way such a notion is characterized. It is probably fair to say that one of the basic tenets within MG

(inherited from the tradition of mathematical logic) is that functions in general are not just pieces of notation, configurations of symbols or procedures for manipulating configurations of symbols. In particular, the functions involved in (1b)-(4b) are probably not entities of this kind, to the extent that they are taken to be ways in which meanings are organized. Configurations of symbols (and procedure to manipulate them) have meanings, and hence cannot be meanings. This is where truth conditions (with respect to a model) come into consideration. The functions involved in (1b)-(4b) are model-theoretically characterizable entities that make up the pieces of a definition of the conditions under which a sentence is true. It is clear, however, that we need a theory of the basic entities (i.e. of the types of functions and arguments) that come into play in the specification of truth conditions for natural languages. Such a theory, among other things, should provide a way to refer to those entities and a way to characterize their logical behavior (i.e. their behavior in inference patterns of various sorts).

The theory of semantic entities and of function-argument structure that we assume is IL_* . Such a theory, as we know from chapter I, provides an axiomatic characterization of some of the fundamental entities needed for the semantic interpretation of natural languages, such as

property, proposition or predication. By doing so, it also sets up the space within which subtheories of other, more specific semantic objects (e.g. events, kinds, action-types, etc.) can be developed. Investigations of axiomatic systems have highlighted the usefulness of matching such systems with model-theoretic interpretations which, among others, allow us to test those systems for consistency and to provide senses in which those systems can be regarded as complete.

The function-argument structure that a speaker associates with a linguistic expression will be given as a map (a translation procedure) from a natural language into IL_* . We will call IL_* (i.e. the language of the translation) 'logical form' or 'semantic structure', and the (reduced) translation of an expression α into IL_* 'the logical form of α '.¹ Since IL_* is model-theoretically interpreted (or interpretable) a translation of English into IL_* will automatically induce a definition of the conditions under which a sentence of English is true. Function-argument structure should provide the (putatively) universal form in which meaning is organized, including a specification of the way in which the meaning of a complex is related to the meaning of its components. We assume that this specification is constrained by two factors. The first factor is what constitutes a possible semantic operation. Our semantic framework provides us

with essentially two basic semantic operations: functional application and λ -abstraction (binding). We will also allow for a limited number of operations that can be defined in terms of the basic ones, such as, in particular, function composition.² The second factor that constrains a specification of the way in which the meanings of a complex is related to the meaning of its components is what Bach calls the rule-by-rule hypothesis. The basic idea underlying it is that for any syntactic rule that combines expressions there is exactly one semantic rule that combines their meanings. I will not try to justify here why it is desirable for grammars to be compositional in exactly this sense,³ except for pointing out at various moments some non-trivial effects of this principle. Other aspects of the syntax-semantics map will be discussed in section 3.

There is an interesting debate among formal semanticists concerning the potential dispensability of logical form. Those who think that logical form is not a significant level of linguistic theory argue that semantics should provide a map from expressions into their non linguistic denotata. Models are abstract structures that help us to get to those non linguistic denotata. Nothing else is needed but a way of referring to them. This may be simplistic but I hope not too misleading. See Cooper (1975) for an interesting defense of this view.

However, there is a way in which the question might be somewhat ill-founded. We do need a way of referring to semantic entities. In fact, if we do semantics via a direct map into a model (as e.g. in Montague's EFL or Cooper (1975)) what we end up with is a map from English into 'ways of referring', often isomorphic to some known logical language (e.g. the typed λ -calculus, to pick one randomly). This is not surprising since one of the reasons logical languages are designed is to have devices that reflect closely (or overtly display) aspects of the "structure of the world" relevant for some purpose. So, in this possibly trivial sense, we need a language to refer to semantic entities. Now such a language might be regarded as a mere convenience, with no other significance but the role it plays in the theory. But can't any theoretical construct within a theory be so regarded? The concepts of mass or electron might be mere conveniences, or so it has been claimed.⁴

It seems to me that claims concerning the dispensability of logical form should be interpreted as a rejection of the view that the syntax of logical form is a linguistically significant level in a theory of grammar. This is evidently a substantive issue that should be eventually decidable on empirical grounds (at least to the extent that any abstract principle within an empirical discipline is). Let me first try to illustrate this point

with an example. Then I will try to sketch the position I will be taking on this issue throughout the present work.

We have claimed that function-argument structure is primarily a semantic (model-theoretic) notion. However in the examples (1)-(4) we have not displayed model-theoretic entities but something like the translation from English into a logical language. The model theoretic entities corresponding to (1b)-(4b) are propositions, which in our "intended" model are primitives. But if propositions are primitives, then how can we recover from them the nesting of functions and arguments represented in (1b)-(4b)? And if we can't recover it, how can we claim that function-argument structure is purely a semantic notion? It would seem that such a notion can be characterized only by crucially appealing to the syntax of our logical form.

We could try to classify into two major groups the possible ways of reacting to this situation. One way to go would be to assume that propositions have an internal structure and that the relevant notion of function-argument structure is to be defined in terms of it. In this direction go recent attempts such as those of Barwise and Perry (1982) or Landmann (1982). One of the reasons we take the notion of proposition as a primitive is because we want to stay open for the time being with respect to exactly how much structure propositions should have. A second way to go is to individuate canonical forms in some

logical language as capable of overtly displaying the relevant characteristics of function-argument structure. For instance one might take the (reduced) translation into IL as representatives of the function argument structure associated with a sentence. Also Kamp's theory of discourse representations (Kamp 1981) seems to be in the spirit of this second alternative. If the first kind of answer is the right one, then the syntax of logical form might not be a crucial level of linguistic theory. If the second kind of answer is right, there are aspects of the syntax of logical form that must be linguistically significant. We will not try to settle this issue here, but we will point out occasionally questions that bear on it. We will therefore use 'semantic structure' or 'logical form' as a sort of neutral way of referring to the theoretical level that defines function-argument structure. In our case this is IL_* , i.e. a model theoretically interpretable axiomatic theory.

Now, there is one aspect that the two approaches to logical form just discussed seem to have in common. This is the idea (built originally by Montague into his UG) that the definition of semantic structure does not employ any notion or principle that operates in the syntactic component of a grammar. We may dub this the thesis of autonomy of semantics. Let me try to exemplify a little. Consider the notion of NP. NP is the label of a syntactic

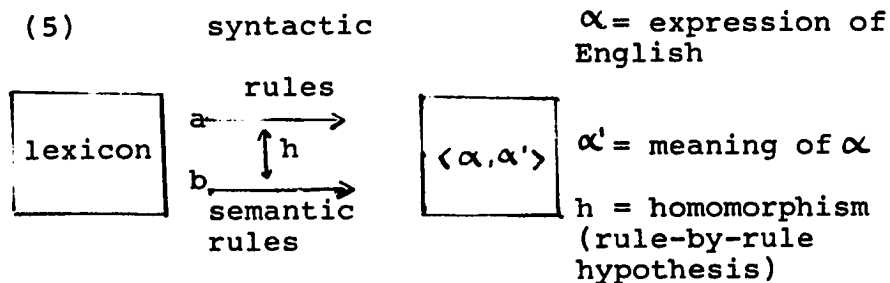
category (Montague's T). NPs may be systematically related to entities of a certain sort, let us say, for the sake of discussion, individuals. I do not believe, however, that we can decide whether something is an NP just by inspecting the semantic entity associated with it (granted that we could do so). We have rather to consider the syntactic distribution of such a linguistic entity, the syntactic processes it enters, and so on. Whether a given construction is an NP is, I would think, a typical question of "autonomous syntax". And syntactic notions, such as NP-hood, though systematically related to semantic ones, do not enter as such in defining semantic structure.

Consider next the notion of subject, and suppose it can be defined by employing some mixture of semantic and syntactic notions (say, as the last NP argument of a verb -- see Dowty (1982a) for discussion). Then autonomy of semantics would require not to use subject in defining semantic structure. Of course, this does not imply that there should not be relevant notions which have to be defined using terms from the theoretical vocabulary of both syntax and semantics. It does mean, however, that any such notion cannot be used in defining semantic structure.

Consider finally principles like subjacency. Subjacency has been proposed as a condition on rules (or, possibly, on representations) and says roughly that no

rule can span across two bounding nodes (let us say NP and S in English). Syntactic dependencies such as wh-phenomena are said to obey subadjacency. Then, the thesis of autonomy of semantics entails that no semantic rule (e.g. no rule involving λ -abstraction, in terms of which binding is defined) should obey anything like subadjacency.⁵

Analogous examples could be multiplied, but these should suffice to illustrate the type of consequences that descend from the thesis of autonomy of semantics and the fact they are not trivial. We could schematize this discussion as follows:



The diagram illustrates that syntactic and semantic rules generate in parallel English expressions with their associated semantic structure. Each of the two arrows (a and b) obey autonomous, though related, principles in terms of which syntactic and semantic generalizations are cast. There need not be principles that apply across-the-board to (a) and (b).

Many current theories make opposite claims. For instance, within GB it is assumed that there are many

levels of grammatical representation, and certain rules or principles (such as move α or the Empty Category Principle) do apply across all such levels. If this were the case, then the thesis of autonomy of semantics would be wrong.

Things are less clear cut with respect to what in LFG is called 'functional structure'. This structure contains primarily information about the nesting of functions and arguments, which we take to be the heart of semantic structure. However, Bresnan's functional structure does not represent quantifier scope and contains various pieces of syntactic and morphological information. Also, within functional structure crucial reference is made to grammatical relations. So, on the model of grammar we are assuming, a level of representation like Bresnan's functional structure would be disallowed.

To sum up, the framework we adopt contemplates devices for building directly expressions of English together with their semantic structure. One of the central aspects of semantic structure is the notion of function-argument structure. Such a structure is the way in which universal meanings come together, and may be encoded in various ways across different languages. We leave it undecided whether function-argument structure can be regarded as a purely model theoretical construction or should rather be defined on the basis of an intermediate

(semantically motivated) level of representation, such as the language of the translation.

2. The Format of Syntax.

We will adopt as a syntactic framework a version of categorial grammar close to the one developed by Montague in UG. Many of the points we will make would hold also within closely related approaches to syntax, such as, in particular, GPSG. We will point out in the course of subsequent discussion theoretical and empirical items where, in principle, a GPSG and a categorial approach might diverge.

Montague's framework has been elaborated and constrained by many authors. Some of the work whose influence will be most directly reflected here is represented by Cooper (1983), Dowty (1978, 1982a), Partee (1979a) and especially Bach (1980a, 1983).

Sets of expressions are taken to be labelled by categorial indices which encode information about the combinatorial power of expressions. Category-labels are defined in terms of a very small set of basic labels by the following schema:

- (6) if A, B are category labels, then so are A/B and $A//B$.

We take B_A to be the set of lexical entries of category A and P_A the set of phrasal expressions of category A .

We assume that for any category A , $B_A \subseteq P_A$. We assume, moreover, that each category label may carry further feature specifications. For illustrative purposes, we will adopt the following (simplified) categorial base for English. The basic categories are going to be NP (noun phrases, corresponding to Montague's T), S (sentences, corresponding to Montague's t), CN (common nouns: student, horse,...), and Pred_A , where $A = \text{PP}$ (predicative prepositional phrases: there, on the roof,...), or $A = \text{ADJ}$ (predicative adjectives; red, drunk, good,...). Some of the derived categories that we will use are:

- (7) IV = S/NP: run, walk,..
 TV = IV/NP: kiss, hit, seek,..
 ADV = IV//IV: slowly, easily
 ADV_S = S//S: necessarily, possibly
 DET = NP//CN = the, a, every,..
 Prep = $\text{Pred}_{\text{pp}}//\text{NP}$ = in, with,...
 IV/S = believe, say,..
 IV/IV = try, wish,..
 (IV/IV)/NP = promise,..
 TV/IV = force, persuade,..

The use of the single vs. double slash notation will be explained in section 3.1. For each category A , given the set B_A , the set P_A is recursively defined in terms of a set of rules of the form:

- (8) $\langle F_\gamma, A, B, C \rangle$ where: F_γ is a (binary syntactic operation
 A and B are the input categories
 C is the output category

The schema in (8) is to be interpreted as saying that if $\alpha \in P_A$ and $\beta \in P_B$, then $F_\gamma(\alpha, \beta) \in P_C$. I will assume

that in general a syntactic operation F_{γ} is the composition of two functions: a bracketing function L_{γ} and a combinatorial function O_{γ} . Each of these components of F_{γ} is selected from a small universal set. The role of L_{γ} is to determine the (surface) labelled bracketing of the output of each syntactic rule. The role of O_{γ} is to determine the way expressions are combined. The core combinatorial function is represented by the following template:

AFFIX α : affix α after the n^{th} constituent of β .

The level of constituency to which AFFIX α has access must be further constrained and is also subject to parametric variation. By setting $n = \text{last}$, we get right concatenation (RC), by setting $n = 1$ (and constituent = word) we get what Bach calls right wrap (RW), by setting $n = 0$, we get left concatenation (LC). Syntactic theory has to determine in general how the parameters n and constituent in AFFIX α may covary with a particular choice of the categorial base. There are further general conditions on rules concerning feature matching and percolation to which we shall return in section 4.

It is worth noting already at this point how the rule-by-rule hypothesis together with the very small set of semantic operations that we allow contributes to narrowing down considerably what may constitute a possible

syntactic rule. Given that, in the unmarked case, we allow just three semantic operations (function application, function composition and λ -abstraction), the rule-by-rule hypothesis forces there to be exactly three corresponding rule types in the syntax. Two of them will be of the following form:

- (9) Functional application (FA):
 a. $\langle F_{\gamma}, A/B, B, A \rangle$
 b. Corresponding translation rule: if $\alpha \in P_{A/B}$ and $\beta \in P_B$, then $F_{\gamma}(\alpha, \beta)$ translates as $\alpha'(\beta')$.
- (10) Function composition (FC)
 a. $\langle F_{\gamma}, A/B, B/C, A/C \rangle$
 b. Corresponding translation rule: if $\alpha \in P_{A/B}$ and $\beta \in P_{B/C}$, then $F_{\gamma}(\alpha, \beta)$ translates as $\alpha' \circ \beta'$.

As for binding (λ -abstraction) we will adopt essentially the storage mechanisms proposed by Cooper (1975). For instance, quantifying in will be dealt with by a pair of rules along the following lines. First, each NP when combined syntactically with some function, say a TV can be optionally stored. What this means is that in the semantics, the TV-meaning is combined with a variable whose index is put in a store along with the NP-meaning. Then there is an S-level rule which allows us to λ -abstract over the stored variable and to quantify in the corresponding NP. This is exemplified in what follows:

(11) Syntax	Semantics	Q-store
love every man [store-in NP]	love'(x ₃)	\langle every man', 3 \rangle

Mary loves every man [subject-verb rule]	love'(x ₃)(m)	<every man', 3>
Mary loves every man [store-out NP]	every man' (λx_3 [love'(x ₃)(m)])	∅

We may regard binding rules like this one illustrated in (11) as a sort of discontinuous rule made up of two necessarily related parts (store-in and store-out). Phenomena such as wh-movement are assumed to be treated in a related fashion. See Cooper (1983) and Engdahl (1983) for discussion.

All this is taken to be implemented in the following way. We assume that a grammar is made up of a set of rules and a set of entries. Each entry is an n-tuple specifying at least the following information: a morphophonemic representation of the word, a categorial index, a translation into IL_{*}, a set of syntactic features, a set of presuppositions, a quantifier store. Each grammatical rule is a pair of a syntactic and a translation rule, subject to general conditions on features. Rules map entries of the grammar into new (phrasal) entries, thereby providing a simultaneous recursive definition of what constitutes a well-formed phrase of each category and what its associated meaning is.

It is worth emphasizing that on the present view, rules of particular grammars are a function of two simple

variables: (a) the categorial base of a language, (b) the language particular values of the free parameters in AFFIX α . If we assume, as I think is plausible, that both (a) and (b) range over fairly small, finitely specifiable domains⁶, then we can conclude that modulo a limited number of choices, the rule-by-rule hypothesis predicts that language particular rules can be read off the lexicon of that language. Once the lexical category of an item is established, the grammar will automatically specify how to fill in the universal schemata in (9), (10) and (11) with respect to that category. Rules that cannot be obtained by this automatic procedure have to be added as special 'housekeeping' rules. This is of course very sketchy and, as things stand now, largely programmatic.

Now, Chomsky (1982, pp. 29 ff.) points out that in the standard theory of Chomsky (1965) there is a certain unavoidable redundancy between subcategorization frames and phrase structure rules.⁷ He proposes the projection principle as a way of eliminating this redundancy. Roughly, that principle requires that the subcategorization features of an item should be met at any level of "syntactic" representation. So, given a general format for PS-rules (say, \bar{X} -theory) the projection principle determines what language particular rules are going to be in terms of the lexicon (modulo some choices concerning the ordering of constituents, as in our case). It is

interesting to note that on the present theory this particular role of the projection principle is subsumed by (or can be derived in terms of) the rule-by-rule hypothesis.

In agreement with the general spirit of Chomsky (1982), I assume that the exact format and function of each piece of information represented in a grammatical entry is determined by the interaction of a set of relatively simple subtheories, such as a theory of government and agreement, a theory of binding, a theory of control. Aspects of some such theories will be discussed below, as they turn out to be relevant for the main topic of the present work.

One of the central ideas that the present variety of categorial grammar is meant to formalize is that the combinatorial power of an expression is somehow a reflex of semantic function-argument structure. Categories are labels of sets of expressions; expressions themselves are not functions (i.e. are not sets of ordered pairs); they are semantically associated with functions. So, speaking of an expression as a function is to be understood as a derivative use of the term 'function'. The only genuinely syntactic functions are the family of operations whose format is represented by AFFIX α . These operations are responsible for phrase structural configurations. Hence on the present view, phrase structural configurations are

a surface epiphenomenon ('phenogrammatcs' to use Curry's terminology -- see Dowty 1982) determined by language particular choices concerning the parameters made available by the general template AFFIX α . Syntactic rules, on the other hand, are determined by semantic function argument structure: they give the general form in which semantic function argument structure is encoded (via syntactic operations, i.e. AFFIX α) into language particular phrase structural configurations (which constitutes what Curry called 'tektogrammatcs'). It follows, then, that any attempt to provide universal characterizations of grammatical processes (and relations) must ultimately take the form of a parametrized function on semantic argument structure. It also follows that any configurational characteristic of tree-geometry must be an observable characteristic of surface structure (and hence, it will turn out to be language particular). In the unmarked case, "abstract" syntactic configurations are not expected to arise on the present view of grammatical theory.

The explicit formulation of the central character of semantic (model-theoretic) function argument structure in grammatical theory was formulated for the first time (to the best of my knowledge) in Keenan (1974). Since then it has been further elaborated by a number of people.⁸

Essentially, the same general point can be made from a slightly different perspective. An interesting evaluation metrics on grammars is constituted by Partee's (1979a) well-formedness constraint. According to it, a grammar should ideally provide a way of directly generating well-formed surface structures. The closer a particular grammar comes to this ideal, the more high it is valued by universal grammar. Principles as general as this can be evaluated only on the basis of overall theory-comparisons, not on the basis of their intrinsic plausibility or empirical support. As a general methodological principle the well-formedness constraint can be regarded as a version of Occam's razor. As a guiding principle of linguistic metatheory, it can be regarded as a very strong locality requirement: syntactic structure has to be determined as soon as possible and as locally as possible; once it has been determined, it cannot be altered. It follows that transformations, restructurings, indexings, filters, phonetically unrealized items are to be regarded as marked options (within the limits given by the attainability of descriptive and explanatory adequacy). Seeing how far one can get with these assumptions I would think is per se an interesting enterprise.

3. The Syntax-Semantics Map.

We have already discussed some general effects of the rule-by-rule hypothesis, which is one of the fundamental constraints on the syntax-semantics map that we are assuming. In the present section we are going to consider other principles that govern the relation between syntax and semantics.

3.1. Categories vs. types: the functional correpondence constraint.

An important claim that Montague's semantics makes is that semantic categories (i.e. logical types) associated with expressions are a simple function of their syntactic categories. This idea is interesting for several reasons. First, it narrows down enormously the class of possible syntax-semantics maps, and it has far reaching empirical consequences; some of them were discussed in chapter I. Second, we know that syntactic categories are recursively built from a small set of primitives. The type assignment function will then be able to simply parallel the recursive schema of category formation. This means that all that has to be learned concerning a language particular type assignment (call it τ) is a specification of its values for the basic categories, the recursive part of τ being modelled on the universal categorial recursion.

So, Montague's constraint on the category-type correspondence represents a substantive hypothesis on the way meaning is related to syntactic structure; in ch. I, we dubbed this the transparency principle. Now one of the points made in chapter I was that given Montague's type theory, the transparency principle cannot be maintained. It was also argued that possible ways of weakening this principle yield undesirable results. In the present section we want to argue that the framework provided by IL_* allows us not only to maintain the transparency principle, but also to strengthen it considerably.

I am aware of three arguments against what we call here the transparency principle. One is due to Williams (forthcoming). A second is due to Rooth and Partee (1983). We will consider them both in section 3.2. The third one has to do with the analysis of nominalization phenomena, and we have considered it at length in chapter I. There we have tried to show that this third kind of problem simply does not arise in the setting of IL_* . For instance, we can let IVs be 1-place propositional functions (i.e. $\tau(IV) = \langle e, p \rangle$); then the particle to (or, maybe the infinitival marking) maps properties into their individual correlates, as we will see in more detail in the next chapter. However, it should be pretty clear that as far as nominalization phenomena go, IL_* allows us to maintain the transparency principle. Further

considerations suggest that we can make this principle even stronger, in that the number of choices in lexical type assignments can be narrowed down essentially to one.

In IL_* we have fewer types than in IL . Essentially, we have three logical types: individuals, 2nd order propositional functions, and 3rd order functors. Strictly speaking, this is not quite correct. Propositional functions in IL_* are also typed with respect to their adicity. Functors are typed as to their adicity and as to whether each argument is an n -place propositional function or an individual. So, even though the space of logical types associated with IL_* is certainly more limited than the one associated with IL , still strictly speaking we have an infinite set of types and hence an infinite set of possible type assignments to basic categories. Of course, we might easily cut down to a reasonable size the number of possible type assignments by relying on "well known facts", such as, for example, the very small adicity (possibly ≤ 3) of verbs. The problem is whether we can find a principled way of carrying this on, rather than simply having to state a collection of facts.

One of the central ideas built into the version of categorial grammar that we adopt is that syntactic combinatorial capacities of phrases are a reflex of semantic function-argument structure. A very natural way

of constraining and making clear the notion of being a reflex of employed in the latter statement would be for example assuming that something is syntactically a function (i.e. belongs to a category of the form A/B) iff it is semantically a function. It is easy to conceive type assignments that do not meet this constraint. For instance, Bennett's (1976) modification of Montague's type assignment does not, since CN is taken as a primitive (i.e. non functional) category and yet mapped into a functional type (namely $\langle e, t \rangle$). We can state therefore the following general principle:

- (12) Functional correspondence: a functional type must correspond to a functional category; a functional category must correspond to a functional type.

This principle is extremely strong, and within our framework certainly very natural. Let us assume its truth and explore some of its consequences. A first immediate consequence of it is that a lexical type assignment to basic categories must be a finite function. It is easy to see why. Basic categories are, by definition, non-functional. Hence, by functional correspondence, they must be mapped into non functional logical types. But in IL_* there are only two of those: e, the type of (logical) individuals and p the type of propositions. So, each of the basic categories must be mapped either into e or into p. Obviously, category S must be mapped into p. Strictly

speaking, members of p are o -place propositional func-
tions. There is a clear sense, though, in which o -place
propositional functions are not really functions: they
lack the most typical feature of functions, namely the
capability of taking arguments. So propositions are sort
of intermediate creatures. With individuals, they share
the property of being non argument-taking. With functions
they share the property of being nominalizable (i.e.
projected as individuals). From now on, let us use the
names "state of affairs" or "eventuality" to refer to
nominalized propositions. States of affairs are entities
in the world; propositions are not. Propositions are the
values of declarative sentences. States of affairs will
turn out to be, among other things, the values of
that-clauses.

If we ignore, for the time being, quantification, NPs
are the prototypical arguments of propositional func-
tions. So, given that propositional functions have the
form $\langle e_n, p \rangle$ (cf. ch 1, sec. 3.4 (56)), NPs will have as
their logical type e . So, if a language has NP and S
among its basic categories, we have essentially no choice
as to what their corresponding types must be. So, we can
hypothesize that the following statements are part of
universal grammar:

$$(13) \quad \begin{array}{l} \tau(\text{NP}) = e \\ \tau(\text{S}) = p \end{array}$$

The problem, then, is what might be the type of other hypothetical primitive categories. Evidently, it won't be p. Something associated with p can't take arguments and can be nominalized. What but Ss can have these characteristics? So, we are left only with e. All primitive categories except S must be associated with e. The type assignment in (13) can therefore be changed to:

$$(14) \quad \begin{array}{l} \tau(S) = p \\ \text{for all primitive categories } \alpha, \text{ where } \alpha \neq S: \\ \tau(\alpha) = e \end{array}$$

But how is that possible? Among the basic categories we assume there are CN and Pred which (as suggested by the name itself) gather a number of items (CNs, ADJs, PPs) that are associated with predicates of individuals. So, how can the logical type of Preds be e as (14) requires? In what follows I would like to suggest that Preds being individuals might explain some of the ways in which they differ from verbs in a language like English.

The point is that Preds too (like Ss) are somehow intermediate creatures between functions and arguments. They all can act as predicates but, as it were, not by themselves. Similarly to verbs, they can assign case to their subjects, but only with the help of the copula be. The copula does not seem to contribute in any fundamental way to the meaning of the items it operates on, apart from being the vehicle of tense, aspect, possibly case-marking. So, common nouns have the potential of being

predicated; but cannot be predicated of anything directly, in the way verbs can. In the framework of IL_* there is a fairly natural way of representing this situation. Properties have individual correlates. So, we might expect to find syntactically basic items that are associated with properties, as well as syntactically basic items that are associated primarily with individual correlates of properties. In English, verbs are items of the first kind; common nouns, adjectives and other predicables items of the second kind. Items associated with properties will have the power of being nominalized via the embedding function ' \wedge '; items associated with individual projections of properties will have the power of being predicativized via the inverse of ' \wedge ', namely ' \vee '. Just as the particle to might be regarded as a reflex of the operator ' \wedge ', the copula might be a reflex of the operator ' \vee '.

This idea can be implemented by assuming that the logical type of Pred is e, viz. the type of individuals. But of course, Preds will have to be special sorts of individuals, namely predicable ones. In section 3.4. we will see how to guarantee this formally. The point here is that the fact that Pred (being a primitive category) must be associated with e is precisely what we were forced to by the functional correspondence principle.⁹

A little bit of terminology will come in handy: let us call kinds the individuals associated with CNs, qualities the individuals associated with ADJS, locations the individuals associated with locative PPs, periods the individuals associated with temporal PPs. Given the double nature of properties, we might expect the way lexical types of property-denoting items are encoded in a grammar to be subject to parametric variation crosslinguistically. So, for instance, there should be languages where CNs and ADJs can be predicated directly of NPs (and hence can assign case directly to them), as well as languages where verbs cannot be predicated directly but need some special operator for that to be possible. And indeed both cases seem to be attested. Russian might be an example of a language of the first kind; languages where the difference between CNs and Vs is said to be fuzzy (such as Nootka) might be possibly regarded as languages of the second kind.¹⁰

So, the type assignment in (14) (given the framework of IL_*) seems able to shed some light on the behavior of some predicative expressions in natural languages. They jointly predict the existence of two sorts of predicable entities: some able to take arguments directly, others capable of taking arguments only with the help of the predicator 'v'. Now such a strange situation really seems to be a peculiarity of natural languages, one that

is not shared by any other known communication system. The fact that IL_* and the functional correspondence principle (of which the type assignment in (14) is a consequence) allow us to begin to make sense of such a situation seems to provide some support for both.

The functional correspondence principle, however, yields a further bonus. If we can provide a recursive type assignment for complex categories similar to Montague's (and we will be able to do so), it follows that type assignments are fully determined by universal grammar, viz. they don't have to be learned. What this means is that a language particular type assignment (similarly to language particular rules) can be read off the categorial structure of items (modulo a finite set of choices in the categorial basis). All the language acquisition device has to do is to determine the lexical category of an item. Then universal grammar supplies it with that structural aspect of meaning which is captured by its logical type.

In order to see how the recursive part of the type assignment function is going to work, we may recall that there is a certain hypothesis about possible semantic universes built into IL_* , namely the three layers hypothesis. A possible way of looking at this hypothesis is the following. Syntactic categories fall into three natural classes, determined by the semantic type asso-

ciated with them: arguments, predicates and functors. This means that there are grammatical processes that involve exactly these three classes of categories, or boolean combinations of them. The categorial notation we adopt is designed to represent this. The derived categories will be of two sorts: some will be associated with propositional functions, some with functors. We use the double slash notation to mark this difference. An item of category A/B will be associated with a propositional function; an item of category A//B with a functor. The double slash can be regarded as a semantically motivated feature (something like +functor). On this basis, the type assignment to complex categories can be simply given as follows:

$$(15) \quad \begin{aligned} \tau(A/B) &= \langle e, \tau(A) \rangle \\ \tau(A//B) &= \langle \tau(B), \tau(A) \rangle \end{aligned}$$

A few examples of how this type assignment will work for our sample categorial system might be useful:

$$(16) \quad \begin{aligned} (TV) &= \langle e, \tau(IV) \rangle = \langle e, \langle e, p \rangle \rangle \\ (IV/IV) &= \langle e, \langle e, p \rangle \rangle \\ (IV/S) &= \langle e, \langle e, p \rangle \rangle \\ (IV//IV) &= \langle \tau(IV), \tau(IV) \rangle = \langle \langle e, p \rangle, \langle e, p \rangle \rangle \\ (DET) &= \langle \tau(\text{Pred}_{CN}), \tau(NP) \rangle = \langle e, e \rangle \\ (Prep) &= \langle \tau(NP), \tau(\text{Pred}_{pp}) \rangle = \langle e, e \rangle \end{aligned}$$

Transitive verbs are going to be two-place propositional functions. The logical type of try- and believe-type verbs is going to be the same. Of course, believe and try are going to be different sorts of two-place propositional

functions, as sanctioned by their different categorial status. On the other hand, the logical type of IV-adverbs, which in PTQ was the same as try-type verbs, here comes out to be the type of functors that map properties into properties. At first sight, the logical type associated with determiners and prepositions might be somewhat shocking. They are viewed as functors that map individuals into individuals. This can make sense only because we have a domain of basic individuals which has got much more structure to it than Montague's D_e . The case of prepositions can be made quite intuitive, at least in some cases. A preposition like under would map an entity such as the table in front of me into a location, the location under the table. Locations are predicable individuals, that is they are the correlates of propositional functions. So, with the help of the copula they can be turned into propositional functions in sentences like John is under the table. Within the present framework, the logical form of the latter sentence would be something like:

(17) 'under'(the table')(j)

To see how determiners can be cast as functions from individuals into individuals we must discuss in greater details the way quantification and quantified NPs are going to be treated, which will be the task of section 3.3.

To sum up, we have proposed two constraints on the relation between types and categories: Montague's transparency principle and the functional correspondence principle. The interaction between the two (within the framework of IL_*) yields two main results. First, it sheds some light on the behavior of non verbal predicative expressions (in languages like English). Second, it narrows down possible category-type mappings to essentially one, which can be taken, therefore, as universally given. It might be that subsequent considerations will induce us to weaken the functional correspondence principle. But the results mentioned above can be achieved by regarding such a principle as a constraint on lexical type assignments. This would mean that types associated with phrasal categories might depart in some limited way from types associated with lexical categories.¹¹ It is not implausible that in this form functional correspondence will hold against a range of data wider than the one considered here.

3.2. Some sortal distinctions.

We have seen in chapter I that there is a trade off between what aspects of meaning should be built into type structure and what aspects should be treated via a sortal articulation of the domain of individuals. Here, we are

going to consider some of the sortal distinctions that grammars seem to make.

We are going to assume that there exists a subset B of the domain U of logical individuals, called the base, containing all the entities we "usually" talk about. B will play a role similar to that of the set D_e of basic entities in Montague's system, but there will be more structure to it. Besides "ordinary" things (i.e. entities such as you, Pavarotti, my typewriter, etc.) B will contain entities like groups, kinds, qualities, locations, periods, eventualities, possibly stages, and more. It will contain the individual correlates of the properties associated with VPs. We shall call the denotata of nominalized VPs "conditions".¹² So, for example, to run, to love John, to be there are all conditions, and are going to be in B . We will use e_B to refer to the type of individuals of sort B . As a rule of thumb, we may suggest the following: a member of the base is any entity of which a basic VP (i.e. run, bother, be fun, etc.) can be predicated.

Now, given the set B , it is possible to define properties of properties of entities of type e_B . Such properties can be used to represent quantifiers over the set B . Among them, there will be the values of quantified NPs such as every man or some man, as Montague has shown. So, every man is a property that the property of running

has just in case every man runs. On this view, the property associated with every man could be represented as follows:

$$(18) \quad \lambda x \forall y [\text{man}(y) \rightarrow \text{run}(x,y)]$$

However, in our system propositional functions (i.e. elements of type $\langle e_n, p \rangle$) can apply only to logical individuals (i.e. elements of type e), so they could never apply to NPs if they were interpreted in the way shown in (18). But properties like the one in (18) (let us call them quantifier-properties) will have individual correlates. Let us call Q (for quantifiers) the set of individual projections (under the embedding map $\hat{}$) of the set of quantifier-properties construable on the base B . Members of Q are the natural candidates in our system for interpreting quantified NPs. Let e_Q be the (sorted) type of individual projections of quantifier-properties (on basis B). We will use superscripted variable of the form x_1^Q, x_2^Q, \dots as ranging over members of Q and superscripted variables of the form x_1^B, x_2^B, \dots as ranging over members of B . In general, quantified NPs will be associated with "non existent" individuals of type e_Q . In this way, Montague's theory of quantification for natural languages can be fully implemented in our framework, and so can all recent developments in the theory of generalized quantifiers.¹³ These developments have brought to light the fact that not all the logically

conceivable quantifiers on a given basis B are possible values of NPs, and have proposed ways in which the class of "natural" quantifiers may be characterized. Any result in this field will be inherited by our framework along the lines just sketched.

In order to actually implement this treatment of quantification into the sample grammar we are gradually constructing, we will introduce, along with a type assignment τ , a sort assignment σ . While the domain of τ is the set of categories, the domain of σ is the set of expressions of a given category. σ specifies for each expression the sortal space of its meaning (which is represented as usual by its translation into IL_*).¹⁴ We will allow, though, statements of the form $\sigma(P_A) = s$, meaning every member of category NP must be an entity of sort s . In our domain of individuals we will have correlates of properties and propositions. It is convenient to have some uniform notation to refer to nominalized properties. Let us stipulate that if \underline{a} is the type of some propositional function, \bar{a} is the sort of the corresponding individual projection. So, \bar{p} will be the sort of eventualities, $\langle \bar{e}, \bar{p} \rangle$ the sort of nominalized 1-place properties, and so on.

For any expression α , we assume that whenever the value of $\sigma(\alpha)$ is not specified in the lexicon, then it is not subject to any sortal restriction. The sort-as-

signment associated with our sample grammar for the moment will be simply the following:

- (19) (i) $\sigma(\text{Pred}) = \langle \overline{e_B}, \overline{p} \rangle$
 (ii) $\sigma(B_{IV}) = \langle e_B, p \rangle$
 (iii) $\sigma(\text{eat}) = \sigma(\text{kick}) = \dots = \langle e_B, p \rangle$
 (iv) $\sigma(\text{seek}) = \sigma(\text{need}) = \dots = \langle e_Q, p \rangle$

(i) states that Preds are individual projections of non-intensional properties (i.e. properties of members of the base). (ii) represents the idea that there are no intensional basic IVs, though there may be intensional IV-phrases (e.g. to be sought). (iii) and (iv) sort out transitive verbs into extensional and intensional ones. You can see that the sort of extensional verbs is different from the sort of intensional verbs in a way parallel to their difference in types in Montague's system.

It is worth noting that in some sense the sets B and Q represent jointly "what is usually talked about". In the standard communicative situation usually assumed in linguistic theory, an utterance will have members of B or Q as its subject matter. Members of B represent "extensional objects", objects which are taken to be "fully specific". Besides ordinary things (this table, Pavarotti, etc.) and abstract entities (the kind of dogs, the action of running, etc.), the set B may also include non actual possibilia (such as Hamlet, Sherlock Holmes, etc., as in Montague's system). We then follow Montague in

assuming that intensional objects (objects of search, desire, etc.) can be modelled by members of Q (i.e. quantified NP denotations). Modelling intensional objects as members of Q, we can represent their lack of 'specificity'. Let us illustrate this by means of an example.

- (20) a. John kicks a unicorn
 b. $\exists x^B [\text{'unicorn'}(x) \ \& \ \text{'kick'}(x)(j)]$
 c. John seeks a unicorn
 d. $\text{'seek'}(\lambda x \exists y^B [\text{'unicorn'}(y) \ \& \ \text{'x'}(y)])(j)$

kick' is a relation between ordinary individuals (i.e. of sorted type $\langle e_B, \langle e_B, p \rangle \rangle$). Hence the (reduced) translation of (20a) will be (20b), which says that there is some specific entity that John kicks. seek' is a relation that can have as one of its relata an intensional object (as by the sort assignment (19 iv)). The member of Q denoted by $\lambda x \exists y^B [\text{'unicorn'}(y) \ \& \ \text{'x'}(y)]$ is taken to represent the object of John's search. There is no specific unicorn that John is seeking. Even though there is some "thing" that John is seeking, such a "thing" does not correspond to an ordinary, 'fully specific' individual. There are members of Q that correspond to members of B, namely those denoted by $\lambda x \text{'x'}(y^B)$ (= Montague's $\lambda P[\lambda y P(y)]$).

Montague's treatment of intensional entities is problematic in many respects. But to my knowledge there are no alternatives that allow for an equally general and compositional approach to this highly intricate problem.

In this regard it is worth pointing out that recent theories of non-existent objects (e.g. Parsons 1981) model such objects roughly as sets of properties, which in some sense correspond to Montague's NP denotations (here reconstructed as members of Q). In other words, there is an evident isomorphism between the logical category of what we call quantifiers and the logical category of Parsons' non-existent objects. This correspondence is, I think, suggestive. Maybe it is possible to rethink Montague's treatment of intensional verbs like seek along the lines of Parson's approach.¹⁵ Be that as it may, it should now be clear in which sense B and Q provide, within the present approach, a representation of extensional and intensional objects respectively, i.e. a representation of the objects we usually talk about. We will call the union of B and Q the referential frame (RF) associated with a grammar.

Something worth considering briefly is that since believe and eat come out as having the same logical type on the present theory, one might wonder whether their being functions of a different sort should be reflected by the sortal assignment σ . It is very easy to do so if one wants it. However, that believe cannot combine with NPs of the form every sandwich is already built into its syntactic category; reproducing the same information in the sort-assignment wouldn't do much work. On the other

hand, believe does combine with NPs like every lie, which might suggest that a 'sortal incorrectness' approach to John believes every sandwich might be preferable, after all. At any rate, I will not pursue this here.

Following essentially Partee and Rooth (1983), I will assume that each function sortally specified as taking members of B has a predictable variant taking members of Q. If α is a B-taking function, $[\alpha]_Q$ will be its Q-taking counterpart. In the case of verbs, the operation $[\alpha]_Q$ is so defined:

$$(21) [\alpha]_Q = \lambda y^Q \lambda x^B, \dots, \lambda x^B [\cup y^Q \lambda z^B [\alpha(z)(x_1) \dots (x_n)]]$$

The sort lifting operator so defined could be generalized along the lines suggested in Partee and Rooth (1983) Appendix A. For the moment we will leave open the exact nature of the sort-lifting process that the operator $[\]_Q$ brings about. It might be built into a rule that could turn out to be the semantic counterpart of some syntactic or morphological process; or it could be an optional rule, analogous to quantifying in. This second view might be justified by the fact the sort lifting rule just considered can indeed be regarded as a special case of quantification (again, see Partee and Rooth (1982), Appendix A for discussion). At any rate, if an unlifted verb, say run is applied to a member of Q, say every man, directly, the result will be sortally deviant. Run can

apply to every man either via quantifying in, or via
 $[]_Q$.

It might also be possible to deal with extensional verbs by means of meaning postulates, as in PTQ. However, Partee and Rooth have argued that such an approach runs into trouble in connection with the analysis of conjunction. To solve these problems they were led to abandon Montague's transparency principle on the relation category-types and to assume a novel interpretive strategy. Essentially, their proposal is that any item is interpreted at the lowest type-level available, whenever possible. So an NP like John would usually be interpreted as an entity of type e , and a transitive verb like eat as an entity of (Montague's) type $\langle e, \langle e, t \rangle \rangle$. If we want to combine eat with a quantified NP such as every sandwich or to conjoin eat with an intensional verb, we will have to resort to the sort lifting rule. In this way they argue that a very general and simple treatment of conjunction becomes possible. In our system the difference between simple and quantified NPs is reconstructed as a sortal distinction in the domain of individuals. So we can adopt essentially Partee and Rooth's proposal concerning conjunction, without giving up Montague's transparency principle, which concerns types (and not sorts).

So far, I haven't said much about the nature of prepositions. Montague thought that there are intensional

prepositions; about would be one of them (in e.g. John is thinking about a unicorn). However, I tend to disagree with Montague's intuitions. With verbs like talk or say, about does not seem to be intensional. This would suggest that, if anything, it is the verb think about that is intensional. I am not able to think of any clear cut case of an intensional preposition.¹⁶ If this is so, then the easiest way to account for it would be assuming the following sortal specification:

$$(22) \quad \sigma(\text{Prep}) = \langle \overline{e_B}, \langle e_B, p \rangle \rangle$$

We can then either extend the definition of our operator $[]_Q$ to cover the case of prepositions, or simply assume that prepositions can combine with quantified NPs only via quantifying in. For the moment we will follow the latter alternative. We will discuss the issues involved in this choice somewhat more in chapter V.

It might be convenient at this point to see how all this might come together with a couple of examples. Before doing it, we need to spell out a little bit more the version of binding rules that we will be assuming. They will be the following:

- (23) s100. STORE IN: $\langle \text{Id}, \text{NP}, \text{NP} \rangle$ where Id = the identity map
 T100. NP translates as x_i and $\langle \text{NP}', i \rangle \in \text{QS}$
 where QS = quantifier store
- (24) S101. STORE OUT: $\langle \text{Id}, \text{S}, \text{S} \rangle$
 T101. $\text{'NP'}(\hat{\ } \lambda x_i \text{'S'})$

Store in is a rule that maps NPs into NPs, replacing their meanings with a variable, and storing such meanings along with the variable index. Store out is an S-to-S rule.

Its semantic effect is to pull the NP-meaning out of QS, forming a λ -abstract over the corresponding variable, and then applying the NP-meaning to such a λ -abstract.

It should be noted that we haven't imposed any sortal restriction on NPs. However, quantified NPs will have their meanings in Q, because of the determiners' meanings, examples of which are given below:

$$(25) \text{ a' } = \lambda x \hat{\lambda} y \exists z [\text{'x}(z) \ \& \ \text{'y}(z)]$$

$$\text{ every' } = \lambda x \hat{\lambda} y \forall z [\text{'x}(z) \rightarrow \text{'y}(z)]$$

The functions in (25) map CN-meanings into members of Q. Pronouns (like he or it) might simply translate as (unrestricted) individual variables; this means that their values can range over both B and Q. We might want to sortally restrict proper names so that their denotata will be members of the base:

$$(26) \sigma(\text{John}) = \sigma(\text{Mary}) = \dots = e_B$$

We can now proceed to give a few sample derivations of simple sentences involving sort-lifting rules and quantification. In what follows we present four derivations of the same sentence. Two of these turn out to be sortally deviant, and hence are ruled out as ungrammatical by the semantics.

(27) Mary loves a man

love a man, IV, love' (a man') sortally deviant
 love, TV, love' a man, NP, $\lambda y \exists z [\text{man}'(z) \ \& \ \text{love}'(z)(y)]$
 a, Det $\lambda x \lambda y \exists z [\text{love}'(z)(y) \ \& \ \text{man}'(z)]$ man, CN, man'

(28) Mary loves a man, S, $[\text{love}']_Q(a'(\text{man}'))(m)$

Mary, NP, m love a man, IV, $[\text{love}']_Q(a'(\text{man}'))$
 love, TV, $[\text{love}']_Q$ a man, NP, a'(\text{man}')
 | love, TV, love' a, Det, a' man, CN, man'

(a) $[\text{love}']_Q(a'(\text{man}'))(m)$

(b) $\lambda y^Q \lambda x [\text{love}'(z)(x)] (\lambda y \exists z [\text{man}'(z) \ \& \ \text{love}'(z)(y)]) (m)$

by def. of $[\]_Q$ and the meaning of a'

(c) $\lambda x [\text{love}'(z)(x)] (\lambda y \exists z [\text{man}'(z) \ \& \ \text{love}'(z)(y)]) (m)$

by λ -conversion

(d) $\lambda x [\lambda y \exists z [\text{man}'(z) \ \& \ \text{love}'(z)(y)] (\lambda w \text{love}'(w)(x))] (m)$

by cancel -cancellation and alphabetic change of variables

(e) $\lambda x \exists z [\text{man}'(z) \ \& \ \text{love}'(z)(x)] (m)$
 by λ -conversion

(f) $\lambda x \exists z [\text{man}'(z) \ \& \ \text{love}'(z)(x)] (m)$

by cancel -cancellation and λ -conversion

(g) $\exists z[\text{man}'(z) \ \& \ \text{love}'(z)(m)]$

by λ -conversion

(29) Mary loves a man, S, 'a man' (λx_1

$\text{love}'(x_1)(m)$)

Mary loves a man, S, $\text{love}'(x_1)(m)$ <a man', 1> \in QS

Mary, NP, m love a man, IV, $\text{love}'(x_1)$ <a man', 1> \in QS

love, TV, love' a man, NP, x_1 <a man', 1> \in QS

a man, NP, a man'

'a man' ($\lambda x_1 \text{love}'(x_1)(m)$) = $\exists x[\text{man}'(x) \ \& \ \text{love}'(x)(m)]$

(30) Mary loves a man, S 'a man' (λ

$x_1[\text{love}'_Q](x_1)(m)$) sortally deviant

Mary loves a man, S, $[\text{love}'_Q](x_1)(m)$ <a man', 1> \in QS

Mary, NP, m love a man, IV, $[\text{love}'_Q](x_1)$ <a man', 1> \in QS

love, TV, $[\text{love}'_Q]$ a man, NP, x_1 <a man', 1> \in QS

love, TV, love' a man, NP, a man'

Before concluding, it is worth discussing briefly a further potential problem that a theory of category-type correspondence has to deal with. Williams (forthcoming) points out that there seems to be a distinction among NPs that might force us to abandon Montague's transparency principle. We will try to translate his argument into our

framework without distorting it too much. Consider the following pattern:

- (31) a. John is a teacher
 b. A teacher is smart
- (32) a. John considers Bill a fool
 b. John gave Bill a book
- (33) a. Every acorn grows into a tree
 b. Every tree grows from an acorn

In the b cases the underlined NPs can be replaced by a referring expression like John or he. For this reason we may call the positions occupied by those NPs 'referential positions'.¹⁷ In the a cases the underlined NPs cannot be replaced with a referring NP, or, if they can, there is a change in the interpretation of the main verb. For instance, in (32a) if we try to substitute him or Mary for a fool the result is bad. If we try to do the same in (31a) the 'meaning' of the copula seems to change. We call the positions where this happens 'predicative positions'.

Now, it would seem that NPs in predicative and referential positions must have different semantic values. In fact, a plausible way of accounting for patterns like those in (31)-(33) would be to follow the traditional view that there are fundamentally two kinds of copulas; the first is the be of identity, the other the be of predication. Some philosophers have found this

bipartition in the meaning of the copula disturbing and have tried to reduce one be to the other. I don't know of any such attempt that is fully successful.¹⁸ According, then, to this traditional view, the be of predication is simply an overt marking to flag that predication is taking place. So this be must be always followed by something that can be predicated, viz. a property-denoting item. The be of identity instead essentially means '=', and hence does not have to be followed by a property-denoting item. On the basis of this, we can analyze the be in (31a) as the be of predication. Hence, what follows it must be a property-denoting item. But NPs in general do not denote properties. In Montague's system where the difference between properties and NP-denotations is represented purely as a difference in type structure, this would have to yield a weakening of the relation between categories and types. One could try to avoid this conclusion by claiming that the underlined constructions in (31a)-(33a) are not NPs (but some other category, say, predicate nominals) or by introducing +PRED as a feature of the category NP. Williams argues against both possibilities.¹⁹

We have tried to argue above that the meaning of be involves essentially the predicator '∩'. It is clear that regarding the role of the copula in this way commits

us to the traditional view, since ' \cup ' (being the reverse of ' \cap ') can apply meaningfully only to (nominalized) properties. If, e.g., $\text{be}' = \lambda x \lambda y[\cup x(y)]$, then be' would turn out to be semantically deviant if applied to an ordinary individual (e.g. John') or to a quantifier (e.g. every man') (cf. the definition of ' \cup ' in ch. I, sec. 3.4. (58)). Hence the be of predication must be flanked by a property denoting expression, if be' is its meaning. So we have to recognize that there are predicative NPs semantically associated with properties. We assume, therefore, that the category NP is partitioned into two subcategories by the feature +Pred, pace Williams. This does not yield, however, a violation of the transparency principle. Both kinds of NPs would be related to (logical) individuals, but of course they would be related to different sorts of individuals. Predicative NPs would be associated with individual projections of properties. They could not be associated with ordinary things (like John or my pen) or with members of Q (i.e. quantifiers). Referential NPs can, instead, be associated with any member of $B \cup Q$. So, again the more constrained type structure that our theory provides seems to allow us to maintain (and strengthen) Montague's transparency principle.

To sum up, in this section we have shown how the distinction introduced by Montague between basic entities and NP-meanings (i.e. generalized quantifiers) reemerges in the present framework as a sortal distinction in the domain of individuals. In this way we can both have our cake and eat it. We can maintain the fundamentals of Montague's treatment of quantification and the advantages that we believe our framework to have with respect to Montague's original one.

3.3. A note on semantic deviance.

In the preceding section, I have been rather elusive on precisely how the sortal articulation of the domain should be carried on. What is the status of sortally incorrect predications? The usual choice here is whether to regard them as logically false or as truth-valueless. The second alternative might have more intuitive appeal. To actually work it out, one would have to implement IL_{*} with truth value gaps, which presents no technical difficulty whatsoever. However, since it is completely routine (and one that would take us out of the main stream of the present work) I would rather leave it as "an exercise for the reader".

This raises the general issue of the status of semantic deviance in a theory of grammar. On the present approach there will turn out to be three main sources of semantic deviance: type-mismatch, sort-mismatch, and meaning postulate violation. As things stand now, "meanings" deviant because of the first reason are non-meanings, their deviancy being established at the level of the logical grammar of IL_* . Something like $x(P)$ is simply not a wff of IL_* .²⁰ Meanings deviant because of sort-mismatch or meaning postulate violations result in logically false formulae (but see below).²¹ All forms of semantic deviance always arise from trying to apply a function to an argument which lies outside of its domain. With enough ingenuity, applying a function to an argument not in its domain could always yield an "undefined". However, I don't see what exactly is to be gained by such an effort.

Somebody might argue that the present situation is infelicitous because it assimilates semantic deviance to logical falsehood, but there are plenty of logically false sentences which grammatically are just fine. I think that the issue here is how logical falsehood is to be defined. Let us agree to call a sentence logically false iff its negation is a consequence of the logical axioms (LA) alone of our theory (in the case at hand the axioms of HST*). So we have the following definition:

φ is logically false iff $\vDash_{LA} \neg\varphi$

Let us instead stipulate that a sentence is semantically deviant iff its negation is a logical consequence of LA plus a set of grammatical meaning postulates (MP). So, we will have:

(34) φ is semantically deviant iff $MP \vDash_{LA} \neg\varphi$ &
it is not the case that $\vDash_{LA} \neg\varphi$.

If the set MP is non empty the set of semantically deviant sentences will be a proper subset of the set of logically false sentences, which is what we want.

Actually, this definition will not quite do in the general case. For suppose that we have a sentence which is semantically deviant according to (34). For the sake of discussion, let us assume that John is promised to come is a sentence of this sort. Then John is not promised to come or John is promised to come and it rains would not turn out to be semantically deviant according to (34), which is obviously unfortunate. It seems that any sentence whose truth or falsity can be determined just by means of meaning postulates (without looking at the world) results deviant at some level. So we can simply stipulate that any sentence whose logical status is determined by our meaning postulates is semantically deviant. Furthermore, any sentence which contains a semantically deviant sentence is deviant too. We propose, then, the following definition of semantic deviance.²²

- (35) a sentence S is semantically determinate iff either $MP \models_{LA} S'$ & not $\models_{LA} S'$ or $MP \models_{LA} \neg S'$ & not $\models_{LA} \neg S'$
- (36) a sentence S is semantically deviant iff it contains as a subsentence (not necessarily proper) a sentence which is semantically determinate.

It should perhaps be noted that, simplicity aside, the present approach to semantic deviance might be appealing because it has some sort of functional basis. A sentence whose falsity can be established just looking at meaning postulates is false solely in virtue of a statement which is part of the grammar, i.e. of the implicit knowledge of the ideal speaker. The speaker "knows" in virtue of the grammar he masters that such a sentence could not possibly be true in any circumstances. So what reason could there be to utter it? And, by a parallel reasoning, why should he be so uncooperative as to utter its negation? Or to conjoin it with some other sentence? Etc. On the other hand, if the truth or falsehood of a sentence is determinable on logical grounds alone (without making use of meaning postulates), then its logical status is not conditioned in any way by the grammar as such. Hence there are no reasons to regard it as being semantically deviant in any way.

It should also be noted that different sets of meaning postulates could yield different kinds of deviancy so that the present approach could be easily extended in

such a way that John was promised to come and John is widespread would both turn out to be semantically deviant but, say, to a different degree. Meaning postulates narrow down the class of admissible models for English. It is quite easy to imagine that instead of simply determining which models are not suitable for interpreting English, meaning postulates could also determine something like a scale of suitability for interpreting English.

We could certainly do something more to semantically deviant sentences, e.g. painting them some ugly colour. Maybe we should. But at the risk of being rather cavalier in handling a tricky question, I would be inclined to relegate this to the (in)famous realm of "problems of execution". The substantive issue seems to me how to identify in a proper, simple and effective way semantic deviance. And the present theory seems able to do so no less than other theories which would prefer having undefined functions all over.

So in the spirit of what I just said, I will assume throughout this work that a sentence is ruled ungrammatical iff either it is not generated by the syntax or is semantically deviant in the sense of (36) above. I will even take the liberty occasionally of saying that a certain function is "undefined" for some arguments. Since we have not officially introduced truth-value gaps in IL_* , the reader should understand this as a shorthand

for "ultimately yielding a semantic deviant wff in virtue of definition (36)".

4. Features.

In PTQ, Montague gives rules roughly like the following:

- (37) Sn. Concatenate NP with VP and mark the main verb in VP as being the 3rd person singular of the simple past
 Tn. H[VP'(NP')]

Interspersing syntactic and morphological operations in this way represents a violation of the strict lexicalist hypothesis embodied in some current version of \bar{X} -theory. According to such a hypothesis, words are first processed by a morphological component of the grammar and then inserted in autonomously generated PS-markers. Why did Montague do that? The point seems to be that a strict lexicalist hypothesis is incompatible with the rule-by-rule hypothesis. It is easy to see why. Morphological features carry meanings. It is fairly uncontroversial that the past tense morpheme in English carries information about the temporal location of the situation being described. Syntactically, the past tense morpheme is attached to the verb. Semantically, the past tense operator has scope over the entire verb phrase, if not over the entire sentence. So if we want to maintain the rule-by-rule hypothesis, there seem to be only two

routes to take. We can attach morphemes to words into a separate morphological component, store their meanings and pull them out at the right point in the syntactic derivation. Or we allow syntactic rules to apply to partially specified words (archiwords). Attachment of inflectional features could then be regarded as a grammatical (i.e. syntactic and semantic) operation on phrases. Montague apparently chose this second route. Bach (1983) has proposed a general framework to spell out precisely what Montague was doing. We will try to cast our theory within such a framework.

I think there are at least two general reasons for doing things this way. First, using storage devices represents an elegant escape path not from the letter but certainly from the spirit of the rule-by-rule hypothesis. It is one that is necessary to handle binding, I think, but it has to be limited as severely as possible. The second reason is that there are a host of phenomena which are clearly "morphological" (in the sense that something happens to the internal structure of words) and yet clearly have a phrasal domain. Bach (1983) gives good examples of such phenomena. For instance, in Amharic the affix for definiteness is attached to the first adjective, or to the verb of a relative clause, or else to the noun itself (if there are no modifiers). Another clear case is the enclitic conjunctive particle *-que* in Latin, or the

postcliticization of pronouns within Italian infinitives. Also the English possessive constructions or the verbal gerunds might be seen as phrasal phenomena. So there are morphological operations that must be defined on phrasal domains. It might be, then, that all operations of inflectional morphology have a phrasal character. This would allow us to give a uniform and general characterization of inflectional phenomena. According to such a view, some morphological operations would be "genuinely" phrasal in that (like in Amharic) a morpheme is infixes in some specified target in a phrasal domain. Other operations, instead, would apply "across-the-board" to all the items in a phrase (granted that they may be undefined for some items). The general picture that emerges from this is the following. Syntactic operations bring together elements of various categories. Categories, however, do not contain fully formed words but archiwords. We index archiwords with a number of morphological operations (features) under conditions of government and agreement (sec. 4.2). These operations will come into play as soon as their domain is reached in the course of the syntactic derivation. The theory of government and percolation is a set of principle determining how indexing of archiwords by morphological operations interacts with syntactic processes.

So the rule-by-rule hypothesis leads us naturally to predict the existence of morphological processes with phrasal domains. It is a consequence of the fact that the semantic contribution of some such processes must take place at a phrasal level. If morphological processes are somehow associated with phrasal semantic rules, and if for any such rule there must be exactly one syntactic counterpart, then those morphological processes must be phrasal too. This seems to be the intuition Montague was building on. In fact, we will try to argue that greater generality can be achieved in a theory of government by regarding all non category-changing morphological phenomena as phrasal. It is interesting to point out how apparent syntax-semantics asymmetries that would seem to constitute a problem for compositionality would turn out to provide striking confirming evidence for it on the present approach.

4.1. On the nature of morphological operations.

In what follows, we will sketch a categorial system of morphological features. We will draw heavily on Bach (1983) to which we refer for both general motivation and formal details.²³ In several points we will propose some modifications of Bach's system and try to motivate them.

Syntactic constituency, on the view adopted here, is taken to be a way of encoding information about semantic function-argument structure. So we are naturally lead to

expect function-argument structure to determine also the domain of other syntactic phenomena, such as, for example, government and agreement. This expectation is justified by the following considerations. On the one hand, it is fairly intuitive to say that functions are "sensitive" to their arguments. By definition, functions are devices for manipulating items and so some properties of those items will obviously be crucial for the way they are supposed to operate. This may be reflected in the form of a function agreeing in some respects with the form of its arguments. So, when properties of an argument determine, in this sense, some properties of a function, it will be regarded as an agreement phenomenon.

On the other hand, functions are, in general, selective: they look for arguments of a specified sort. An argument must qualify as having certain properties to fit a particular "slot" in a function. From this perspective, the form of an argument may depend on the form of a function, i.e. the function may govern certain properties of its argument. So, whenever a function triggers certain properties of its arguments we will have a case of (functional) government. There are two ways in which functional government may be realized. Either a certain class of arguments is intrinsically partitioned into various subsets and functions select one such subset as their domain. Or a class of arguments is not intrinsical-

ly partitioned, but the function carries an instruction as to how to perform such a partitioning. Features are ways of encoding information on morphological properties of items. On the first of the two ways of realizing functional government, it would be natural to construe feature systems as devices for checking consistency of the properties of arguments with what functions require of them. On the second way of realizing functional government, features would be regarded as actual operations of words. I think that both views of features are ultimately necessary. I also believe, though, that the operational view of features is in some sense more fundamental and that the role of features as checking devices can be derived from the former view, as I hope will become clear below.

Before going on, it should be pointed out that not all cases of government can be analyzed as cases of functional government. For instance, the vocative case in Latin (tu quoque, Brute 'you too, o Brutus') cannot be governed by some function, because NPs in the vocative simply do not appear to be operated on by any function. One is forced to assume, then, that government phenomena are also triggered by specific syntactic constructions. We will call this 'structural government'.

We will distinguish three kinds of things: morpho-phonemic operations, morphological features and morpho-

syntactic features. Morphophonemic operations are defined on morphophonemic representations of words. Typically they will result in affixations of various sort. They will, in general, have access to syntactic information only very limitedly (e.g. in the form of conditions on their domains). Morphological features (such as 3rd person, nominative, masculine, etc.) are regarded here as syntactic and semantic operations (or "parts" of operations) on (archi)words that make use in their definition of morphophonemic operations. To emphasize the operational nature of features of this sort, I will put them in square brackets. For instance, in Italian [3^{rd} MS SN] is an operation defined, among others, on the category CN and involves a morphophonemic operation that affixes -o to noun stems. So, [3^{rd} MS SN] (AMIC-) = amico ('friend'). Semantically, [3^{rd} MS SN] carries the information that we are speaking of a single (i.e. non group) male friend (different from the speaker).

The semantic counterpart of [3^{rd} MS SN] (let us indicate it as [3^{rd} MS SN]') could be regarded in this case as a function from properties into properties (which would map, e.g., the property of being a friend into the property of being a male friend, etc.). The same morphophonemic operation of -o affixation which is involved in the definition of [3^{rd} MS SN] can also enter other morphological operations. For example, it is involved in

the definition of [1st PRES IND] (AM-) = amo 'I love'). In "semi-inflectional" languages like Italian a single morphophonemic operation (like -o infixation) may convey information about gender, number and person simultaneously. In 'agglutinative' languages like Finnish, each piece of information would be associated with a different operation. (This is what we mean by saying that a single feature is "part of" a morphological operation). Morphological features are grouped into classes by morphosyntactic features (such as case, gender, number, etc.). Formally, we can regard morphosyntactic features as partial functions that assign to categories a domain of morphological operations (i.e. the set of morphological operations defined for each category). Given that the set CAT of categories is recursively defined, each morphosyntactic function will have to be recursively defined as well. Let us illustrate simply by means of an example how a specification of morphosyntactic functions may be provided in the case of Italian. For any morphosyntactic feature F and any category A, we adopt as a general convention that F(A) is undefined unless otherwise specified. The following matrices provide a specification of some morphosyntactic maps in Italian:

(38) CASE: NP → {ACC, NOM}
 IV → { S/ NP }
 NOM

$$\text{TV} \rightarrow \left\{ \begin{array}{l} \text{IV/ NP} \\ \text{ACC} \end{array} \right\}$$

$$\text{Prep} \rightarrow \left\{ \begin{array}{l} \text{Pred}_{\text{pp}} // \text{ NP} \\ \text{ACC} \end{array} \right\}$$

TENSE: IV \rightarrow { PRES, PAST, FUT }

ASPECT: IV \rightarrow { PERF, IMPERF, ESSERE, AVERE }

$$\text{IV/IV} \rightarrow \left\{ \begin{array}{l} \text{IV/ IV} \\ \text{AUX ASPECT} \\ \text{MOOD} \end{array} \right\}$$

MOOD: IV \rightarrow { SUBJ, COND, IMP, INF, IND }

GENDER: NP, Pred_{CN,ADJ,DET} \rightarrow { MAS, FM }

PERSON: NP, IV \rightarrow { 1st, 2nd, 3rd }

NUMBER: NP, Pred_{CN,ADJ,DET,IV} \rightarrow { SN, PL }

Obviously there are many ways in which a specification of morphosyntactic maps could be improved; but for the general points we want to make this is unnecessary. So, morphosyntactic features are functions that (a) specify the domain of morphological operations and (b) determine what the set of governors is going to be. For instance, CASE(IV) establishes that IVs govern nominative case on their arguments. Similarly, MOOD and ASPECT as applied to AUX establish that AUXs govern the mood and aspectual form of their arguments.²⁴ The particular morphosyntactic map that we adopt incorporates Bach's (1980b) idea that tense and aspect are essentially VP-operators (i.e. functions that apply primarily to IVPs). However, this is not forced upon us by our theory of morphosyntactic features. What such a theory does

require is that morphosyntactic features determine a set of governors. The set of governors is going to include all those items specified as being of the form A/B by some morphosyntactic map.

We now need to say more about morphological operations themselves. The central idea is quite simple. These operations involve, in general, instances of $AFFIX \alpha$, that is to say affixations of various sorts. Syntactic $AFFIX \alpha$ does not alter the internal structure of lexical entries. Morphological $AFFIX \alpha$ does. We pointed out in the introduction to sec. 4 that the rule-by-rule hypothesis leads us to expect morphological operations to be defined on phrases. We assume that in general there are two kinds of such operations. Operations of the first kind are similar to syntactic operations, i.e. something is attached to the n^{th} constituent of a specified type. The parameters that need to be set are the values of n and the type of the constituent, just as for syntactic $AFFIX \alpha$. To give a simple example of such an operation, let us consider the formation of verbal gerunds in English. Such an operation might be defined as follows:

- 39) If $\alpha \in P_{IV}$, $GER(\alpha) = M_{RW}(\alpha, ING)$. E.g.:
- $GER(SING \text{ the national anthem}) = M_{RW}(ING, SING \text{ the national anthem}) = \text{singing the national anthem.}$

The superscripted M (for 'morphological') in (39) informs us that rightwrapping ing in α must result in the creation

of a new word within α . There are various obvious ways in which this might be made explicit. For instance we might imagine that each member of a basic category B_A is bracketed in some way which is different from the way members of the corresponding phrasal category P_A are bracketed. The bracketing component of a morphological operation would then yield lexical bracketings. An operation like the Amharic definiteness marking would be of the same family as (39). Operations of the second family would instead apply sequentially to all the (archi)words of a phrase. A morphological operation γ of this second kind would scan a phrase $\alpha_1, \dots, \alpha_n$ and for each α_i , if α_i is in its domain, γ applies to it, otherwise it leaves it unchanged. This is somewhat reminiscent of proposals concerning across-the-board applications of syntactic operations (see Williams 1977). So, morphological operations such as $[3^{\text{rd}} \text{ MS SN}]$ can be thought as being recursively specified as follows.

- (40) (i) if $\alpha \in B_A$, $[3^{\text{rd}} \text{ MS SN}] (\alpha) = M_{\text{RC}}(\alpha, o)$,
whenever
PERSON(A), NUMBER(A) and GENDER(A) are
defined, and $[3^{\text{rd}} \text{ MS SN}] (\alpha) = \alpha$
otherwise

(ii) if $\alpha \in P_A$ and $\alpha = \alpha_1, \dots, \alpha_n$, then

$$[3^{\text{rd}} \text{ MS SN}](\alpha) = [3^{\text{rd}} \text{ MS SN}]$$

$$(\alpha_1), \dots, [3^{\text{rd}} \text{ MS SN}](\alpha_n)$$

Example:

a. $[3^{\text{rd}} \text{ MS SN}](\text{AMIC}) = \text{amico}$

b. $[3^{\text{rd}} \text{ MS SN}](\text{L MI BRAV AMIC}) =$

il mio bravo amico

the of mine clever friend (ms).

It seems plausible to assume that an item to which a certain morphological operation has applied can never be operated on again by the same operation. In other words, given that morphological operations involve morphophonemic operations (in the preceding example -o affixation) their domain and their range will in general be disjoint. We don't want to keep affixing the same morpheme indefinitely. Inflectional morphology by its very nature does not seem to be recursive.

A schema like (40) provides only the syntax of a morphological operation. Such a schema should in fact be matched by a specification of its semantic counterpart. This will also be done recursively, but it will have to be a recursion of a different type than the one in (40). The reason why this is so is that in general we assume that if an item α is in B_A then it is also in P_A , but its meaning α' is always the same. To assume otherwise would

constitue a violation of the well-formedness constraint, since nothing happens to the syntax of α in passing from the lexical into the corresponding phrasal category, and hence has to be considered a marked option.²⁵ To give a rough idea of how the semantic counterpart of (40) might look like, consider the following.

(41) (i) if $\alpha \in P_{NP}$, then $[3^{rd} \text{ MS SN}]'(\alpha')$ = add
to the presuppositions of : $MS'(\alpha')$ &
 $\neg \text{speaker}'(\alpha')$

(ii) if $\alpha \in P_{CN}$, then $[3^{rd} \text{ MS SN}]'(\alpha')$ = add
to the presuppositions of $\alpha' \lambda x[\alpha'(x) \&$
 $MS'(x) \& \neg \text{speaker}'(x)]$

etc.

Obviously this should not be taken too seriously. The idea, however, is quite clear: morphological operations have both a syntax and a semantics. The latter is an operation on meanings (possibly on presuppositions, like in (40), but not necessarily so).

At any rate, something like (40) is in many ways an hypersimplified statement. Various complicating factors will eventually have to be taken into account. But I don't think that it will endanger the effectiveness of the overall approach. For instance, a problem is that morphophonemic operations involved in morphological

operations have exceptions (e.g. there are a few masculine words in Italian that do not end in -o). There are various conceivable ways in which exceptions might be handled; see Bach (1983) for discussion. Also, it won't in general be the case that a certain morphological operation will have the same morphophonemic (or semantic) interpretation across all categories in its domain (e.g. [3rd SN] will involve different morphophonemic operations on verbs and nouns). So the recursion in (40) has to be modified accordingly.

A further problem concerns the fact that in general features are parts of morphological operations. The grammar has therefore to specify which are the "complete" morphological operations. For instance [MS] or [MS PL] are not complete operations in Italian. Something like [PL] will be associated with a class of morphological operations, namely all those that involve PL (e.g. [NOM 1st PL MS], [NOM 2nd PL MS], [2nd PL MS PRES] where the first two are defined, among others, for NPs but not for verbs, while the latter is the other way round). So single morphological features are associated with classes of morphological operations (at least in inflectional languages). The general strategy that we adopt is then the following. We will freely index each lexical entry with a set of features consistent with their category. "Consistency" is determined by morphosyntactic maps. So,

for example, to index an NP with PAST is inconsistent because TENSE(NP) is undefined. Features may be thought of as sitting in a store in lexical entries. When expressions are combined, features are combined too under conditions of government and agreement (sec 4.2). In this process the class of operations associated with a set of features may be narrowed down to, eventually, a singleton. This is because features combinations will correspond to operations on classes of morphological operations. For instance, when we combine [NOM] with [1st MS SN] we intersect the corresponding class of operations. When a complex of features reaches the point where it denotes a single operation it is ready to apply (i.e. to be pulled out of store). This is going to be made more precise in the next section.

4.2. Government and percolation

Morphosyntactic maps determine simultaneously the domain of morphological features and the set of governors. Within the limits established by those maps, features are freely instantiated in lexical entries. I will follow the traditional view that certain features must be instantiated on members of certain categories (i.e. they are "intrinsic" to those categories) and certain features do not have to be. Items that may carry non intrinsic features are not specified for them. So,

for instance in Italian, common nouns must be specified for gender. So they will carry either the feature MS or the feature FM. Adjectives agree in gender with nouns; this means that they do not carry an intrinsic gender feature. So the lexical entry for casa 'home' and bello 'beautiful' will be respectively:

- (42) <CAS, CN, [FM SN 3rd] , house' >
 < BELL, ADJ, ∅, beautiful' >

Some CNs have both a masculine and a feminine variant like amico and amica. Since gender is an intrinsic feature of nouns, a gender feature must be instantiated on CN lexical entries. So either of the following will be taken as the input to syntactic derivations:

- (43) a. < AMIC, CN, [FM SN 3rd] , friend' >
 b. < AMIC, CN, [MS SN 3rd] , friend' >

Tense and aspect are regarded as intrinsic features of intransitive verb phrases. Tense features will therefore be instantiated on lexical entries as follows:

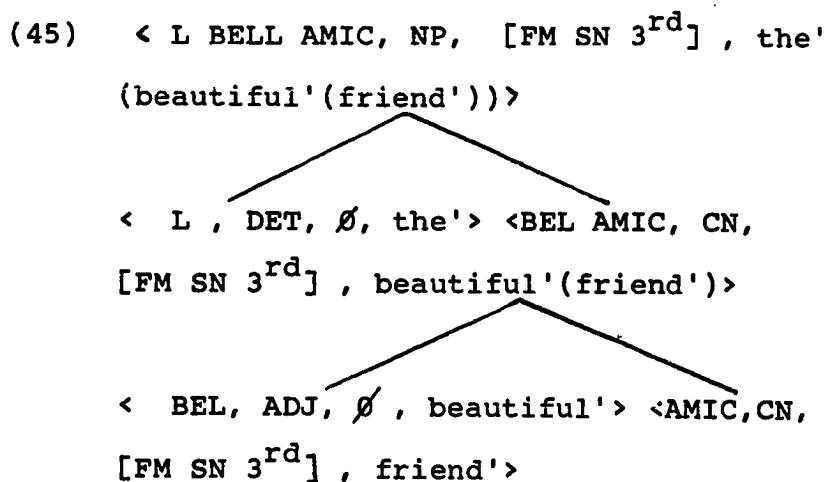
- (44) a. < CORR, IV, [PRES IND] , run' >
 b. < BACI, IV/ NP, ∅ , kiss' >
 PRES ACC
 IND

As is evident from (44b) transitive verbs as such do not carry tense features. However, they are functions from NPs into functions that carry such features. This is the intended interpretation of the notation IV/NP; [PRES] and
 PRES ACC
 IND

[IND] are features of (i.e. members of the feature-store of) the function that results from applying baciare to an NP.

In general we assume that there is a function FT that associates morphological features with lexical entries along the lines just sketched. What we need, evidently, is some recursive specification of the function FT for phrases. We will consider here just cases of function application. The hypothesis we wish to put forth is that in the unmarked case there are just two classes of function applications.²⁶

The first class is when the function is not a governor. This means that we have an item of some functional category A/B that does not impose any requirement on its arguments. In this case the features of both the function (if there are any) and the argument are simply passed up. This is illustrated in the following sample derivation:



Of course this is a particularly simple case for two reasons. First, there is a unique (partial) operation in feature store, while conceivably there might be more than one. Second, the functions not only are non governors, but furthermore do not carry any intrinsic feature. There are various ways in which this simple percolation mechanism could be generalized to cope with more complex cases, but we will not explore them here. In general, thus, we assume something like the following:

(46) Percolative function application (PFA)

If $\alpha \in P_{A/B}$ and $\beta \in P_B$, then $F_\gamma(\alpha, \beta) \in P_A$ and

$FT(F_\gamma(\alpha, \beta)) = FT(\alpha) \sqcup FT(\beta)$

Function application will not work this way when the function is specified as being a governor. In this case what happens can be roughly described as follows. First we compose the governed features with the features of the argument, and we apply the resulting morphological operation to the argument. Second we extract the agreeing features from the argument, we compose them with the intrinsic features of the function (if there are any) and we apply the resulting operation to the function. Then we combine function and argument so modified in the way specified by the syntax. Finally, we pass up the features of the function that are defined for the output category. Let me try to make this clear with an example, then I'll try to make it more precise. Suppose we want to combine a

governing function (say a TV like the one in (47) with the phrase in (45).

(47) <BACI, IV / NP, Ø , kiss'>
 PRES ACC
 IND

First we put together the governed feature ACC with the features of the argument and we get [ACC FM SN 3rd]; this is specified by the grammar as being a complete operation and we apply it to the NP (i.e. the argument):

(48) [ACC FM SN 3rd](L BEL AMIC) = la bella amica

Then we put together the intrinsic features of the function (in this case there are none) with the features of the argument that are defined for the function. Again, in this case there are none, since in Italian there is no object-agreement. So nothing happens to the function. We then put together function and argument in the way specified by the relevant syntactic rule (in this case, right concatenation). The result will be:

(49) <BACI la bella amica, S/ NP, [PRES IND] ,
 NOM

 kiss'(the'(beautiful'(friend')))>

Also, since the function BACI has no intrinsic features, nothing is passed up. Suppose we want now to combine (49) with a subject noun phrase, say:

(50) <PRO, NP, [MS SN 1st] , I'>

We proceed as in the former case. We first build the function [NOM MS SN 1st] and apply it to PRO:

(51) [NOM MS SN 1st](PRO) = io

Here PRO is simply an (abstract) morphophonemic representation of a personal pronoun unspecified for gender number and person. Then we compose the features of the argument defined for the function. What we get in this case are two functions, namely [PRES IND] and [MS SN 1st]. These individuate two different morphophonemic operations. In concatenative morphology these operations will have to be intrinsically ordered. What I mean by that is that for example the tense morpheme in Italian is more internal than the person morpheme. We can capture this fact by saying, for example, that the operation that attaches person-number morphemes (in the case at hand [MS SN 1st]) has at its domain the range of the operation that attaches tense morphemes (in this case [PRES IND]). In general, I assume that the order in which morphological operations apply (when there is more than one) is determined by their definition. So in our case we get:

(52) [MS SN 1st]([PRES IND](BACI)) = bacio

The outputs are then combined by the relevant syntactic operation (left concatenation in this case) and the features of the function defined for the resulting category are passed up:

(53) <io bacio la bella amica, s, [IND] ,
kiss'(the'(beautiful'(friend')))(I')>

The general form of governing function application can be given more or less as follows;

(54) Governing function application (GFA):

If $\alpha \in P_{A/B}$ and $\beta \in P_B$, then $\gamma \in P_A$, where γ is

$F_\delta ([FT_{AG}(\beta) \cup FT(\alpha)] (\alpha) , [FT(\beta) \cup F](\beta))$

and $FT(\gamma) = FT_A(\alpha)$

In this schema, F is the set of governed features, FT_{AG} is the set of features of β (the argument) that trigger agreement on the function (i.e. α), and $FT_A(\alpha)$ are the set of features of α which are defined for the output category A and passed up.

This completes our sketch of how Montague's insight concerning the "operational" nature of morphological inflections can be developed into an explicit feature system within categorial grammar. In essence we have built a simple mechanism of feature matching and percolation into the definition of our core rule, namely function application. There are two fundamental forms of function application, which are determined by the status of the function. If the function is not a governor, all the features of both the function and the argument are simply "passed up". If the function is a governor, the relevant morphological operations are applied to the function and the argument and only some of the function's features are passed up to the output.

4.3. Consequences and comparisons

Even if the feature system we have outlined is far from being exhaustive or fully explicit, it is nevertheless sufficiently worked out to allow us to individuate some of its general consequences. We will focus on five quite different kinds of phenomena which we think provide some support for the general line we are adopting.

4.3.1. Possible governors. Morphosyntactic maps select the set of governors among the functions. So what can count as a possible governor is determined by the categorial structure of the language, and hence by the principles that underlie the choice of a particular categorial system among those that universal grammar makes available. Arguments can never be governors. Nor can inflectional features (such as say [3rd PL]) since they do not have a categorial status. These are all straightforward consequences of the basic line that we are taking, namely that function-argument structure determines the domain of government.

It is very difficult to compare locally theories that differ so much in basic assumptions, such as (extended) MG and GB. So what I will say here must be taken with a grain of salt. Still there are a few differences in the consequences of theories of government which can be individuated quite clearly. Configurational theories of government require, among others, that the governor must

c-command the governed item (Chomsky 1982) and that maximal (\bar{X} -) projections block government. It would follow, then, that the subject position in an English-like language would always be ungoverned (and hence could never receive case via government) since the only thing that seems to c-command the subject is the VP (i.e. a maximal projection). To avoid this undesirable consequence, two moves are usually made. An abstract INFL constituent is assumed under S, and +AGR (one of the possible expansions of INFL) is stipulated to be a governor. Neither of these moves is necessary or possible on our approach. This is a desirable result, since a theory that resorts to similar statements seems to lack any principled basis for rejecting thousands of analogous, no less conceivable stipulations such as "-AGR is a governor" or "+SN governs NP" (granted the possibility of resorting to some suitable abstract configuration), and so on. In this regard, the present theory seems to provide a principled basis for constraining the class of possible governors and governed elements more severely than other current approaches.

4.3.2. Possible agreement features. Only functions can agree on the present theory. Arguments, such as NPs, may have features that are intrinsic to them, because such features carry information which, possibly for semantic reasons, pertains in some intuitive sense primarily to the kind of entities associated with NPs. Consider for

instance plurality. Being a singular individual or being a group seems to be a property of entities. NPs are the typical way of referring to entities. Thus we might expect that in the unmarked case information concerning plurality is carried by NPs. It is possible to conceive a language where information about plurality is carried by, say, the verb (possibly in the form of 'iterative' aspectual markers -- Jespersen 1924 p. 287). However, such an information would have to reduce ultimately to the information that we are speaking about an individual or a group, i.e. to something concerning the make up of entities. Hence it seems reasonable to expect a language where NPs bear no feature of plurality or singularity to be marked, just because of what these features mean. It then follows that PL and SN are possible agreement features, i.e. can trigger agreement phenomena on NP-taking functions such as verbs, being 'intrinsic' to arguments.

By a parallel reasoning, there might be features that are 'intrinsic' to functions, maybe for semantic reasons. Temporal and aspectual markers might be features of this sort. Semantically they correspond to operators on propositional functions (possibly, propositions). Hence, we might expect them to appear as operators on propositional functions, at least in the unmarked case. It is well known that there are languages where something like tense

and aspectual features shows systematically up on nominals (e.g. Kwakiutl, Haisla, Nootka).²⁷ But those are precisely the languages where the distinction between the verbal and nominal systems is fuzzy. To the extent which tense features are intrinsic to functions, the present theory predicts that they never trigger agreement, because arguments do not agree. So, for instance, tense or mood features should not be able to show up on NPs as a function of properties of their governors, i.é. verbs should not be able to govern tense features on NPs. For all I know, this prediction is borne out. In other words, government and agreement are intrinsically directional phenomena, because function application is. Hence, NPs can never "agree" with NP-taking functions, and functions cannot govern features which are intrinsic to them.

So, one consequence of the present theory is that some features are not expected to occur as agreement features on certain categories (e.g. NPs cannot carry, say, mood features, modulo the caveats above), a generalization which does not seem to be accounted for as naturally by other available theories of government.

4.3.3. The phrasal character of inflectional processes. The theory of government and percolation developed in sec. 4.2 ensures that inflectional markings are applied always to maximal phrases. This is a consequence of the fact that only governing functions (and not

endocentric modifiers or determiners) force the feature store of their arguments to be emptied. In this way a uniform definition of morphological operations (including the Amharic definiteness marking) becomes possible, since no further relevant material can be added to phrases after a governing function has applied to them. We saw that the rule-by-rule hypothesis leads us to expect the existence of phrasal morphological operations. The present theory provides a way of implementing this idea in a uniform way.

4.3.4. The domain of government and agreement.

Function-argument structure determines the (putatively) universal domain of government and agreement. This establishes the local character of government. A particular consequence of it is, for example, that government phenomena have to be clause-bound. So, a verb cannot in general govern case on NPs within an embedded S. This is so because those NPs will have already been operated on by the verb(s) of the embedded S; so they will already be case-marked, and hence they will not be in the domain of case-marking morphological operations (since inflectional phenomena are in general non recursive). If we have a function-argument nesting of the form $\gamma(\beta(\alpha))$ where β and γ are both governing functions, then γ will not be able to trigger any government phenomena in α . All of α 's features will have been pulled out at the point when

combines with β . On the other hand we can have indefinitely long nestings of the form $\gamma(\beta_1(\dots(\beta_n(\alpha)\dots))$ where β_1, \dots, β_n are non governing functions, since non governing functions are totally transparent with respect to feature percolation. So, in a sense, governing functions create islands. NPs or Ss are not islands in absolute terms. For example, Ss can be operated on by endocentric modifiers such as S-adverbs. However they will have to be islands when they occur as arguments of a governing function (e.g. a verb).

The following (pointed out to me by F.R.Higgins) represent a rather clear counterexample to the principle of locality just sketched. In Hungarian verbs agree in 'definiteness' with their objects. E.g.: lat-om ot 'see-1sn.subj.-def.obj. him'. However in constructions of the form $V_1[V_{2inf} NP_{def}]$ the definiteness agreement marker shows up on V_1 rather than on V_2 thereby yielding an example of agreement across a governing function, which on our theory should be impossible. An example follows:

- (55) Akar-om megverni ot
 want-1sn. subj.-def.obj. beat him

These constructions need a special analysis on the present theory. But the same seems to be the case for any other approach. See Szabolcsi (1983) for relevant discussion of

these constructions within both a categorial and a GB approach.

While function-argument structure determines the kind of locality relevant to government, the span of government in surface structure (i.e. the linear distance between the governed item and its governor) may vary across languages. The range of variability is determined by parametric variability in categorial systems, however this might be characterized (see sec. 4.1. and the references in fn. 6 for some discussion). So the present semantically based approach provides us with a theory of government and agreement which can aim at universality, modulo independently needed parametricizations of syntactic categories and operations. It is difficult to see how a configurational theory of government could achieve similar results. If one starts off by defining government in terms of configurational notions such as c-command (as in GB), then some abstract English-like syntactic structure will have to be created for languages that have little or no configurationality. No such move will be necessary (or possible) on the present approach. Semantic structure provides us with all the 'virtual' structure (tektogram-matics) that is needed.

4.3.5. Long distance agreement. Natural languages seem to have what looks like long distance, potentially

that agreement phenomena like those exemplified by (56) are closely related to the kind of semantic treatment of gender just hinted at. The starred variant of (56) seems to be bad because a feminine pronoun is used to talk about somebody which has been introduced as a male. Whatever theory will account for the badness of (60) (used to talk about a male) should also be able to account for the badness of the starred variant of (56). The same can be argued about (58)-(59). The ungrammaticality of their starred variants might be attributed simply to the fact that the information concerning the sex of the referent of the pronouns and their antecedents is contradictory. There are well-known problems in working this out, especially if one wants to consider grammatical (as opposed to natural) gender. However, Cooper (1983) has shown that (at least for natural gender) it can be done quite elegantly. He has provided an explicit fragment of English where starred sentences analogous to those in (56)-(59) systematically give rise to presupposition failures which accounts for their ungrammaticality. His treatment of these cases could easily be extended to the present framework (which in many ways is similar to his).

While a semantic account along the lines just sketched might be quite plausible for (56), (58) and (59), a different story might be needed for (57), i.e. for agreement in control structures. The latter kind of

agreement appears to be in some intuitive sense more 'local' than the one illustrated by, say, (59). We will discuss in more detail agreement control in ch. V after having introduced our treatment of control.

Also worth pointing out is that while the semantic approach to (unbounded) gender agreement is independently needed and extends straightforwardly to number agreement, a different approach has to be taken within the present theory with respect to case on wh-words. The natural move to make is to assume that case is assigned to a wh-word by the usual mechanism (sec. 4.2) and stored along with the rest of the wh-word (cf. sec. 2). When the wh-word is pulled out of store, case will be too and provide the correct case marking. See Cooper (1983) and Engdahl (1983) for relevant discussion.

4.3.6. Conclusions. We have considered so far a number of both general and specific consequences of the theory of features we are proposing. The cases we have discussed show, we think, that a Montague-like categorial approach to features does not merely provide a valuable and highly explicit descriptive apparatus (which I regard as a noble enterprise per se). It also contributes to narrowing down substantially the logical space of possible morphological processes, thereby contributing to a better understanding of them.

4.4. Governing functions, heads, and the "case filter".

We would like to discuss here some further consequences of the present approach which we regard as more speculative than those in the preceding section, but still worth pointing to.

Our approach is crucially based on a dichotomous partitioning of functions between governing and non-governing. The empirical basis for this partitioning is provided by the actual morphological paradigms that languages display along with the ways in which the latter interact with syntactic processes. Now, it is pretty obvious that there is a close resemblance or isomorphism between the notion of governing function and the notion of head. Within classical categorial syntax the notion of head has no theoretical status. There is nothing in the way categorial syntax works that allows us to individuate a particular item within a phrase as the head of that phrase. On the other hand, \bar{X} -syntax is designed precisely to accommodate the view that phrases have heads. Once the notion of head is made available (via a certain PS-rule format) then it can be used to state a number of generalizations, including generalizations concerning feature matching and percolation. A typical example of the latter is the head feature convention (see e.g. Gazdar, Pullum and Sag 1981). The route we are taking is sort of the opposite. We first implement a categorial grammar with a

system of features. And we now would like to make the (perhaps obvious) suggestion that the notion of head can be defined in terms of it. A plausible first approximation would be the following:

(61) Definition of head of a phrase:

In combining a function with an argument, the argument is the head of the resulting phrase iff the function is non governing.²⁸

It is easy to see that under definition (60), given standard assumptions, head of a phrase is roughly coextensive with the corresponding notion in most current \bar{X} -theories. It is also easy to see that the head feature convention follows from the present theory (granted definition (61)). If the argument is a head, it will determine basically the features of the phrase, till the point at which it is combined with a governing function. As in the case of \bar{X} -theories, once the notion of head is made available (this time by definition), it can then be used to state generalizations of various sorts (e.g. about word-order).

As usual, it is difficult to see whether at some suitably abstract level our theory and \bar{X} -theories are just notational variants, or whether we are really "explaining" some features of \bar{X} -syntax in terms of the "deeper" notion of function-argument structure. Both cases might turn out to be true. One advantage that the present approach might have over transformationless

\bar{X} -theories of syntax concerns the existence of morphological phenomena like the Amharic definiteness marker (cf. Bach 1983 and sec. 4.1. above). It would seem that phenomena like those should lead to a radical revision of the lexical integrity hypothesis. On the other hand, we have tried to show that, in principle, such cases can be handled by the present theory on a par with other inflectional phenomena.

Let us now turn briefly to a different but related topic. Why do highly configurational languages like English still maintain overt, morphologically evident traces of case marking? Let us see what the present theory might have to say on this. Imagine that there was no morphologically overt trace of case marking in English. Then there would be no reason to consider verbs as governing functions, for they wouldn't govern anything on their arguments. Hence verbs would be expected to act as endocentric modifier in that they would combine with their arguments by what we call percolative function application. Hence, given our percolation mechanism, the features of all NPs of an S would climb up to the matrix S level, and would then be set free all at once at that level. This is what the mechanics of our theory, as we have developed it so far, would yield. What would happen then? Our theory makes no provisions, and it is not immediately evident what provisions could be made. Hence

for language with no overt case marking (Chinese?) our theory seems to make wrong predictions.

In order to extend the present approach to deal with languages which lack any overt form of case marking, the simplest thing to do would be to stipulate that verbs are governing functions (in spite of the fact that they govern nothing) and hence they combine with their arguments by governing function-application. another way to put it is to say that in those languages verbs govern an abstract case, and this is why governing function applicaton has to be used. So, in some sense our theory seems to make the claim that NPs have to have case (overt or phonetically unrealized), which is of course reminiscent of Chomsky's case filter. In our theory, however, such a "filter" is not separately stipulated, but follows from an independently motivated theory of government and percolation.

It should be noted that introducing phonetically unrealized items (e.g. abstract case) represents a violation of the well-formedness constraint. This means that languages where we have to resort to e.g. abstract case marking are predicted to be marked (in relation to case) by the present theory. In other words, it would be no historical accident that even highly configurational languages like English (or Italian) tend to maintain overt traces of case marking, even if such a device lacks apparently any functional basis. Case marking tends to be

preserved because verbs are universally governing functions, i.e. they trigger the emptying of the feature-store of NPs.

Notes

¹The term 'logical form' has a wide variety of uses (not always very clear) in current linguistic theories. Sometimes it is used to mean something like Montague's notion of 'disambiguated language' (see e.g. the discussion in Dowty 1982b). In current GB theories, logical form is the result of various rules (Quantifier construal, reflexive interpretation, etc.) that apply to S-structure. The use we propose is problematic in many respects. We will try to justify it more in subsequent discussions. See also on this the discussion in Chierchia (1982).

²If f and g are two functions, their composition $f \circ g$ is defined as $\lambda x[f(g(x))]$. The proposal to use systematically function composition in grammatical theory goes probably back to Geach (in Davidson and Harman 1972). See also Levin (1976) and the references quoted in fn 6.

³A very good recent discussion of the issues involved can be found in Partee (1983). See also the references therein.

⁴See e.g. Hempel (1966).

⁵The proposal that quantifier scope obeys subadjacency was put forth in May (1977). More recent proposals claim that there are ECP effects on logical forms. The strictest version of the autonomy of semantics thesis would rule out these as possibilities.

⁶Some discussion can be found in Abe (forthcoming), Ades and Steedman (1982), Flynn (1982), Bach (1983).

⁷This point was already noted by F. Heny in his review of Chomsky's "The Logical Structure of Linguistic Theory," which appeared in Synthese (1979).

⁸See especially Bach and Partee (1980), Bach (1983), Dowty (1982).

⁹Of course, I am not claiming that all the semantic differences between CNs and verbs can be reduced to the one I am suggesting.

10 It is conceivable, though, that being associated with (predicable) individuals represents the unmarked case for CNs and being associated with propositional functions the unmarked case for verbs. My semantic hypothesis on the difference between CNs and IVs is still too rough to yield an explanation for this.

11 The idea that lexical items and phrases of the same category might be associated with different types has been systematically exploited in Klein and Sag (1982) to deal with control.

12 The term "condition" is also used by Bealer (1982) but with a different meaning.

13 See e.g. Barwise and Cooper (1981), van Bentham (1982).

14 We can of course have redundancy rules on sortal assignments. But we will not try to investigate such rules here.

15 Cocchiarella (1982) discusses various ways of reconstructing Parsons' theory within HST*.

16 E. Bach points out also the following candidates for intensional prepositions:

- (i) I bought a chicken for dinner
- (ii) Instead of a symphony, I wrote a quartet

17 The notion of referential position (and, derivatively, of referential NP) is orthogonal to (and should not be confused with) the notion of quantified NP. Quantified NPs (such as a man or every man) can occur in referential position (and in this sense will be regarded as "referential"). The terminology adopted here might be a little confusing, but it is pretty standard and I will stick to it.

18 One such attempt is Montague's. He translates the copula as $\lambda P^{\langle e \rangle}(\lambda x[x=y])$. This applied to the meaning of a man will yield $\lambda y[\text{man}(y)]$, and applied to the meaning of John will yield $\lambda x[j=x]$. However this translation will not work for things like dogs are mammals (given Carlson 1976; see also Chierchia 1983). Moreover, Montague's analysis of the copula does not extend to adjectival predicative phrases.

19 I should perhaps point out that while I find Williams' arguments for the NP-hood of predicative NPs quite convincing, I think that his arguments against a +PRED feature are less stringent. But I will not go into this here. Also, Williams has a second argument for postulating a category-type mismatch based on raising structures. We will discuss raising somewhat in chapter IV and offer an analysis that I believe doesn't suffer from the shortcomings Williams points out in connection with other semantically based analyses of this phenomenon.

20 If we do semantics via a direct map into a model, type-mismatches reduce to trying to apply a function to arguments that lie outside of its domain.

21 On the line we are taking, sort mismatches reduce to meaning postulate violations. Sorts act in fact as a shorthand for sets of meaning postulates.

22 I think that (35) and (36) can be restated in our framework in purely model theoretic terms (i.e. in terms of propositions) along the following lines. Let W^* be the set of all worlds that verify the set of axioms of IL^* and W_* the set of all worlds that verify the axioms of IL_* and the set of all meaning postulates. Clearly $W_* \subseteq W^*$. We can then give the following recursive definition of semantic deviance:

- (i) p is semantically deviant if either there is some $w \in W^*$ such that for some $j \in J$, $w(p, w) = 0$ but there is no $w' \in W_*$ such that for some $j \in J$, $w'(p, j) = 0$ or there is some $w \in W^*$ such that for some $j \in J$, $w(p, j) = 1$ but there is no $w' \in W_*$ such that for some $j \in J$, $w'(p, j) = 1$.
- (ii) if p is semantically deviant, then so are $g_{\forall, 0}(p)$, $g_{\neg}(p)$, $g_{\square}(p)$, $g_{\text{H}}(p)$, $g_{\text{W}}(p)$, $g_{\rightarrow}(p, q)$ (for any proposition q), and $f(p)$ (where f is the Fregean embedding).
- (iii) Nothing else is semantically deviant.

23 Cf. also Landmann and Moerdijk (1983).

24 I am not sure of how to deal with auxiliary selection (which in Italian is an extremely complex phenomenon). Currently I would be inclined to regard it as a special case of government. But see Bach (1983) for discussion.

²⁵Not everybody agrees on this. Barwise and Cooper (1981) propose that members of B_{NP} should denote individuals and members of P_{NP} quantifiers.

²⁶See Bach (1983) fn 7 for suggestions concerning function composition.

²⁷Even in English things like ex- or former might be regarded as temporal operators on CNs. Cf. Jespersen (1924 pp. 282 ff.).

²⁸In a preceding, unpublished version of Bach (1983) there was a principle which provides the direct inspiration for this definition.

CHAPTER III

DISTRIBUTIONAL PROPERTIES OF PREDICATIVE EXPRESSIONS AND THE SEMANTICS OF NOMINALIZATION

In the present Chapter we are going to investigate what the semantic framework we have developed might have to say about some semantic and syntactic properties of English predicative expressions and especially about infinitives and gerunds. We will try to show first that our semantic theory, although it can legitimately be regarded as a simplification of Montague's original framework, makes distinctions previously unavailable and allows us therefore to shed some new interesting light on the semantics of infinitives, gerunds and some related constructions. Furthermore, and perhaps more interestingly, given the tight syntax-semantics relation built into our theory of grammar, our hypothesis on the semantics of infinitives and gerunds will reflect back on the syntax leading us to interesting predictions that, we will try to argue, go in the right direction.

One problem we will address is the vexed question of the constituency of infinitives and gerunds. In considering the syntactic category of infinitives and gerunds in relation to their semantic type, there are four main logical possibilities that come to mind. Infinitives and

gerunds might be syntactically VPs of some sort¹ semantically associated with properties; this is what we call the VP = P hypothesis. Or infinitives and gerunds might be syntactically VPs semantically associated with (possibly 'open', i.e. containing a free variable) propositions. Or they might be syntactically clausal (i.e. Ss) semantically associated with propositions. Or finally, they could be syntactically clausal and semantically associated with properties. All these four possibilities are in fact realized in current linguistic theories. The situation can be represented schematically as in the following chart.

1)	syntax	semantics	theory
(A)	VPs	properties	standard MG
(B)	VPs	propositions	Bresnan's LFG (see Halvorsen, 1982); Bach and Partee (1980)
(C)	Clauses	propositions	GB
(D)	Clauses	properties	Williams' predication theory

Hopefully, it should be possible to decide among these hypotheses on empirical grounds. However, it is quite obvious that no hard and fast evidence is going to do the job. The recognition of the fact that theories are underdetermined with respect to empirical data is a well-known (unfortunate?) achievement of modern epistemology. So, choosing among (A)-(D) is going to be more a question of internal elegance and compactness of the

theory or capability of achieving explanatory results and similar not fully understood notions.

There is little doubt that within the general theory of grammar that we adopt hypothesis (A) is by far to be preferred, given the functional correspondence between categories and types and the well-formedness constraint. It also would seem that even on theory neutral grounds (to the extent that there are such things) hypothesis (A) should constitute some sort of unmarked case. For on anybody's account, it would seem, infinitives and gerunds have to be analysed at some level as being some sort of VPs associated with whatever VPs are associated with in their role as predicates. Hypothesis (A) claims that this is sufficient to account for the properties of infinitives and gerunds in general. Hypothesis (B)-(D) claim that some additional apparatus (a 'subject' in the syntax or in the semantics) is needed.²

There are a varieties of arguments that have been offered in favour of (B)-(D). It is impossible and probably unrewarding to consider all of them in detail. What we will do instead is to assume hypothesis (A) (i.e. the VP = P hypothesis) as it represents the unmarked hypothesis within our theory of grammar and see what consequences follow from it on the basis of the semantic theory of properties and predication developed in Chapter I. Then we will draw some comparisons between our theory

and theories that embody hypotheses (B)-(D) and argue that hypothesis (A) gives better results.

In the present chapter we will consider mainly arguments that have to do with distributional characteristics of English predicative expressions and, in particular, with the behavior of inflectional features (i.e. [+AGR]). In subsequent chapters we will consider arguments that have to do with control and anaphora. First we will try to present our arguments in a form which is as far as possible independent of the particular 'execution' of the theory of properties and predication developed in Chapter I that we will adopt. Then we will consider in some detail various possible implementations of the general claims that will be made and their empirical consequences.

1. On the Constituency of Infinitives and Gerunds

1.1. A non argument.

One of the classic arguments for the clausal nature of infinitives has to do with the fact that they have what appears to be a COMP position or, in more neutral terms, they have +WH counterparts. We know that Ss have a COMP position (or have +WH counterparts, i.e. interrogative clauses), hence, the argument goes, by assuming that infinitives are Ss the fact that they have +WH counterparts follows automatically. It is quite evident, how-

ever, that this argument is far from being conclusive. By similar reasoning, for example, one might argue that VPs are NPs (or viceversa), for they both can take PPs as modifiers. If we assume that NPs are VPs this distributional phenomenon would follow automatically, a way of reasoning we don't want to make large scale use of. Furthermore, there are many differences in distribution between tensed wh-Ss and wh-infinitivals, which would appear to lack a principled explanation on the hypothesis that infinitives are Ss. For example, wh-infinitives are a late development in the history of English (F.R. Higgins p.c.). So at some stage, wh-movement to COMP was allowed with tensed Ss but disallowed with infinitives. Furthermore, there are no nonrestrictive infinitival relatives. Also free relatives are disallowed in the case of infinitives. Thus the distribution of infinitival wh-constructions appears to be much more restricted than that of non-infinitivals. It has also been argued repeatedly (e.g. Horn 1975, Reuland 1983) that ACC-ing constructions, such as the ones in (1), have to be analysed as Ss.

- (1) a. I hate John always being late.
 b. John eating a sandwich is a horrible sight.

However, these Ss do not have +WH counterparts:³

- (2) a. *I hate who__being always late.
 b. *I remembered which sonatas Rosa playing__
 c. *The boy (who) Mary dating__is nice.

So being an S is not even a sufficient reason for having +WH counterparts. This further weakens the aforementioned argument. If one assumes that infinitives are VPs, then it is obviously necessary to acknowledge that they have +WH counterparts (or something like a COMP position), while gerunds do not. This constitutes, however, no loss of generality greater than saying that PPs can modify both VPs and NPs. What strengthens this even more is the fact that on anybody's account Ss and VPs will have to form somehow a natural class. On some theories, both Ss and VPs are headed by Vs. On our theory they are both associated semantically with propositional functions. And the list could be made much longer.

There is an area of Italian syntax that provides an interesting kind of support for the present view. The following pattern (from Rizzi, 1982) illustrates the typical parallelism of sentential and infinitival relatives in Italian:

- (3) a. Cerco un uomo al quale presentero Maria.
I am looking for a man to whom I will introduce Maria.
- b. Cerco un uomo che/*il quale presentero a Maria.
I am looking for a man that/*whom I will introduce to Maria.
- (4) a. Cerco un uomo al quale presentare Maria.
I am looking for a man to whom to introduce Maria.

- b. Cerco un uomo da/*il quale presentare a Maria.

I am looking for a man "da"/*whom to introduce to Maria.

It is quite evident from this paradigm that da with infinitival relatives behaves similarly to che with sentential relatives: both are in complementary distribution with respect to the wh-phrase. Che is usually analyzed as the unmarked complementizer of relative clauses in Italian (see the discussion in Rizzi, 1982). Hence, to account for the pattern in (3)-(4) (whatever the account turns out to be) one has to assume that da is the (unmarked) COMP of infinitival relatives. Now, interestingly, the complementizer da behaves as a clitic on the verb, as shown by the following example:

- | | | | |
|-----|-------------------------------|---------------|------------|
| | | possibilmente | |
| | | se possibile | |
| (5) | Cerco un uomo <u>al quale</u> | forse | presentare |
| | Maria. | domani | |

I am looking for a man to whom possibly/if possible/perhaps/tomorrow to introduce Maria.

- | | | | |
|-----|-------------------------|---------------|--------------|
| | | possibilmente | |
| | | se possibile | |
| (6) | Cerco un uomo <u>da</u> | forse | presentare a |
| | Maria. | domani | |

I am looking for a man "da" possibly/if possible/perhaps/tomorrow to introduce to Maria.

Rizzi also points out that da-relatives do not count for subjacency. These facts could be easily accounted for on

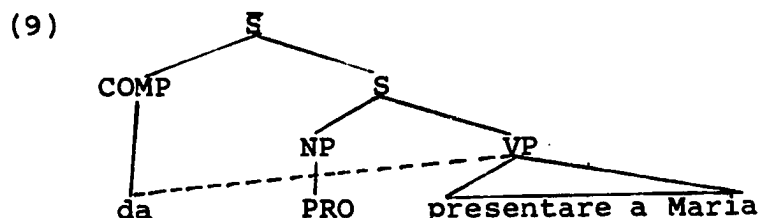
the assumption that the complementizer da forms as constituent with the V, roughly as follows:

(7) $V''[_{\text{COMP}}[da]_V, [presentare\dots]]$

Something like (7) is exactly what we would expect on the hypothesis that VPs have something like a COMP position.⁴ On the other hand, if infinitives are clauses, the structure of infinitival relatives will have to be something like the following:

(8) $S[_{\text{COMP}}[da]_S [PRO_{V''}[presentare\dots]]]$.

So in order to account for the clitic like behavior of da some extra machinery will be needed. Rizzi, for example, suggests reanalysis. The structure in (8) is reanalysed as follows according to Rizzi:



Rizzi proposes to interpret reanalysis as an operation that associates a new analysis with a given terminal string, without wiping out the old one (see Rizzi 1982, p. 98). So the same lexical material will be associated with two distinct analyses and this will allow for the reanalysed structure to have sets of otherwise inconsistent properties. Now, the exact formal nature of reanalysis is quite unclear to me, as is also what plausible constraints one might impose on such a device. Can, for example, as

given construction have three distinct analyses associated with it? Be that as it may, it is quite evident that reanalysis represents a marked option even within theories of grammar that allow for the existence of operations of this sort. However on the assumption that there are +WH (or +COMP) VPs the need for such a marked and hence undesirable choice simply vanishes and the clitic like behavior of da receives a natural and straightforward account.

It is worth pointing out that COMP like particles occur also with infinitival complements such as the following (again from Rizzi, 1982):

(10) Mario pensa di poter partire.

Mario thinks 'of' to-be-allowed to leave.

As expected, particles such as di in (10), be they COMPs or prepositions, display the by now familiar clitic-like behavior:

	difficilmente	
(11) *Mario pensa di	forse	poter partire.
	domani	

Mario thinks 'of' unlikely/perhaps/tomorrow to be allowed to leave.

Nothing special needs to be said about (11) on the present approach. If, however, infinitives are clauses then an appeal to reanalysis will become necessary.

So not only is the classic argument for the clausal nature of infinitives based on the existence of +WH

infinitives at best extremely feeble; it also turns out to induce considerable complications in the grammar as the case of clitic COMP particles in Italian clearly illustrates. The behavior of these particles receives a very simple account on the hypothesis that VPs have something like a COMP position.

1.2. An argument.

A striking empirical generalization about English complementation seems to be the following. English verbs can take both finite and non finite Ss as complements:

(12) John Believes that Mary is home. V - S[+FIN]

(13) John wants (for) Mary to be home. V - S[-FIN]

(assuming for-to clauses to be Ss)

However, verbs can take only non finite VPs (i.e. infinitives or gerunds) as complements:

(14) John tries to run. V - VP[-FIN]

(15) *John tries runs. V - VP[+FIN]

There seems to be a systematic gap in this paradigm that can be schematized as follows:

(16) V - \bar{S} [+FIN]; V - \bar{S} [-FIN]; *V - VP[+FIN]; V - VP[-FIN]

The definition of 'non finiteness' is quite problematic. Below we will try to provide a theoretical characterization of the finite/non finite opposition. For the moment we will limit ourselves to purely observational considerations. Phenomenologically, in English 'non

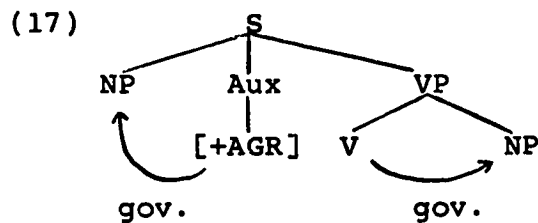
finiteness' can be characterized perhaps as the lack of tense and agreement features on the main verb of a clause or verbal group. In Italian, non finiteness would also be characterised by the lack of inflectional features but in addition to that there is a special 'infinitival mood' morpheme (e.g. -ere in correre 'to run'). In Portuguese non finite clauses may carry inflectional features and are characterized by the presence of an infinitival affix. In other languages (e.g. semitic languages) it might be argued that other non-declarative moods play a role similar to that of 'non finiteness' in English or Romance. It is therefore quite evident that 'non finiteness' may be realized in a variety of different ways and hence its definition will have to be parametrized. However what is interesting about the paradigm in (16) is that modulo these parametrizations in the definition of non finiteness (and modulo the interaction with other relevant factors, such as pro-drop phenomena) something like it seems to be attested pretty steadily crosslinguistically, if the language displays a subordinate/non-subordinate distinction in the verb system. In other words, with the appropriate qualifications, cases analogous to (15) tend to be ruled out in language after language (at least in Indoeuropean and Semitic). Languages tend to mark in some special way subordinate verb forms. It is legitimate to ask why that should be. Why verbal consti-

tents occurring as complements of other verbs (or, if you like subjectless clauses) tend to be systematically marked in a special way?

Within theories that analyze infinitives as VPs the paradigm in (16) hasn't so far received any principled account, to my knowledge. It should be noted that the lack of tense and inflection in infinitival complements cannot be attributed to the lack of an Aux constituent, because of the presence of to. The particle to might be analyzed either as some sort of COMP-like element or as some sort of (degenerate) auxiliary. In either case, its distribution with respect to modals (can, will, should, etc.) can be argued to show that it should be treated on a par with the latter, i.e. presumably generated in Aux position.⁵ So for theories that embody in some form or other hypothesis (A) or (B) of sec. 0 the existence of a systematic gap in (16) seems to constitute a still unsolved problem.

On the other hand, within clausal theories of infinitives and gerunds (and more specifically, within the GB framework) it has been argued that the existence of the gap in (16) receives a principled account in terms of the theory of case and government. Within such theories, the issue under consideration takes the form of what constrains the occurrence of PRO (i.e. the phonetically unrealized pronominal subject of infinitives) to, basi-

cally, just the positions where it actually occurs. In its barest outline, the story is the following.⁶ Case is assigned, in general, under government. Verbs govern and assign case to their objects. However, they do not govern and hence can't assign case to their subjects. What governs and can assign case to subject is the [+TNS +AGR] expansion of the Aux node, which is assumed to be generated under S. This would explain why something like John to run is bad. The lexical NP John has to receive case. But since the clause lacks agreement, there is nothing that can assign case to John. The situation can be schematized in the following way:



The next thing to do, then, is to constrain somehow the occurrence of PRO to positions where it couldn't receive case (and hence a lexical NP could not appear). One way of achieving this within the system just sketched is by saying that PRO must be ungoverned.

Chomsky (1982) tries to derive the fact that PRO must be ungoverned in terms of the binding theory. Again in its barest outline, the binding theory distinguishes between pronominals (he, she, it) and 'anaphors' (reflexives and reciprocals) and states that the former have to

be free within their minimal governing category (principle (B)), and the latter have to be bound within their minimal governing category (principle (A)). Then Chomsky argues that PRO has the typical characteristics of pronominals, since it never has an antecedent within its minimal clause, and furthermore the typical characteristic of anaphors, since it is unable to refer directly to a contextually salient entity. Hence:

- (18) PRO is subject to both the binding conditions (A) and (B). Then PRO is bound and free in its governing category, a contradiction if PRO has a governing category. Therefore PRO has no governing category and is therefore ungoverned (Chomsky, 1982 p. 191).

This is occasionally referred to in the literature as the "PRO-theorem". Now, as pointed out, among others, by Bresnan (1982) there is a clear flaw in this theorem, namely that one of the premises is false. If what characterizes anaphors is their incapability of referring directly to a contextually salient entity, then PRO is not an anaphor in this sense:

- (19) Mary was happy and excited. But to have involved herself in the group was a risky action (from Bresnan, 1982).

So the "PRO-theorem" is in effect a stipulation that PRO has to be ungoverned. For convenience, however, I will continue to refer to (18) as to the "PRO theorem", using the scare quotes as a reminder.

That PRO has to be ungoverned is, of course, not the only stipulation that is made within such an approach. Another stipulation is the fact that [+AGR] is what governs (and hence allows nominative case assignment to) the subject. We have already seen in Chapter II that these kinds of stipulations are impossible within the present theory of grammar, which yields many desirable results. However, even within these limits, there is something appealing about the "PRO-theorem". It represents an attempt to relate the gap in paradigm (16) to mechanisms that have broader application in the grammar, such as case theory. This has interesting empirical consequences. Basically whenever a lexical subject shows up in a non finite clause some special case assigning mechanism is to be held responsible for its case. Hence these constructions are predicted to have a marked status. So for instance in English lexical subjects with infinitives can occur only if there is a complementizer for available (that can be held responsible for case assignment) or with special classes of verbs (essentially believe type constructions). Similarly, in Italian the subject in finite clauses cannot occur after the auxiliary:

(19) *ha Mario lavato i piatti
has Mario washed the dishes

However in certain embedded infinitival constructions, the subject has to occur in post Aux position:⁷

(20) *Mario ritiene la situazione essersi deteriorata.

Mario believes the situation to have deteriorated.

(21) Mario ritiene essersi la situazione deteriorata.

Mario believes to have the situation deteriorated.

So (21) shows that infinitives can have lexical subjects but only if they undergo a process of Aux preposing, a marked phenomenon that does not occur in finite clauses and is triggered by a special class of verbs. Again this special marked Aux fronting phenomenon might be held responsible for assigning case to subjects of infinitival clauses.

The present cursory considerations tend to show that there is something that seems to go in the right direction in the "PRO-theorem" account of the gap in (16). Such an account is based crucially on the assumption that infinitives are clausal. To the extent that no comparable explanation for the pattern represented in (16) seems to be forthcoming on a VP analysis of infinitives, this might be taken as argument in favour for a clausal analysis.

We know that the "PRO-theorem" really is a stipulation in disguise. Yet it is a stipulation that takes care of not just the paradigm in (16) and hence is an interesting stipulation. In what follows we will try to argue that the "PRO-theorem" really is a theorem of semantics. In other words the effects of the "Pro-theorem" (i.e. that

lexical subjects of [-AGR] -- i.e. non-finite -- clauses could not receive case) are predicted to exist by the theory of properties and predication developed in Chapter 1. However, the argument presented in what follows relies, I believe crucially, on the assumptions that infinitives are non-clausal, and in particular on hypothesis (A), sec. O. Hence it provides support for such an hypothesis.

1.3. The semantics of nominalization and the "PRO-theorem."

Our semantics distinguishes two modes of being of properties: as propositional functions and as individual images of propositional functions. Consider now a simple (matrix) subject-predicate sentence:

(22) John runs.

From an intuitive point of view there can be little doubts that runs is a paradigmatic propositional function taking John as an argument. Runs is evidently an unsaturated structure that is looking for an argument to yield a proposition. Propositional functions have individual correlates. Since the distinction between propositional functions and their correlates is so central in our semantics we would expect it to be somehow marked in the syntax of natural languages, given the well-formedness constraint. In other words, in our logical form we have a device, namely '^', that turns predicates into individual expressions. Such a device might be realized in various

ways in the syntax of different languages, but we would expect it to be somehow realized. It may, of course, be realized by some zero morpheme, but the well formedness constraint forces this to be a marked option. How is the nominalizing function to be realized in English? Or, equivalently, what is the individual projection of the propositional function runs? Again there is not much choice as to what the answer can be. Consider:

(23) John to run.

This is not a well formed English sentence. Why? The reason is straightforward if we assume that to run is a way of referring to the individual projection of runs, i.e. a way of referring to the action of running. So to VPs must be the way our nominalizing function is realized on verbs. The reason why (23) is bad is because John and to run are two names of individuals. To run being the individual projection of a propositional function, it cannot, as such, take an argument because individuals don't take arguments. So (23) as a matrix clause is bad for the same reason why (24) is not a sentence:

(24) John Mary.

Something is needed to glue together individuals into propositions. Propositional functions are, typically, the glue. Infinitives, being nominalized propositional functions, have lost their gluing power.

On this hypothesis, the existence of the gap in (16) follows immediately. Consider again (15) reported here for convenience:

(25) *John tries runs.

Tries is a propositional function: it applies to a given number of individuals and gives back a proposition concerning them. But runs is not an individual. It is a propositional function. So tries cannot apply to it, for runs is of the wrong logical type. Propositional functions as such cannot be arguments of other propositional functions. Only their individual projections can. The individual projection of runs is to run, according to our hypothesis. So only to run can occur as an argument of other propositional functions. Whence the badness of (25). On the present hypothesis if something like the paradigm in (16) wasn't attested, we would have to make it up.

Could things be otherwise, on the basis of our semantics? Could we for instance analyze to run as a propositional function and runs as its individual projection? Maybe we could, but it would be crazy. If runs is not as function, why would (22) be good? And if to run is a propositional function, why would (23) not be a proposition? Some quite complicated machinery would have to be devised to account for phenomena which are otherwise very easily accounted for.

Hence, we may conclude that given some very plausible assumptions our semantics not only makes the right distinction (between propositional functions and their individual correlates) but also forces us to analyze inflected VPs as propositional functions and to-VPs as their individual correlates. The facts that the "PRO-theorem" is designed to capture follow from this.

The one we have just given is of course a 'naturalness' argument. Given our semantics it is very natural to look at things in a certain way and very complicated to do the opposite. In presenting the argument we have used very few aspects of the theory of grammar presented in Chapter 2. In this sense this version of the argument is relatively theory independent, in that it relies on the semantic framework presented in Chapter 1 more than on its implementation in a particular theory of grammar. The assumption that we have used is that inflected VPs denote properties that the embedding map defined on properties is encoded as infinitival clause formation. It is easy to see how things would go on the explicit interpretation procedure (based on the rule-by-rule hypothesis) built into our categorial framework. Verb stems would be associated with propositional functions. Inflectional features would not make any specific contribution to the meaning of verbs, while the infinitival morpheme (or possibly the particle to) are interpreted as " \wedge ", and map

propositional functions into their individual correlates. This implies that there is no PRO in the syntax, and in this sense our argument supports the VP = P hypothesis.

One might try to provide an alternative semantic explanation for the "PRO-theorem" phenomenology, based on the analysis of tense, along the following lines. The tense of the embedded verb in (25) would be 'governed' by the tense of the matrix verb (i.e. should be interpreted in relation to the tense of the matrix verb); since however it is an indicative tense it also must be interpreted with respect to the time of the utterance (i.e. indexically), as other indicatives. Thus the interpretation rule for (indicative) tense would yield potentially clashing instructions with respect to the embedded V in (25), whence its ungrammaticality. It should be noted, however, that while there might be something right in this hypothesis (to which we shall return below), in the present form it would be insufficient to explain why there is no subject in (25) and related constructions. Thus such an hypothesis would fail to explain one of the trade mark of the "PRO-theorem" phenomenology.

The explanation we propose is based on the incapability of infinitives of taking arguments (since they are not functions). The argument we have provided above in favour of this view can be made more forcefully within the theory of syntax-semantics encoding developed in Chapter

II. IVs are analyzed as S/NP; by our independently motivated type assignment mechanism, this means that the corresponding logical type has to be $\langle e,p \rangle$, i.e. that of propositional functions. This means that basic lexical entries of category IV (i.e. verb stems) will have to be associated with propositional functions. By our theory of case (Chapter II, sec. 4.2) members of S/NP are necessarily case assigners. Consider next what the contribution of inflection to the meaning of an IV can be. Evidently it can't be ' \wedge ', for otherwise simple sentences like (22) would be ruled out. Whatever the meaning of inflectional features might be, it cannot alter the status of an IV as a propositional function. So our theory really gives us no choice as to what the meaning of runs is going to be.

Consider now the category of TRY, i.e. IV/IV; by our type assignment function its logical type is going to be $\langle e, \langle e, p \rangle \rangle$, i.e. a 2-place propositional function. This means that the members of category IV that TRY is going to combine with will have to be associated with individuals. How is this possible given that the lexical type assignment to the category IV has to be $\langle e,p \rangle$? It is possible only if either the infinitival marking (which in English happens to be a zero morpheme) or the degenerate Aux to are semantically associated with the nominalizing function ' \wedge '.⁸ So again we have no choice as to what

the meaning of to run is going to be. Hence, things like (23) or (25) are impossible. This conclusion is forced upon us by our semantic framework and by the general theory of grammar developed in Chapter II, both of which have been motivated on totally independent grounds.

We have considered so far what we take to be the main aspects of the "PRO-theorem" phenomenology, represented (in a theory neutral way) in the paradigm in (16). There is a second aspect to such a phenomenology, however, namely that occurrences of PRO are essentially limited, in the unmarked case, to the subject position of non-finite clauses.⁹ As a consequence, the control relation (i.e. the relation between PRO and its antecedents) can be characterized, basically, as a relation between a controlling NP and the subject position of a non-finite clause. It can now be seen that also this aspect of the "PRO-theorem" phenomenology follows from our approach. A series of independent considerations brought us to conclude that the infinitival morpheme ([-AGR] in English) or the degenerate Aux to must be semantically associated with the nominalization device 'Λ'. A consequence of this will be that in either case, it is the entire VP-meaning (i.e. a 1-place propositional function) that will be nominalized. In other words, the scope of the nominalization operator 'Λ' must be the entire VP (rather than, say, some subcomponent of it). This is particularly evident if it

is the Aux to that performs the nominalization, since Aux's in the present theory are treated as VP-operators (i.e. IV//IV), which forces them to be associated in the semantics with functions on VP-meanings (see Bach (1980b) for a justification of this view; see also Gazdar, Pullum and Sag (1982) for a closely related approach to the Aux system). If it is the infinitival morpheme that is semantically associated with the nominalizing function ' \wedge ', it also follows from our independently motivated percolation mechanism that such a morpheme has to be 'set free' on the VP and thus whatever is its semantic contribution to the VP-meaning it will have the entire VP as its scope (see Chapter II, sec. 4.2 and next section for details). So in either case our theory predicts that in general only 1-place propositional functions are going to be nominalized by to or [-AGR] and hence can occur as arguments of other propositional functions. The fact that the control relation in general seems to involve crucially only what looks like the phonetically (unrealized) subject position of a non-finite clause is a consequence of this.

Now for the sake of explicitness, let us assume that in English [-AGR], i.e. the infinitival morpheme, is to be associated with the nominalizing function ' \wedge ' and that furthermore infinitives are IVs. In the following section we will discuss in more detail these options, and the various alternatives that there are. For the time being,

let us note that under such an hypothesis it follows that infinitives (being of category S/ NP) are potential case assigners. But, in the unmarked case, they won't be able to actualize their potential because [-AGR] will turn them into individuals and hence they will be unable to take arguments. On the present approach, therefore, the paradigm in (16) is also related to case theory, as in GB, but only in an indirect way, i.e. via semantic function argument structure and the imaging of propositional functions into the domain of individuals. This enables us to derive the effects of the "PRO-theorem" in terms of more general principles.

2. A Typology of Clausal Constructions and the Notion of Finiteness

In the present section we will explore further consequences of the present approach and put forth a general hypothesis on the notion of finiteness. The main idea is implicit in the argument developed in the preceding section. Essentially, we would like to suggest the following (putatively universal) definition of finiteness:

- (26) A clause is finite iff it is gotten by the following rule:
 $S/X + X \Rightarrow S$

What (26) says is that finiteness is crucially related to being a propositional function, (i.e. an unsaturated predicative structure).¹⁰ For a clause to be finite it has to be obtained by saturating a (1-place) propositional

function by means of an appropriate argument. There are a variety of ways in which finiteness might be encoded. The presence of [+AGR]¹¹ on verbs might represent the unmarked way of encoding finiteness. In the preceding section we have argued that this is precisely what follows from our semantics and our theory of syntax-semantics encoding. We are going now to consider several properties of various clausal constructions, trying to see what the present hypothesis on finiteness might have to say about them.

2.1. Finite clauses

On our theory of syntax-semantics encoding Ss are associated with propositions, whatever the latter may be. Propositions are identified with 0-place propositional functions and as such they are not individuals. It follows that propositions as such cannot be arguments of verbs. Hence the badness of (27) and related examples is predicted:

(27) *John smokes bothers Mary.

Something has to map propositions into states of affairs (i.e. their individual projections). The obvious candidate is the complementizer that. Hence the logical form of (28a) would be (28b):

- (28) a. that John smokes bothers Mary
 b. bother'(m) (^ smoke'(j))

Syntactically that might be regarded as functor mapping Ss into Ss (or, possibly, NPs). Believe type verbs being categorized as IV/S would be associated with the logical type $\langle e, \langle e, p \rangle \rangle$, which gives the desired results¹²:

- (29) a. Mary believes that John smokes.
 b. believe'(^smoke'(j))(m)

An interesting case is represented by 'subjunctive' clauses. In English these clauses lack apparently, tense and agreement features and look pretty much like infinitivals:

- (30) a. Mary requestd that Bill's proposal be rejected.
 b. *(that) you be here is imperative

It should be noted that there is general agreement on the fact that subjunctive clauses are clauses (unlike say NP-VP [-FIN] sequences in the complement structure of believe type constructions or of verbs of perception). This entails, within the present framework, that the VPs in subjunctive clauses have to be functions (and hence can assign case). The question that arises is whether the bare infinitival VPs that show up in subjunctive clauses are the same infinitivals that we find in to-VPs. In other words both the VPs of subjunctive clauses and the VPs of to-VPs carry in English a [-TNS -AGR] feature; the question is whether it is the "same" [-TNS -AGR] feature in both constructions. If it is, then [-TNS -AGR] cannot

be associated with the nominalizing function ' λ ', as hypothesized in the preceding section. This would not undermine in the least the results achieved there (i.e. our account of "PRO-theorem" effects). It would merely mean that it is the particle to rather than [-TNS - AGR] that nominalizes VPs (i.e. that is associated with ' λ ').

On the basis of the data available in English it is not easy to decide whether 'subjunctive' VPs and bare infinitivals are the same thing. It has often been noted that the negation not precedes 'subjunctive' VPs, unlike what happens with finite VPs, and that do-support does not occur in subjunctive clauses. These facts might be taken either as supporting the view that subjunctives are simply infinitives or the view that there is a \emptyset 'subjunctive' modal that takes an infinitive. On the first of these views, infinitives would have to be functions, on the second they could be nominalized properties. Since the available data is not sufficient to choose between these two hypotheses, we have to look at other constructions. The only other structures where bare infinitives occur is in the complement system of perception verbs:

(31) Mary saw John run.

It is not clear whether the sequence John run in (31) forms a clause or not. If it is a clause we would have a straightforward account of the well known fact that the "object" John does not passivize in (31), for the category

of see would be IV/S. Furthermore, it would be easier to provide an adequate semantics for constructions such as (31),¹³ because see in these constructions seems to express something like a relation between an individual (Mary in (31)) and an event or situation, rather than a relation between two individuals and an action.¹⁴

However, if John run is a clause it would be very difficult to explain its non-constituent like behavior with respect to standard constituency tests (e.g. clefting, conjunction, right node raising etc., see Gee (1977)). At any rate, the issue that concerns us here is whether bare infinitivals in (31) and 'subjunctive' VPs in (30) are the same thing or not. With respect to this if run in (31) is an argument of the matrix verb (i.e. if John run is not a clause), then it would have to be semantically associated with an individual in our theory, while we have already established that 'subjunctive' VPs must be functions. If, on the other hand, John run is a clause and run, therefore, a function we would have to assume that bare infinitives as complements of perception verbs assign a different case (i.e. accusative) than 'subjunctive' VPs¹⁵ (i.e. nominative). This would seem to constitute some evidence that bare infinitives and 'subjunctive' VPs are not the same construction. Hence we could maintain that the former are functions while the latter are not. This in turn would entail that subjunctives and related non

indicative constructions (e.g. imperatives) are finite clauses in the sense of definition (26). It should perhaps be noted that in languages with a less skeletal inflectional system (e.g. Romance, German, Latin, Greek, etc.) non indicative forms usually display fairly rich mood, tense, and agreement features, which clearly supports the idea that clauses of this sort are finite.

We thus conclude, somewhat tentatively, that subjunctive and related clauses in English are finite. Consequently, we assume that bare 'subjunctive' VPs carry a diacritic feature [+SUBJ] (i.e. subjunctives are [+SUBJ -AGR] while infinitives are [-SUBJ -AGR]). In English [+SUBJ] happens to be phonologically unrealized, while in many other languages it shows up as a lively morpheme. [+SUBJ -AGR] (unlike the infinitival 'morpheme' [-SUBJ -AGR]) is not semantically associated with the nominalizer '∩'. As already pointed out above, not much of the present theory hinges upon the details of the analysis just sketched. What is necessary for the present approach is some way of differentiating infinitives from subjunctives, which can be argued to be necessary on any account. We have of course to explain, then, why subjunctives are odd as matrix clauses. But it is pretty evident that this is to be related to the different contributions to the meaning of sentences made by non indicative and indicative mood in general. The literature on the topic

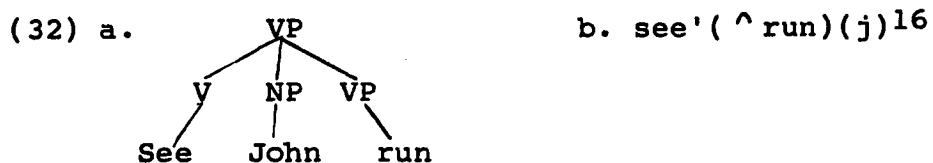
is fairly rich. A recent interesting discussion can be found in Huntley (1982). Huntley proposes to analyze indicative clauses as containing implicit indexical reference to the actual world, to be accommodated along the lines of Kaplan's theory of indexicals. Non indicative clauses would not contain such an implicit reference to the actual world and hence would be associated with propositions that are not 'about' the actual world but are merely envisaged as possibilities. His proposal could be implemented within the present framework, but we will not try to carry this out.

Summing up, our theory leads us to identify two types of finite clauses in English: declarative (characterized by the presence of [+AGR]) and non declarative, characterized by [+SUBJ -AGR] (and [+IMP -AGR]), where the latter, unlike the infinitival [-AGR] does not constitute a nominalization device. Furthermore the present theory leads us to expect the existence of a device (the complementizer that) to map propositions (i.e. the value of finite clauses) into their individual correlates (i.e. states of affairs).

2.2. NP-to-VP and for-to clauses.

According to what we have established so far, to-VPs must be semantically associated with nominalized properties. The question which is still somewhat open is whether it is the to or the infinitival [-AGR] that

"nominalizes" VP meanings. An equivalent way to put it is to ask whether bare infinitive VPs are functions or individual projections of functions. In the preceding section, we have tentatively established that 'subjunctive' VPs are functions but are different from infinitival VPs. So the question whether bare infinitives are functions or not is still open. Now, in English it seems very difficult to find decisive evidence one way or another, because the only contexts where bare infinitives show up seems to be following modal verbs and in the complement structure of causative and perception verbs. With respect to the latter constructions, if we let surface structure be our guide, the available evidence seems to suggest the NP-VP sequences are not constituents as complement of causative and perception verbs. Within the present framework the most natural way of accommodating this is to assume that a verb like see combines with an IV and with an NP to form an IV. But this entails that the (bare infinitival) IV is an argument of see. So, the surface structure of the relevant constructions and the relative function argument structure, would be as exemplified in (32a-b) respectively.



But if the bare infinitival VP run is an argument of see in structures like (33), then it would have to denote an individual, since in our system only individuals can be arguments of propositional functions. Hence the infinitival [-AGR] must be semantically associated with the nominalizing function 'λ'. Then, we are forced to maintain that whatever the contribution of the particle to to VP meanings might be, it does not alter their status as nominalized propositional functions. For sake of explicitness, we might take the following tentative analysis of the relevant items:

(33) a. [-AGR] = $\lambda P[\lambda P]$ b. 'to' = $\lambda x[x]$

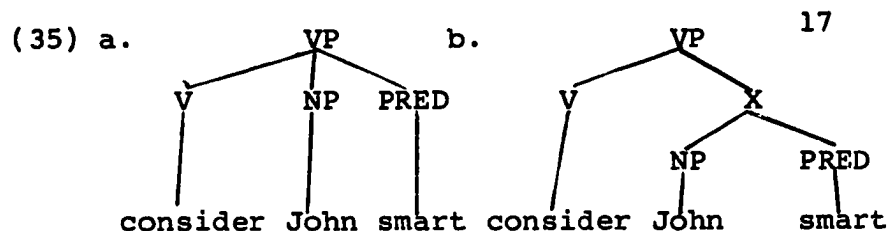
A conceivable alternative would be to analyze verb stems as being semantically associated with individual projections of predicates; inflection (i.e. [+AGR]) would then 'predicativize' verb-meanings. So [+AGR] would be semantically associated with the predicativizer 'λ', in pretty much the same way we have hypothesized the copula be is (see Chapter II, sec. 3.1.). Such an approach could presumably achieve similar results (with respect to "PRO-theorem" phenomenology) to the ones obtained above. However, it also seems to encounter several problems within the grammatical framework that we are assuming. For one thing, we would have to assign (intransitive) verbs to a new primitive category, say IV_{pred} . IV_{pred} is going to be associated, by the type assignment function to

e, i.e. the type of individuals. The functional correspondence principle, which as we have tried to argue in Chapter II gives several explanatory results, forbids that IV_{pred} be defined as, say, S/NP.[+AGR] would then map IV_{pred} into S/NP's. There are two undesirable consequences of this fact; first, inflectional affixes do not, in general, change the category of the items they operate on in the way derivational affixes may do; second, inflected and uninflected IVs would end up being of different syntactic categories. Of course some syntactic feature may be devised to turn IV_{pred} and S/NP into a natural class; but this would be clearly ad hoc. A further undesirable consequence of the hypothesis we are considering is that it would make IVs too similar to CN's, PPs, ADJs etc. We put forth the hypothesis in Chapter II that one possible way of making sense of the way in which IVs differ from other 'predicate-like' items (e.g. CNs) is by assuming that the latter "start off" as individual correlates of properties (i.e. kinds, qualities, locations, etc.) while IVs "start off" as propositional functions. In other words being a propositional function, on this hypothesis, would be the unmarked state for verbs but not for CNs, PPs, ADJS etc. All these seem to constitute sufficient grounds for rejecting the hypothesis that IVs start off as individuals and are predicativized by [+AGR], at least as things stand now.

It is interesting to point out that the hypothesis that non verbal PREDs (i.e. ADJs, PPs, predicative NPs) are not functional categories (i.e. there no categories X,Y such that $PRED = X/Y$) has further nice consequences. Consider so called "small clause" constructions such as the ones exemplified in what follows:

- (34) a. Mary wants John dead.
 b. Mary considers John smart.

A priori there are two possible syntactic structures for these constructions illustrated in what follows:



(35b) constitutes the "small clause" analysis. Now our hypothesis on the nature of PREDs clearly selects something like (35a) (or its equivalent in a categorial framework) as the unmarked analysis for these sort of constructions. For how could (35b) be obtained if both NP and PRED are non functional categories? The rule schemata universal grammar provides cannot combine two non functional categories. Non functional categories can only be arguments, which is what they would be on analysis (35a). Now the empirical evidence available clearly favors (35a) over (35b). The NP-PRED sequences in "small clauses" systematically fail all standard constituency tests. With

consider type constructions the NP behaves like an object (e.g. it passivizes). Furthermore the matrix verbs in the constructions in question impose "selectional restrictions" on the PREDs in their complement structure (see Williams, 1983 for further relevant discussion). All these facts support empirically analysis (35a) over (35b). But on the present approach analysis (35a) was deduced from abstract principles (constraints on syntax-semantics encoding) motivated in a totally independent way, which seems to provide strong support for our approach. That there can't be "small clauses" follows in a principled way from it.

Let us now go back to the analysis of to-VPs. We have established thus far that these constructions have to be semantically interpreted as nominalized properties by trying to argue that a number of explanatory results (essentially the "PRO-theorem" phenomenology) can be derived from this hypothesis. We want now to investigate several other consequences of interest that descend from it.

To-VPs, being semantically individuals, cannot be related directly to NPs in the way propositional functions relate to their arguments. Nominalized propositional functions can be related to an NP argument only within some (higher) relation that has both as arguments in the way illustrated in what follows:

(36) R(NP', to VP')

To-VPs have the potential of being predicated, being nominalized properties, but cannot take arguments by themselves, being nominalized properties.

The next question to ask, then, is how R in (36) could be realized. One obvious answer is as a verb. Given the rule-by-rule hypothesis, we would in such a case expect the NP-to-VP sequences not to form a constituent just as in the case of 'small clauses'. Are these cases attested? Yes; believe-type constructions provide the relevant evidence. Their complements are not for-to clauses and, as expected, fail all the classic tests for constituency. On our theory the syntactic structure of (37a) must be (37b) and its function argument structure must be (37c)¹⁸:

- (37) a. Mary believes John to be home.
 b. [_S[_{NP} Mary]_{vp}[believes_{NP}[John]_{vp}[to be home]]].
 c. believe'(^ be home')(j)(m)

Within the GB framework the projection principle forces the NP-toVP sequences in the complement structure of believe to be an S. A lot of extra machinery is then needed to explain why the object NP behaves like objects do (i.e. receives accusative case, must be disjoint in reference from the subject etc.) and why the NP-to-VP

sequence systematically fails all standard constituency tests. See on this Bresnan (1982) for relevant discussion.

How else could R in (36) possibly be realized? Well, R could be some sort of predication relation or predication marker, an instruction to apply a property to an argument. In such a case we would expect the NP-TO-VP to form a constituent together with whatever the realization of R will be. This is evidently the case of for-to clauses. The complementizer for seems to act just like a predication marker; more precisely, for must be a two-place functor that maps NPs and to-VPs into \bar{S} s, i.e. states of affairs. Now, if we take indicative clauses as prototypical propositions, and that-clauses as prototypical nominalized propositions, then it seems natural to assume that within for-to clauses there is no intervening S node. For creates directly \bar{S} s. If we assume, instead that for creates Ss, then something would have to guarantee that such Ss are always nominalized, an unappealing analysis (cf. also fn. 4). The function argument structure of for-to clauses, a plausible meaning for for and an example are given in what follows.

- (38) a. $\text{for}'(\text{NP}')(\text{VP}')$ or possibly
 a'. $\text{for}'(\text{VP}')(\text{NP}')$
- b. $\text{for}' = \lambda x \lambda y \wedge [\text{ }^{\vee}y(x)]$ or
 b'. $\lambda x \lambda y \wedge [\text{ }^{\vee}x(y)]$
- c. $\text{for}'(\text{John}')(\wedge \text{to do the dishes}')$ = $\wedge [\text{do the dishes}'(j)]$

- c'. for'(^ to do the dishes')(John') = ^ [do the dishes'(j)]
- d. s[NP[Mary][vpwants s[for John to do the dishes]]]¹⁵
- e. want'(^ do the dishes'(j))(m)

There are various ways in which this could be executed. Either for is assigned to category (S/NP)//IV and then the primed version of the analysis of for given in (38) must be chosen. A sample derivation is provided in (39a). Or else, for is assigned to category (S/IV)//NP and then the non primed version of (38) must be chosen. A sample analysis is provided in (39b). Another possibility would be to have a rule that combines to-VPs (i.e. IVs) with NPs and inserts for syncategorematically. This latter approach is compatible with either analysis of for in (38). All such analyses put for into a class with with (in e.g. with John in such a shape, we can't lose) and similar constructions which form S-like structures (cf. for discussion McCawley 1983).¹⁹

- (39) a. for John to do the dishes, S
 For to do the dishes, S/NP (by 'rightwrap') John, NP
 for, (S/NP)//IV to do the dishes, IV
- b. for John to do the dishes, S
 for John, S/IV to do the dishes, IV
 for, (S/IV)//NP John, NP

I tend to favour the analysis illustrated in (39b). There is some evidence that at surface structure for-NPs in

for-to clauses form a constituent. Emonds (1976) argued for this on the basis of the fact that in for-to clauses (unlike what happens in that-clauses) adverbials cannot intervene between the for and what follows:

- (40) *Mary wants for in the kitchen John to do the dishes.

Mary believes that in the kitchen John is doing the dishes.

Chomsky (1982) uses similar data to argue for a requirement of 'connectivity' to which case assignment is subject. However, Bresnan (1982) has pointed out several problems into which the 'connectivity' hypothesis seems to run. Assuming that for is 'cliticized' into the NP with which it is in construction would offer a simpler explanation. Also, there are speakers that accept constructions like:

- (41) I would prefer for Mary and for Bill to pass the exam (from Horn 1975)

If for-NPs in for-to clauses are constituents, it would be possible to analyze (41) without resorting to (an analogous of) right node raising. These facts, however, are far from being conclusive.

Be that as it may, our theory seems to make a number of quite striking predictions. On the basis of the fact that to-VPs must be nominalized properties, a conclusion to which we are forced by a series of independent considerations, we are led to predict that NP-to-VP construc-

tions can form a constituent only if some sort of predication marker (i.e. the complementizer for) meshes them together. On the other hand if no for is available, as with "raising to object" constructions or with so called 'small clause' constructions, then NP-to-VP (or NP-PRED) strings cannot form a clausal structure. All these predictions seem to be borne out.

2.3. On gerunds.

In the tradition of generative grammar, at least three classes of gerundive constructions, with different properties, have been individuated (see e.g. Roeper and Wasow 1972, Horn 1975). They are illustrated in (42)-(44).

- (42) a. John hates someone being killed.
 b. John disapproves of him helping Mary.
 c. Mary didn't have to pay, her husband owning the place.
- (43) a. We deprecate John's always being late.
 b. We attribute great importance to John's being elected.
 c. John's quitting that job surprised me.
- (44) a. The killing of a prisoner of war is a crime.
 b. Loud singing disturbs me.
 c. It took many readings for me to understand that paper.

The gerunds in (42) are known in the literature as ACC-ing constructions. We will refer to the gerunds occurring in ACC-ing constructions as V-gerunds (for 'verbal gerunds'). The constructions in (43) are usually referred to

an POSS-ing constructions. In general, they are analyzed as being NPs with the internal structure of VPs. We will refer to the gerunds that occur in POSS-ing constructions as NP-gerunds. Finally, the gerunds in (44) not only appear to be NPs but also their internal structure is that of NPs, their peculiarity begin that the head of these NPs is a V. We will refer to the gerunds occurring in (44) as CN-gerunds (for 'common noun' -- see below for a justification of this terminology).

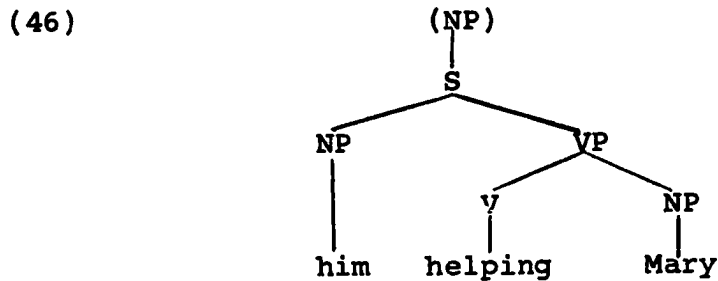
There are other verbal structures where V-ing forms show up. some of them are illustrated below:

(45) a. John, always being late, can really be irritating.

b. John is running.

The V-ing in (45a) is usually analyzed as a participial construction, and the one in (45b) as a progressive form. In the following discussion we will mostly ignore the latter two kinds of V-ing forms. I do not know whether a unified analysis of all V-ing forms is possible, or whether one simply has to assume that there are several-ing morphemes. It certainly seems hard to provide an integrated analysis of (42)-(44) and (45) for these constructions appear to have very different properties. In what follows we will outline an analysis of ACC-ing and POSS-ing constructions and say something about how CN-gerunds relate to the former two.

There seems to be general agreement on the idea that ACC-ing constructions are clauses. On most analysis, their structure would be represented as something like (46).²⁰



The structure in (46) raises interesting issues with respect to our approach. First, on the basis of (46) the subject NP must be an argument of the VP and hence V-gerunds must be functions. However, ACC-ing constructions can never occur as matrix clauses but only in subordinate position. These clauses, furthermore, never take a complementizer, which, in our framework, are analyzed as nominalizing devices. In fact, it has been argued, for this reason, that -ing in ACC-ing constructions plays a role analogous to that of COMPs (see again Horn, 1975 for relevant discussion). So it would seem that V-gerunds are not really propositional functions, but functions from individuals into nominalized propositions (i.e. states of affairs). In other words, V-gerunds are functions that create nominal propositions. Thus the function argument structure associated with (46) has to have the effects illustrated in what follows:

$$(47) \text{ helping Mary}'(\text{him}') = \hat{\ }[\text{help}'(\text{m})(\text{him}')]$$

This in turn means that the gerundive morpheme *-ing* has to map propositional functions into special sorts of propositional functions, say functions from individuals into states of affairs or eventualities. Let us call the latter special sorts of predicates 'eventuality functions'. *-ing* would therefore be some sort of predicate modifier whose meaning might be represented as follows:

$$(48) \text{ a. } \text{ing}' = \lambda P \lambda x \hat{\ }[P(x)]^{21}$$

$$\text{ b. } \text{ing}'(\text{help}'(\text{m}))(x) = \hat{\ }[\text{help}'(\text{m})(x)]$$

Eventuality functions are of course 'argument taking'. Hence, given our theory of case, they are case assigners, a property that they share with 'ordinary' propositional functions. According to the definition of finiteness suggested in (26), then, ACC-*ing* constructions would be regarded as finite, since they are obtained by means of a special sort of propositional functions, i.e. what we call eventuality-functions. We might call clauses obtained by means of eventuality-functions 'semifinite'. Semifinite clauses would share with finite ones the property of being obtained by applying a case assigning VP to an NP. However, since semifinite clauses involve eventuality-functions, one might expect that the case assigned to the subject might be different from the one subjects get in "genuinely" finite clauses. Now, in general the subject of ACC-*ing* constructions bears accusative case.

However, in the absolutive use of ACC-ing constructions, nominative is also possible:

(49) Mary didn't have to pay, ^{he} _{him} owning
the place.

One way of capturing this fact would be by assuming that V-gerunds assign some sort of unmarked case (UC); unmarked case is subject to an 'elsewhere condition': it is assigned whenever no other positive instruction is provided by the grammar. So, in English the unmarked case would be accusative, and the subject of clauses in adjunct position (i.e. say, when they are adjoined to Ss, as ACC-ing constructions in their absolutive use presumably are) could receive a structurally assigned nominative case. When such a structurally triggered nominative case assignment takes place, the unmarked case would not be assigned by the elsewhere condition. In the present framework, then, V-gerunds would be assigned to the following category:

(50) IV = S/NP (where UC = accusative)
ing UC

This simple hypothesis seems able to account for most properties of the construction under consideration.²²

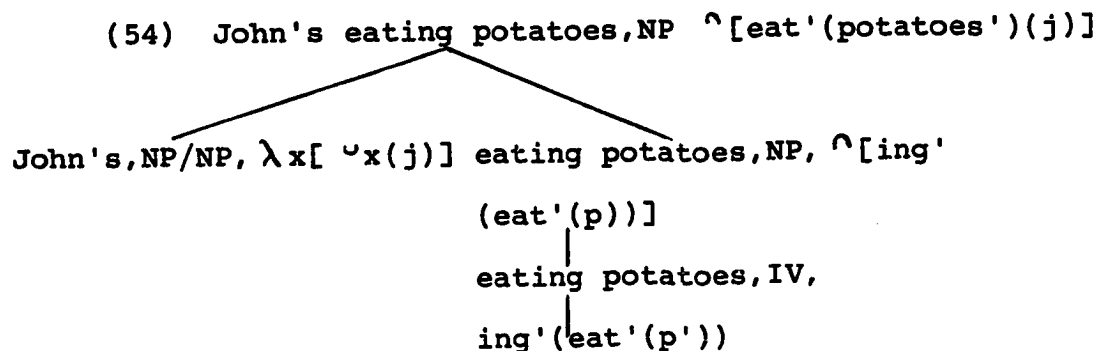
Let us now turn to a consideration of NP-gerunds. NP-gerunds, with or without possessives, behave typically as NPs. NPs in our framework are semantically associated with individuals, hence NP-gerunds must be semantically

associated with individuals of some sort. The obvious move is to treat them as nominalized properties. In particular, this can be achieved by a category changing rule of the following type:²³

- (51) a. if $\alpha \in IV$, then $F_{\gamma}(\alpha) \in NP$, where F_{γ} is
 ing
 the identity map
- b. if $\alpha \in IV$, then $F_{\gamma}(\alpha)$ translates as $\hat{\alpha}$.
 ing

So on this view, (subjectless) NP-gerunds would be associated semantically with nominalized eventuality-functions. They differ from infinitival VPs in the syntax because the latter presumably aren't NPs, and in the semantics because infinitival VPs denote conditions. Concerning the semantic distinctions between NP-gerunds and infinitival VPs, I do not claim that the present approach does justice to them. On this issue, I simply let the theory decide. There seem to be good reasons for treating V-gerunds as special kinds of propositional functions (i.e. eventuality-functions); and if NP-gerunds are derived from V-gerunds (as it seems plausible to maintain), then NP-gerunds will be nominalized V-gerunds. The present approach, for instance, does not represent any difference that arguably might be detected in the way NP- and V-gerunds and infinitival VPs relate to tense and aspect.

Given that NP-gerunds are nominalized properties, they will not be able to take arguments, just like



It is unclear to me whether a unified analysis of possessives in English can be provided that could account also for POSS-ing constructions.²⁴ It is commonly held, within the MG tradition, that possessives are CN modifiers (i.e. determiner like), while quite clearly NP-gerunds are NPs and not CNs (i.e. they do not take determiners). So it may be that the possessive marking on the subjects of gerunds is simply some sort of a reflex of the configuration $NP[NP \alpha]$. At any rate, the analysis in (53) is only meant to be a first approximation, and I will leave things open on this issue. What is important to note is that the present approach leads us to expect NP-gerunds to be nominalized properties and hence unable to relate directly to a subject; from this perspective, the appearance of a possessive marker playing the role of a predication marker receives a principled account. This also allows for an essentially integrated analysis of ACC-ing and POSS-ing constructions, which, to my knowledge, has always encountered serious difficulties in the tradition of generative grammar.

Something which is worth pointing out is that there are verbs that take subjectless V-gerunds as complements. This is illustrated in what follows:

- (55) a. John began singing the national hymn.
kept
- b. John detested singing the national hymn.
practised

There are reasons to believe that the gerunds in (55a) are V-gerunds or participle while those in (55b) are NP-gerunds.²⁵ For instance, the constructions in (55a), unlike those in (55b), do not passivize and are subject to the 'double ing' filter:

- (56) a. *John was keeping swimming.
- b. John was practising swimming the Channel.

If the matrix verbs in (55a) actually take V-gerunds, given that V-gerunds are associated semantically with propositional functions (or rather eventuality-functions), we have to assume that those verbs are in this regard functor-like, i.e. capable of applying directly to properties. This means that the category of begin and keep must be IV//IV. The logical type associated with the latter category, let us recall, will be $\langle\langle e,p\rangle,\langle e,p\rangle\rangle$. We should perhaps notice that there other functor-like verbs of this sort, namely Aux's. On the assumption, motivated in Bach (1980b), that Aux's are IV-modifiers, it is quite plausible to maintain within the present framework, that they (or at least most of them) are predicate-forming functors

(i.e. IV//IV rather than true predicates (IV/IV). So verbs like those in (55a) would have to be regarded as being similar to Aux's in that they take a gerund (or a participle) to form a complex predicate.

CN-gerunds are obtained by a (presumably lexical) process that maps verbs into something like common nouns. So, for instance, killing in (44a) would be a CN obtained from the verb kill; its extension would be a set of events. So CN-gerund formation plays a role which appears to be fully parallel to morphological nominalizations like those exemplified by the pair destroy-destruction. A CN-gerund can denote a set of events, states or processes according to the aspectual nature of the verb it is derived from. This phenomenon is, however, quite intricate and it is impossible for us to consider it here in any detail. We will discuss a little bit more the nature of events and related creatures in the chapter on control. But perhaps we have to say right away something about the difference between what we call states of affairs (or eventualities) and events (and states and processes). States of affairs, in the present framework, are individual correlates or projections of propositions. In this sense, they are abstract or non-existent individuals. States of affairs do not exist in space or time. If propositions are purely mental entities, an option which our theory of properties leaves open, i.e. something

like "thoughts", then nominalized propositions would be abstract images (in the set-theoretical sense) of those mental entities. States of affairs (or eventualities) are something like possible facts. (Actual) facts are states of affairs that are actually the case. Events, states and processes are instead in space and time: complexes of individuals-doing-or-being-something in the world. This is, of course, reminiscent, of the "complex particulars" or "complex occurrences" of Russell's 1911 "Knowledge by Acquaintance and Knowledge by Description." John's love for Mary is a state involving John, Mary and loving. Similarly Brutus' killing of Caesar is an event involving Brutus, Caesar and killing. That John loves Mary or for John to love Mary are eventualities or states of affairs. According to Huntley's (1982) suggestion, indicative that-clause are in some sense "about" the actual world; non indicative nominalized proposition do not 'contain' any overall reference to the actual world.

To sum up, we have provided in the present section the essentials of a unified analysis of ACC-ing and POSS-ing constructions, and said something about how these constructions are related to what we call CN-gerunds. V-gerunds denote a special sort of propositional functions, i.e. eventuality functions. NP-gerunds (without possessive subjects) are associated with nominalized eventuality-functions, and in that regard they differ from

to-VPs. ACC-ing constructions, POSS-ing constructions (i.e. NP-gerunds with possessive subjects), for-to clauses and that-clauses all denote states of affairs or eventualities. The semantic differences among these constructions are not represented in the present system, though some suggestions have been made to the extent that those differences ought to be related to the working of the tense, aspectual and modal systems of English.

2.4. Conclusions.

We have investigated several distributional properties of English predicative expressions and tried to show that the semantic framework developed in Chapter I and the theory of grammar presented in Chapter II lead us to a number of interesting and quite detailed predictions concerning the way finite and non finite constructions behave. We have tried to relate the notion of finiteness to that of propositional function, non finite constructions being constructions that involve crucially nominalized propositional functions (i.e. individual projections of propositional functions). The distinction between propositional functions and their individual images, which is central in our approach, has thus been shown to provide interesting explanatory results.

3. A Case Study: Lexical Subjects in Italian Infinitives

The complementation system of Italian has many properties in common with the English one. In particular, infinitival constructions, as we will try to show below, have to be analyzed as nominalized properties, which predicts for them the impossibility of taking lexical subjects (i.e. of taking arguments). There are however some stylistically marked constructions where infinitival clauses with an overt subject do occur, which, therefore, represent an interesting challenge for our theory. In what follows we will try to consider these constructions in some detail. We will first present the relevant data, drawing heavily from Rizzi (1982). Then Rizzi's analysis, which is cast within the GB framework, will be presented in its essentials. Finally we will provide an alternative analysis within the present framework and attempt some comparisons between the two.

3.1. The data.

In Italian, bare infinitival phrases can freely occur in subject position, as illustrated in what follows:

- (57) a. ballare la rumba e bello
 dance the rumba is beautiful
- b. ballare la rumba non diverte Maria
 dance the rumba does not amuse Mary

Within the theory that we have been developing, it is possible for infinitival VPs to occur in argument position

only if they have been nominalized. Hence we are forced to assume that in Italian the infinitival morpheme -re must be semantically associated with the nominalizing function 'λ'. Hence, Italian bare infinitives will denote individual images of propositional functions (i.e. conditions). If in English there is some doubt as to whether it is [-AGR] or the particle to that nominalizes VP-meanings, the examples in (57) make it pretty clear that no such doubt can arise in Italian.

With this in mind, let us now turn to the relevant constructions. In considering the data, we will differentiate two dialects, indicated by R (a mnemonic for Rizzi) and C (a mnemonic for Chierchia)²⁶. It was already pointed out in sec. 1.2. that in Italian subjects cannot, in general, follow the auxiliary. There is, however, a certain class of verbs (supporre 'to suppose', affermare 'to claim', ritenere 'to believe') that cannot take as complements infinitival clauses with subjects if the subjects occurs clause initially (and this is expected according to our theory) but allow for infinitival complements to have a subject, if this subject is dislocated after the Aux. This is illustrated in what follows:

(58) a. *Maria ritiene/afferma/suppone la situazione essere insostenibile.

Maria believes/claims/supposes the situation to-be unbearable.

(59) Maria ritiene/afferma/suppone essere la situazione insostenibile.

Maria believes/claims/supposes to-be the situation unbearable.

Sentences like (59) are acceptable only at a stilistically marked, literary level. I'll call the property displayed by the infinitival clauses in (59) 'preposed Aux' (PA). A second characteristic of these constructions is that they allow, as it were, for the subject in post-Aux position to be extracted:

(60) Le persone che suppongo non essere state _____ messe al corrente sono molte.

The people that (I) suppose not to-have been _____ informed are many.

I will call this property WH-extractability (WHE), and assume that the complement clauses of the ritenere class is uniquely characterized by these two properties.²⁷

Interestingly enough, in my dialect there are complement constructions of a different class of verbs (aspettarsi 'to expect', preferire 'prefer', desiderare 'to wish') that are characterized solely by the WHE property and not also by the PA property. This is illustrated below:

(61) *Preferirei/mi aspetterei/desidererei molte persone non essere messe al corrente.

(I) would prefer/expect/wish many people not to-have been informed.

(62) *Preferirei/mi aspetterei/desidererei/non essere molte persone messe al corrente.

(I) would prefer/expect/wish not to-have many people been informed.

(63) (R*) Le persone che preferirei/mi aspet-
terei/desidererei non esser state messe al
corrente sono molte.

The people that I would prefer/expect/wish
not to-have been informed are many.

A similar phenomenon seems to take place in French, where the infinitival complements of croyer 'to believe' displays the WHE-property but not the PA-property (see Rizzi 1982, p. 88).

Now for the GB approach these constructions are problematic because the complement clause lacks agreement; hence the subject of these clauses is in ungoverned position (only [+AGR] being a governor); hence no case can be assigned to the subject; hence these structures should be ruled out by the case filter. In our theory these structures should be ruled out as well, because infinitival VPs, being nominalized properties, are intrinsically unable to apply to some argument. So both theories have to do something special about these constructions and the problem is what. Now the empirical generalization seems to be that if some class of verbs has a complement structure that displays the PA-property it also will display the WHE property, but not viceversa. This seems to be what any adequate account of the structures under consideration has to explain.

3.2. Rizzi's account.

Rizzi's solution to the problem presented by the structures in (59)-(60) (recall that (63) is ungrammatical for Rizzi) can be perhaps summarized in four major points, schematically given in what follows.

3.2.1. Aux-to-COMP. Rizzi argues that there exist a rule that moves Aux to COMP in Italian which is responsible for the otherwise inadmissible position of the subject in the structures under consideration. This rule applies in the following constructions:

- (a) Infinitival clauses, as by the example in (59)
- (b) nominal infinitival (i.e. infinitivals preceded by the definite article):

(64) L'essersi la situazione deteriorata e stato risentito da molti.

The to-have the situation deteriorated was resented by many.

- (c) Infinitival adverbs:

(65) Il giudice ha assolto l'imputato per non aver questi commesso il fatto.

The judge acquitted the prisoner for not to-have he committed the deed.

- (d) Adverbial gerunds:

(66) Essendo lui molto in ritardo, abbiamo deciso di incominciare.

Being he very late, we decided to begin.

- (e) Subjunctive clauses:

(67) (C*) Avesse lui capito al volo, non ci sarebbero stati problemi.

Had he understood immediately, everything would have gone smoothly.

The evidence that Rizzi gives to back up the claim that the position of the subject in the relevant clauses is actually determined by an Aux-to-COMP rule relies crucially on the data in (67) which in my dialect (but not in Rizzi's) is ungrammatical. At any rate, I will bear with Rizzi in assuming that there exist some such process. For Rizzi's solution, as we will see, it is quite crucial that the Aux be actually in COMP position at S(urface) structure. This is not so for the account I will propose.

3.2.2. Exceptional case marking. Rizzi assumes that nominative case assignment is structural, i.e. is a purely structural property of certain syntactic configurations, while non nominative case assignment has to take place via a governing lexical head. A special stipulation is then needed to extend the configurations that allow for nominative case assignment to those created by the Aux-to-COMP rule. This can be done as follows:

(68) Assign nominative case in the context Aux ____.

In this way the exceptional case marking is related to the Aux-fronting phenomenon that takes place in the structures under consideration. It might seem that this very same stipulation could account also for the WH-extractibility property of these constructions. The subject, having received case from the Aux in COMP, can then be moved by

WH-movement. Rizzi rejects, however, this solution on various grounds. The main reason is that such a solution would leave unexplained dialects like mine (or corresponding constructions in French) where there are structures that display the WHE-property but not the PA-property. In such cases there is no Aux-preposing phenomenon that could be held responsible for assigning case to the wh-phrase, and hence the wh-phrase (whether moved or not) could not receive case, and the structure would be rule out by the case filter.

Moreover, something has to rule out constructions where the subject of the infinitivals in question is moved by passive or by clitic movement, as the ungrammaticality of the following examples shows:

- (69) a. *Carlo e ritenuto ___ essere strano.
 Carlo is believed ___ to be strange.
- b. *Carlo lo_i ritiene ____i essere strano.
 Carlo him-believes ___ to be strange.

Rizzi wants to rule out (70) in terms of the binding theory. In order to accomplish that, he proposes the following modification to the definition of governing category:

- (70) In the context of structural case assignment, the governing category of α is the minimal NP,S containing α .

Consider now the following configurations:

- (71) a. ...afferma s' [COMP[aver]] s_i sempre
fatto... (wh mov.)
- b. *...era affermato s' [COMP[aver]]
p_s [t_i sempre fatto... (passive)
- c. *...lo afferma s' [COMP[aver]] s_i [t_i
sempre fatto... (cl-mov.)

Since Aux in COMP is a context of structural case assignment, the governing category for the trace t_i will be the embedded S. Rizzi wants to rule out (71b-c) in terms of principle (A) of the binding theory, that requires anaphors to be bound in their governing category. But then, unless one can guarantee that clitic traces and NP traces are anaphors but wh-traces are not, also (71a) would be ruled out. In the latter case it becomes crucial that whenever wh-movement takes place, some other mechanism assigns case to the wh-phase, for movement of Aux to COMP would not be sufficient to guarantee that.

3.2.3. Case assignment to traces in COMP. In order to account for the WHE-property of the constructions in question, Rizzi adopts a suggestion of Kayne's, namely that certain verbs allow, exceptionally, the assignment of case to traces in the COMP position of their complement Ss. So in Italian (at least for Rizzi's dialect) the verbs that can assign case to traces in COMP must be all and only those verbs whose complements can undergo Aux-to-COMP; we will see in a moment how Rizzi tries to guarantee that. The fact that ritenere type verbs allow

for this novel case assignment device is what rescues (71a) and allows for wh-movement to take place in the relevant cases. The crucial derivation is represented by the following structure:

(72) ...afferma[_{S'}[COMP e_i]_S e_i aver
sempre fatto...

The wh-phrase is first moved into the lower COMP where it is assigned case and then moved to the matrix COMP. Rizzi shows how within his system this is the only derivation that is not ruled out by the Empty Category Principle (ECP) and that therefore can account for the WHE-property of the constructions under consideration when Aux-to-COMP (which Rizzi argues to be an optional rule) does not apply.

3.2.4. Subcategorizations. The next, and final, step for Rizzi is to relate somehow the PA-property to the WHE-property. In order to accomplish this he proposes to adopt the following base rules specifying the COMP position to be optional:

(73) a. S' → (COMP) S
b. COMP → (+WH)(TNS)

He further suggests interpreting +WH as designating interrogative clauses, -WH as designating relative clauses and the absence of a WH specification (which is optional in (73b)) as designating declarative clauses. Now, for declarative (non interrogative and non tensed) clauses Rizzi suggests that there are two options ranked differ-

realized lexically. Hence all the properties of the relevant structures are accounted for in a coherent and elegant way.

3.3. A categorial alternative.

Rizzi's solution cannot be directly translated in our framework for various reasons. First, the theory of grammar that we adopt disallows, in general, stipulations concerning the existence of special empty categories (such as COMP would be in Rizzi's theory). Furthermore, in our theory nominative case assignment to subjects that are arguments of VPs must be functional and not structural. Since the subject of the relevant (exceptional) infinitival constructions clearly is an argument of the infinitival VP, the option of structural case assignment is ruled out. So we have to look for a totally different account for the puzzling behavior of these infinitival constructions.

Before proposing an alternative solution, I would like to point out what seems to me a serious difficulty for Rizzi's theory. The main generalization to be explained is the absence of constructions that display the PA-property without having the WHE-property. Does Rizzi's theory really succeed in accounting for this generalization? I fail to see how. The point is that in Rizzi's theory two different mechanisms seem to be responsible for the two relevant properties: Aux-to-COMP takes care of

the PA-phenomenon, case marking of traces in COMP of WH-extractability. What ties these two properties together is the subcategorization in (75), i.e. the presence of an empty COMP position in the frame of ritenere type verbs. However, for the French croire cases and for my dialect one has to assume that case assignment to traces in COMP must be permitted (to account for the WHE-property of these structures) but Aux-to-COMP must be disallowed (to account for the lack of PA-phenomena). But this means that these constructions must have a COMP. So in particular, preferire (which in my dialect has the WHE-property but lacks PA) would have to have the following subcategorization frame in Rizzi's theory:

(76) preferire _S [COMP[]]_S [NP VP []]

The empty COMP is necessary, because that is the only way traces would be able to be assigned case. However Aux-to-COMP has to be disallowed, because these structures do not have PA. But then the lack of verbs that have PA but lack WHE becomes mysterious. Why couldn't there be a construction like preferire (in my dialect) that allows for Aux-to-COMP but not for case marking of traces in COMP? Why should these two properties systematically cluster together if they depend on separate mechanisms? The fact that they do appear to cluster in the way they do seems to be, then, purely stipulative. With this criticism in mind, let us now turn to an alternative approach.

3.3.1. The PA property. In Italian, as already pointed out, subjects cannot in general occur in post-AUX position. So in order to account for the fact that, in the infinitival constructions under consideration, they have to occur in post-Aux position a special (house-keeping) rule will be needed within our theory. The syntactic part of this rule might be given as follows:

$$(77) \quad \alpha \in \text{NP} \text{ and } (\exists \in \text{IV}, \text{ then RWRAP } (\alpha, \text{ inf} \\ \beta) \in \text{S} \\ \text{PA}$$

(77) will have the effect of inserting the subject NP after the Aux in the IV. Now what is the semantic side of rule (77) going to be? Evidently, it can't be functional application, because infinitival VPs are not functions. At the moment where the NP and the IV are combined, our percolation mechanism will "empty" the feature store of the IV and hence the semantic counterpart of these morphological operations (the nominalization of the IV-meaning in the case of the infinitival morpheme) will be triggered (see Ch. II, sec. 4.2. for details). This means that the semantic counterpart of (77) has to contain a special instruction that enables the nominalized propositional function (i.e. the nominalized IV-meaning) to take the NP-meaning as an argument. There are various ways of executing this. One way is to resort to an abstract (i.e. syntactically null) predication marker,

like, say, the English complementizer for. So the semantic side of rule (77) could be given as follows:

(78) if $\alpha \in \text{NP}$ and $\beta \in \text{IV}$, then $\text{RWRAP}(\alpha, \beta)$
 inf
 translates as $\text{for}'(\beta')(\alpha')$

This is simply a way of saying that what is achieved by an overt predication marker in English, is achieved in Italian by using a special sort of syntactic operation (namely RWRAP which is disallowed in normal subject-verb rules). Notice that our theory predicts that the regular NP + IV rule could not be used to combine an infinitival VP with a subject. The counterpart of the NP + IV rule (as is the case for all unmarked rules) is established to be functional application by universal grammar. But infinitival VPs are not functions and hence cannot be applied to arguments. So in order to combine an infinitival VP with a subject, a special rule will be needed, carrying in its semantic component the instruction to attribute the (nominalized)VP to the NP. We would expect, then, this semantic operation to be somehow marked overly in the syntax. In English this is obtained by means of an overt predication marker, in Italian by using a special operation. Thus the similarity and differences between the two languages receive a principled and uniform account.

3.3.2. Some consequences. There are two interesting predictions that the approach just sketched seems to make. Consider first the WHE-property. In the present

framework wh-movement is treated by means of a Cooper-storage device. A typical lexical entry for a wh-phrase would be as follows:

(79) <e, NP, x_i , i, <WH i> >

(79) allows us to have a phonetically unrealized item (a wh-trace, if you wish) in any NP position, which will be semantically associated with a variable. The actual wh-phrase is stored along with the variable index, to be pulled out at the relevant sites (subject to various island constraints). Given this general approach to wh-phenomena, and given rule (77) it follows that the NP of rule (77) can be realized as a wh-word and hence the subject of the infinitival constructions in question will be wh-extractable. There is no way in a construction introduced by something like rule (77) to block extraction of the subject, given the approach to wh-movement we are adopting. This means that there can't be any structure that has the PA-property but lacks the WHE-property, which is the crucial characteristic of the infinitival constructions under consideration. Above we have tried to argue that no comparably straightforward account is forthcoming on Rizzi's theory.

The present approach has a further extra bonus. Consider the way case is assigned to the subjects of our infinitival constructions in Rizzi's theory, i.e. via stipulation (69) sec. 3.2.2. Given Rizzi's approach, the

grammar wouldn't be any more complex if the case on the subjects of PA-infinitives turned out to be accusative or oblique. What the case on these subjects is going to be is simply stipulated, and so there should be no special reason to expect nominative case to be assigned. This is not so on our approach. The category of infinitives is IV, i.e. S/NP . Since infinitives belong to the same category as non-infinitival VPs there should be no reason to expect, in the unmarked case, that the case governed by infinitives vs the one governed by non infinitives should be any different. Case is an operation associated with categories. Thus, if two items belong to the same category, they are in general expected to govern the same case specification. Now the way case is assigned in clauses generated by our rule (77) is by means of the general mechanism built into (governing) functional application. In particular, the expanded form of rule (77) will be roughly:

$$(78) \quad \text{if } \alpha \in \text{NP} \quad \text{and } \beta \in \text{S/NP} \quad \text{then RWRAP}([\text{CS} \\ \text{aGND} \quad \quad \quad \text{cs} \\ \text{bPRS} \quad \quad \quad \text{inf} \\ \text{cNMB} \\ \text{aGND bPRS cNMB}] (\alpha), [\text{aGND bPRS cNMB}] \\ (\beta)) \in \text{S}.$$

The general mechanism of government developed in Chapter II predicts, thus, that the case assigned to the subjects of the infinitival constructions in question is going to be nominative, rather than some other possible case.

3.3.3. The WHE-property. We have now to say something about constructions, such as preferire in my dialect, that have the WHE-property but lack PA. Evidently something like our rule (77) will not do for these cases. One way to go would be to adopt the following rule for the constructions under consideration:

- (79) a. if $\alpha \in IV/IV$ and $\beta \in IV$, then $LC(\alpha, \beta)$
 $\begin{matrix} +wh \\ \in IV \\ +wh \end{matrix}$
- b. if $\alpha \in IV/IV$ and $\beta \in IV$, then $LC(\alpha, \beta)$
 $\begin{matrix} +wh \\ \in IV \\ +wh \end{matrix}$
 translates as $\alpha'(\wedge [{}^u(\beta'(x_i))])$, where i
 is the index of the WH-store of α .

Let me explain what rule (69) achieves. The basic idea is that verbs like preferire that have the WHE-property carry in their wh-store a wh-word along with a variable index. When they combine with an IV, the variable will be inserted as an argument of the complement IV in the way shown in (79b). The whole expression will now have a WH-phrase in its WH-store, which will then be treated as all regular WH-phrases. So, in other words, items like preferire that have the WHE-property will carry, optionally, a phantom WH-phrase that will be construed in the semantics as the subject of their IV complement. Rule (79) is triggered only by a marked lexical property that some items (preferire type verbs in Italian, croyer type verbs in French) may have.

Our proposal is thus based on two (housekeeping) rules. These days, there is a strong tendency to limit as much as possible what has to be written into language specific rules in favour of developing a parametrized system of 'general principles', a program with which I am fully sympathetic. However, we are studying here a special kind of marked constructions. On anybody's theory, language specific stipulations are going to be needed to handle constructions of these kinds. In Rizzi's theory 4 different devices were argued for (Aux-to-COMP, nominative case assignment in the context Aux____, nominative case assignment to traces in COMP, verbs subcategorized for clauses with empty COMP positions). In our theory we have rule (77) and rule (79). Such rules have been shown to fit naturally within the theory of predication, case and, in general, function-argument structure embodied in universal grammar. They represent rather minimal departures from the unmarked case and can account for language particular variations in the ways infinitives can have subjects in English and Italian, thus providing us with a theory that can be legitimately regarded as "parametric". Furthermore, this is achieved within a framework that can be argued (as we have tried to do in Chapter II) to be generally more restrictive than the GB framework adopted by Rizzi. This would constitute *per se* an interesting result. But we have also tried to show

that our theory offers several empirical advantages over Rizzi's. First our theory seems capable of providing a principled explanation of why there aren't constructions that display PA but not WHE, explanation that does not seem to come as straightforwardly on Rizzi's theory, especially if that theory is extended to cope with dialects (such as mine) where constructions with PA and WHE coexist with constructions that display solely WHE. Second, our theory provides an account of why the subjects in the constructions in question show up in the nominative, rather than any other available option, while such an account is missing in Rizzi's theory.

In conclusion, these considerations tend to provide strong support for the general line of inquiry adopted in the present work and, more specifically, for the theory of properties that underlies our theory of grammar. It is particularly interesting that such support should come from a phenomenon (infinitival clauses with subjects in Italian) that *prima facie* appears to be a counterexample (or at least a problem) for our theory.

Notes

¹There is widespread agreement on the view that gerunds such as the one in (i) are NPs:

(i) John practised playing the violin

The question I am trying to address here is whether such gerunds are clausal in that they have an implicit subject (in the syntax or in the semantics) or are VP-like in that they don't have a subject. As far as the syntax goes, what is at issue is whether the syntactic constituency of gerunds is as shown in (ii) or in (iii).

(ii) NP[vp[playing the violin]]

(iii) NP[(s[])PRO vp[playing the violin]()]

²For a defense of hypothesis (B) see e.g. Bresnan (1982) and Bach and Partee (1980). For hypothesis (C) a good survey can be found in Koster and May (1982). For hypothesis (D) see Williams (1980, 1983).

³I am following here the traditional view that for-to clauses are (untensed) Ss. A classic discussion of these constructions can be found in Bresnan (1972).

⁴Notice that if we analyze also (5) along the lines of (7) (i.e. assuming that there is no intervening S), then we have to assume that 'propositional' adverbs like possibly or if-clauses can also be VP-modifiers. I don't see any special difficulty with this proposal. The syntactic and semantic differences between 'propositional' and 'non-propositional' modifiers could still be captured in various ways. For example, in the semantics the form of 'propositional' VP modifiers like, say, possibly might be roughly $\lambda VP' \lambda NP'[possibly'(VP(NP))]$. In the syntax one could maintain that 'propositional' VP modifiers occupy more 'peripheral' positions than 'non-propositional' adverbials.

⁵Most analysis of the English auxiliary system seem to agree on this point. See e.g. Akmajan, Steele and Wasow (1979), Gazdar, Pullum and Sag (1981) or Koster and May (1982) for relevant discussion.

⁶See Chomky (1982) pp. 188 ff. We will discuss the theory of binding in more detail in Chapter V.

⁷A detailed discussion of exceptional case marking in Italian can be found in Rizzi (1982), from which these examples are taken. Below we will develop an alternative account of the same range of phenomena.

⁸If to-VPs are IVs, it follows that we will have to allow for a limited relaxation of the functional correspondence constraint (see Chapter II, sec. 3.1.) for phrasal categories.

⁹'Tough-movement' constructions involve verbal structures where something like PRO seems to occur in non-subject position. However these constructions on everybody's analysis are analyzed as VPs with a gap, i.e. they are not 'whole' constituents in the way infinitives are (see e.g. Chomsky 1982 or Gazdar and Pullum 1983). "Gapped" constituents require a different kind of treatment.

¹⁰Recall that the logical type associated with a category of the form S/X must be, according to our theory of syntax-semantics encoding $\langle e, p \rangle$, i.e. the type of 1-place propositional functions.

¹¹Or, more precisely, by the lack of an infinitival mood affix that in English is realized as [+AGR].

¹²It is a well known fact that in English complementizers are optional in post-verbal position. I don't have anything to say about this, apart from pointing out that they are always obligatory in Italian (with indicative S complements). For example:

(1) Maria crede *(che) Gianni sta fumando.

Maria believes (that) Gianni is smoking.

¹³For relevant discussion see Akmajan (1977), Gee (1977), Barwise (1981), Higginbotham (1981). Also worth mentioning briefly is that another traditional argument for clausehood is the capability of taking 'dummies' as subjects (e.g. John saw it rain). However, currently this argument has lost a great deal of strength. Cf. Williams (1983) or Klein and Sag (1982) for relevant discussion. See also Chapter IV below.

¹⁴It should be noted that analyzing perceptual reports as relations between two individuals and an action could account for most of the semantic properties of these constructions and would be compatible with the idea that there is a (derivative) sense in which seeing can be

regarded as a relation between an individual and (something like) a situation. In Chapter IV we are going to argue for a general treatment of control based on entailments of the following sort:

(i) Mary sees John run (=see'(^run')(j)(m)) in all the situations compatible with what Mary sees, John runs.

¹⁵Recall that on our theory all members of IV = S/NP are case assigners.

¹⁶The syntactic structure in (32a) is the same that Gee (1977), Bach (1980a) and Williams (1983) argue for. I have nothing to say about the impossibility of passivizing the object NP in (32a). See Bach (1980a) and Williams (1983) for some suggestions. It is useful perhaps to recall that bare infinitive constructions in perceptual reports do passivize in most languages where they are attested (see on this again Gee (1977)). Cmp also the following Italian example:

(i) Gianni e stato visto correre.

Gianni was seen run.

¹⁷In constructions such as I heard it sung by an Italian tenor the NP-PRED sequence might simply be an instance of attributive uses of the participle. The same holds for one of the possible analysis of sentences like I saw John disappearing in the dust (cf. the discussion in sec. 2.3.).

¹⁸Another reason why NP-to-VP sequences in the complement structure of believe couldn't possibly be a constituent on the present theory is because the object NP can undergo passive. Hence believe-to-VP must be a TVP. See on this Bach (1980a).

¹⁹Similarly to the complementizer that, for can (must?) be omitted in various circumstances.

²⁰See e.g. Horn (1975), Reuland (1983), McCawley (1983). Akmajian (1977) argues for a different structure in relation at least to ACC-ing constructions in the complement structure of perception verbs. But see Reuland (1983) for a criticism of Akmajian's proposal.

²¹Actually, ing' maps a property P into a function from individuals into individuals (i.e. something of type $\langle e, e \rangle$); however any member α of ME $\langle e, e \rangle$ has an "iso-

morphic copy" α' in ME $\langle e, \langle e, p \rangle \rangle$ defined as follows: $\alpha' = \lambda x \lambda y [\alpha(x) = y]$; α' is as (two place) 'functional' propositional function (i.e. for any x there is a unique y such that $\alpha'(x)(y)$ holds). So for any property P, ing'(P) corresponds uniquely to a propositional function. In this sense ing' can be understood as a shorthand for a map from properties into properties. Also worth pointing out is that I don't know whether the present analysis of the semantics of ing has anything to say about other semantic properties of ing-constructions (e.g. their factivity).

²²The subjects of ACC-ing constructions (similarly to the subjects of for-to clauses and of subjunctive clauses in Romance) has to be disjoint from the higher subject:

- (i) *he_i hates him_j being hanged.

We will discuss non coindexing constraints at some length in Chapter V, however we will not have any substantive proposal to make about phenomena like the one illustrated in (i). Bresnan (1982) suggests that they are cases of obviation (her proposal is specifically made in relation with for-to clauses). In GB approaches the definition of governing category is adjusted so as to make the matrix S in (i) and related cases the governing category for the subject of the embedded S.

²³Category changing rules of this sort are to be regarded as housekeeping rules, since they by their very nature cannot be provided by universal grammar. Also, applying the embedding map ' ' to eventuality-functions is a slight abuse of notation because, strictly speaking, eventuality-functions are represented as being of type $\langle e, e \rangle$ and the embedding map ' ' is not defined over such entities. However, this is an innocuous abuse for the reasons given in fn. 17.

²⁴For relevant discussion on this issue, see Hellan 91980) and Partee (1983).

²⁵See on this e.g. the discussion in Williams (1983).

²⁶Rizzi was born in Genoa and grew up and lives in northern Italy; I was born and grew up in Rome.

²⁷Rizzi individuates a third property which has to do with subjacency. Rizzi's data, however, does not hold for my dialect. Also the issue of how the constructions under consideration interact with subjacency is tangential

to the explanation of the PA-property and the WHE-property also on Rizzi's account. So, I believe that our discussion will not be effected in any substantial way by disregarding the subjacency problem.

CHAPTER IV

CONTROL AND SEMANTIC STRUCTURE

In the preceding chapters we have argued on the basis of semantic and syntactic (distributional) characteristics of infinitives and gerunds that these constructions should be regarded as being VPs (or, in the case of some gerunds NPs derived from VPs) associated with nominalized properties. Henceforth, we will refer to constructions of this sort with the label 'verbal arguments' (i.e. VPs that occur in argument position). Now, there are in English, as well as in many other languages, a number of constructions where verbal arguments are understood as if they had a subject. They are illustrated in what follows:

- (1) a. John tried to get a job
- b. Bill forced John to go to Alaska for a vacation
- c. Bill believes John to have gone to Alaska
- d. John thinks it is important to be at the meeting

In (1) the NP John bears various grammatical relations to the sentence in which it occurs (in (1a) and (1d) it is the subject, in (1b-c) the object) and acts simultaneously as the understood subject of the infinitival complement (though not necessarily so in (1d)). In the tradition of generative grammar, control theory deals with the principles that determine the ways in which a constituent may be selected as the understood subject of verbal arguments.

In particular, (1a, b, d) are known in the literature as Equi structures while (1c) as a raising (to-object) structure. Raising structures are characterized by the fact that the NP controlling the relevant verbal argument does not seem to be a 'logical' argument of the matrix verb, as the following examples illustrates:

- (2) a. John believes it to snow
 b. *John forces it to snow

The dummy it (a typical example of 'virtual' or 'non logical' argument) can occur as object of believe but not of force. Control theory has represented and represents one of the major concerns of current linguistic debates also because it has important consequences for linguistic metatheory in general. It has been argued, for example, that a theory of control has to make crucial use of phrase structural configurations (Williams (1980), Chomsky (1982), Manzini (1982), Koster (1981)) or of grammatical relations (e.g. Bresnan, 1982). Control theory has also represented one of the traditional bases for the claim that it is necessary to posit the existence of phonetically unrealized subjects either in the syntax (GB and related approaches) or in the semantics (LFG, many versions of MG).

Within the framework we are adopting here, verbal arguments do not have subjects neither in the syntax nor in the semantics, and reference to phrase structural

configurations or to grammatical relations in syntactic and semantic rules is, in general, disallowed. Thus we are forced to look in different directions to provide the basis of a theory of control. In fact, we will try to argue that the crucial empirical generalizations having to do with control phenomena can be captured in a satisfactory way in terms of a very simple set of conditions on what can constitute a proper model (in the sense of model theory) for natural languages. We will try to provide a truth conditionally explicit approach to control that does not have any of the characteristics mentioned above (i.e. crucial appeal to phrase structural geometry or to grammatical relations, postulation of null subjects), and try to compare it with other available alternatives. Our approach can be regarded as an attempt to put together the guiding principles of semantically based theories of control (especially, Bach 1979) with ideas developed within LFG (Bresnan 1982).

For our purposes it is convenient to distinguish two basic kinds of control structures: argument control and modifier control. By argument control we will refer to constructions where a predicative expression lacking an overt subject (i.e. the controllee) is semantically an argument. In this sense (1a-d) are all cases of argument control. Whatever the syntactic constituency of infinitives may be, there is in fact general agreement on the

hypothesis that the semantic function-argument structure of, for example, (1) looks more or less as follows:

(3) try' (get a job')(j)

Where try' is a function and get a job' and j are its arguments. Hence, what we call argument control includes all the cases that within the standard theory were analyzed as Equi, Super-Equi and Raising constructions.

By contrast, modifier control concerns structures where the controlled item is (part of) a modifier, which semantically are analyzed usually as functions. Modifier control concerns, therefore, constructions such as purpose and rationale clauses, adjuncts, infinitival relatives and the like. A consideration of modifier control phenomena goes beyond the scope of the present work, and we will consider only argument control. Throughout, in speaking about control (with no further qualifications) we will mean argument control (unless otherwise specified).

1. Relevant Data.

In what follows we will review some of the facts that any adequate theory of control must deal with. In doing so we will draw from various sources (especially Williams, 1980 and Bresnan, 1982) although the empirical typology of control phenomena we will present differs somewhat from most of these sources. In particular we will suggest that

control phenomena form three natural classes that we call obligatory, semi-obligatory and prominence control.

1.1. Obligatory control.

What we call obligatory control phenomena correspond more or less to what Williams (1980) calls by the same label (though extended here as to include raising structures) or to what Bresnan (1982) calls 'functional control'. Relevant examples are the following:

- (4) a. John tries to leave
- b. Bill persuades John to leave
- c. John promises Bill to leave
- d. Bill believes John to leave
- e. John seems to leave

- (5) a. Bill considers John a friend
- b. Bill thinks of John as a friend
- c. John strikes Bill as pompous
- d. John painted the room red

- (6) a. John likes drinking beer
- b. Mary accuses John of being a communist

There are at least six properties that I think make the constructions exemplified in (4)-(6) a natural class. They can be identified in the following terms.¹

First, the relation between the controlling NP and the controlled verbal argument is strictly local. In particular, it can never cross an S boundary:

- (7) a. *Mary believes that John tried to kill himself
- b. *Mary believes that Bill strikes everybody as proud of herself.
- c. *Mary believes that John likes making a fool of herself

Second, the understood subject of the constructions in question can never receive an arbitrary (i.e. generic or contextually specified) interpretation.

Third, the controlling NP is uniquely determined, even when there is more than one potential controller. For instance objects can obviously be obligatory controllers (see e.g. (4b), (5a) or (6b)). However there are verbs that have objects but do not allow them to control a verbal argument in their complement structure:

- (8) a. *John promised Mary to wash herself
 b. *John strikes Mary as proud of herself
 c. *Mary accused John of killing herself

This means that the controller must bear a specific Θ -role with respect to the matrix predicate (i.e. must be the agent or the goal, etc.) We will call this property 'thematic uniqueness'.

Fourth, the controlled verbal argument can't have more than one controlling NP. This phenomenon is known in the literature as lack of 'split antecedents':

- (9) a. *John persuades Mary to wash themselves
 b. *John considers Mary proud of themselves
 c. *John accused Mary of destroying themselves

Fifth, there has to be a controller. The controller cannot be 'implicitly understood'. Relevant examples are the following:

- (10) a. *It was tried to leave.
 b. *John considered stupid
 c. *Bill accuses of being a communist

There two important generalizations somehow connected to the last point. The first one, known as Visser's generalization concerns the impossibility of passivizing subject control verbs:

- (11) a. *Mary was struck as pompous (by John)
 b. *Mary was promised to come (by John)

Notice that passive with promise is fine when no control is involved:

(12) Mary was promised that she would get the job
 The second related generalization, attributed by Bresnan (1982) to Bach, is represented by the impossibility of detransitivizing object control verbs. Transitive verbs can undergo a lexical process that turns them into intransitives, but object control verbs never can. This is illustrated in what follows:

- (13) a. Mary promised Bill to come
 b. Mary promised to come
 (14) a. Mary persuaded Bill to come
 b. *Mary persuaded to come

That Bach's generalization has to do with control is clearly illustrated by the following examples, where detransitivization is acceptable where there are no control phenomena involved:

- (15) a. John painted something (from Bresnan, 1982)
 b. John painted
 c. John painted it red
 d. *John painted red

- (16) a. I taught them that the earth is round
 b. I taught that the earth is round
 c. I taught them to speak German
 d. *I taught to speak German

There is of course some redundancy in the empirical typology we have set up, and part of the task of our theory will be to cut down on this redundancy by deriving all these properties in terms of general principles. For convenience, let us sum up in a schematic form the properties of obligatory control constructions:

- (17) a. locality
 b. lack of arbitrary reading
 c. thematic uniqueness
 d. lack of split antecedents
 e. obligatory presence
 f. patterning according to Visser's and Bach's generalizations

1.2. Semi-obligatory control.

What seems to be the crucial characteristic of semi-obligatory control constructions is that they have essentially all of the properties of obligatory control constructions except for (17e). In other terms, if a specified argument position of argument control predicates is lexically filled, then they pattern like obligatory control predicates; but such a position doesn't have to be filled. Relevant examples are the following:

- (18) a. John decided to have dinner late
 b. Mary signalled Tom to cross the road
 c. to have dinner late is tough for John
 d. Mary recommended reading War and Peace to
John

It is evident that in these examples an arbitrary interpretation for the understood subjects of the verbal arguments is impossible (17b). Furthermore, the controller is individuated as bearing a specific Θ -role to the matrix predicate; other potential controllers not bearing that Θ -role are excluded (17c):

- (19) a. *Mary signalled Tom to kiss herself
 b. *Mary recommended not making a fool of herself to Tom

It can also be seen that the control relation has to be strictly local:

- (20) a. *Mary said that John decided to kill herself
 b. *Mary said that to enjoy herself was tough for John

Constructions such as those in (18) also sound pretty bad if one tries to force on them a 'split controllers' reading (17d):

- (21) a. *Mary said that John decided to wash themselves
 b. *Mary signalled Tom to entertain each other
 c. *Mary said that to enjoy themselves was tough for John
 d. ?? Mary recommended reading each other's papers to John

However, the controlling NP appears to be optional with these constructions:

- (22) a. It was decided to leave
 b. Mary signalled to cross the road
 c. to have dinner late is tough
 d. Mary recommended reading War and Peace

An immediate consequence of this is that these predicates seem to provide counterexamples of Visser's and Bach's

generalizations. In (22a) we have a subject control verb that does passivize, contra Visser. In (22b) we have an object control verb that does undergo detransitivization, contra Bach. Moreover one can easily verify that whenever the designated controller is absent also the properties (17a-d) observed so far vanish all of a sudden. It is, for instance, pretty clear that the examples in (22) allow for some sort of arbitrary interpretation of the missing subjects of the relevant verbal arguments. Furthermore, in cases such as (22) the controller can appear indefinitely far away from the controlled verbal argument:

- (23) a. The committee announced that it was decided to leave
 b. The students were very happy when their teacher signalled to go home
 c. Mary said that to have dinner that late was tough
 d. The students of that class told me that their teacher recommended reading War and Peace

Finally, when the designated argument is absent, the controller does not have to bear any specific Θ -role and 'split controller' phenomena become possible:

- (24) a. John suggested to Bill to decide to leave together
 b. John suggested to Bill to signal to leave together
 c. Mary told Bill that to live together would be tough
 d. The students of that class told me that their teacher recommended leaving together.

The presence of a group adverbial (together) in (24) shows that the verbal argument modified by the adverbial has a

plural understood subject. It is possible that the understood subject be construed as the 'group' or 'plural individual' made up of the two other NPs occurring in the sentence.

So what seems to be going on with this class of predicates is that if a certain argument is present then it obligatorily controls the verbal argument (with all the relevant characteristics of the obligatory control relation). However, this argument may not be overtly present but rather implicit or understood. In the latter case one has the effect of arbitrary, long distance or split antecedent control. This strongly suggests that such a phenomenon arises precisely because what the implicit argument is may be recovered from the context. So for instance, the long distance control phenomena in (23) may arise because the distant controller is reconstructed as the controlling argument of the relevant verbs. Similarly for the split control phenomena illustrated in (24). The problem is how to make precise this idea and how to account for the similarities and differences between this class of predicates and the obligatory control ones. This is part of the task that any adequate theory of control seems to face.

1.3. Prominence control.

A third class of predicates doesn't seem to have any of the properties listed in (17). What the controller of verbal arguments is going to be seems to depend on a variety of factors that go beyond the limits of formal semantics (at least in the classical, Montagovian sense of this word). The controller of verbal arguments in the constructions we are going to consider can be some NP in the local environment of the predicate, some NP in the neighboring structure or the preceding discourse, or even some entity in the extralinguistic context. All that seems to be required is that the controller be somehow sufficiently prominent. Examples are the following:

- (25) a. John told me that to make a fool of himself
bothered Sue
disturbed
amused
- b. John told me that to make a fool of himself
was dangerous for his wife
- c. The commissioner denounced drinking beer at
ball games (from Horn, 1975)

The examples in (25a-b) show that the infinitival clauses in the subject position of bother and related predicates can be controlled by an NP which occurs in some superordinate position. Notice that it doesn't have to be the first NP to the left (or the first C-commanding NP, for that matter). "Prominence" seems to be, thus, the sole criterion. Example (25c) shows that the verbal argument in the complement structure of denounce and related verbs

(discuss, imagine, etc. see Thompson 1972 for a list) can be controlled by some contextually prominent entity or have a generic interpretation. Notice that the same is possible with bother or be dangerous type predicates. For instance:

- (26) I know that my wife is courageous and adventurous. But getting herself into such a dangerous spot really scared me.

I am not aware of any predicate of the type exemplified in (25c) (i.e. that can escape subject control) which takes an infinitival (rather than a gerundive) complement.

Cases of the latter sort, however, abound in Italian, precisely with the same sort of predicates as discuss or denounce in English. For example:

- (27) Il commissario ha denunciato vigorosamente
prendere tangenti per appalti pubblici
The commissioner denounced vigorously to accept
bribes for public contracts.

Examples (25b) above show that 'dative' complements of adjectives like dangerous (unlike the 'dative' arguments of tough-like adjectives) don't have to control the verbal argument in subject position (contrast 25b with 20b). Of course being in the same local environment as the verbal argument seeking control makes an NP quite prominent and therefore a strong candidate for the role of the controller. This is especially evident when sentences with bother or dangerous type predicates are uttered in isolation:

(28) to travel by plane scares me

But the examples in (25) show quite clearly that 'locality' is by no means sufficient to license control. Given this relative freedom in the control properties of the predicates under consideration, the lack of thematic uniqueness and the possibility of split antecedents is of course no surprise:

(29) John told me that making fools of themselves at that party really bothered his wife.

The understood subject of the gerund in (29) can be the group constituted by John and his wife.

The empirical typology of (argument) control phenomena discussed so far gives us, I think, a good overview of the kind of problems that control theory has to face and identifies in an acceptable way the crucial set of properties of each class of cases. Of course, control phenomena are highly complex and there are many subtleties and details that we haven't considered. Some idealization of the data, however, is always necessary to get an interesting theoretical account started, and the data-idealization presented here is, I believe, not below the standards commonly accepted in much recent literature on the argument.

2. Control as a Semantic Entailment.

I would like to put forth the idea that our intuitions about there being an understood subject of verbal arguments in control structures is grounded on certain entailments licensed by some predicates. However, before getting into details there are a few preliminaries relevant for subsequent discussion that have to be filled in. I am going to assume without discussion that there exists a standard way in which arguments of a verbs are encoded in argument structures.

As in standard MG, a verb V is going to be translated as an n -place propositional function V_n . In a structure of the form $V'(x_1) \dots (x_{n-2})(x_{n-1})(x_n)$, x_n is the argument that corresponds, generally, to the subject, x_{n-1} the one that corresponds generally to the object, x_{n-2} the one that corresponds to second objects or to obliques, and so on (see Dowty, 1978, 1982, Bach, 1980a for arguments and relevant discussion). I am also going to assume, in its essentials, the analysis of passive developed in Bach (1980a) modified as in chapter 1 sec. 2.2. of the present work. Passive morphology will, accordingly, be regarded as something that maps TVPs into IVs, along the following lines:

- (30) a. if $\alpha \in \text{TVP}$ then $\text{PAS}(\alpha) \in \text{IV}$
 b. if $\alpha \in \text{TVP}$, then $\text{PAS}(\alpha)$ translates as $\text{PAS}'(\alpha')$
 c. $\Box \text{PAS}'(\alpha')(x) \leftrightarrow \exists y[\alpha'(x)(y)]$

By-PPs are then regarded as predicate modifiers. This will turn out to be crucial for our analysis and we have argued for it independently in ch. I sec. 2.2. Other details of the analysis of passive, however, are not crucial. So, for instance, it is not crucial whether passive is a purely lexical or both a lexical and a syntactic process.

With this in mind, let us now turn to our analysis.

2.1. Conditions on admissible models of English.

Consider the following relation:

$$(31) R = \lambda P \lambda x [P \text{ is the property of being equal to } y]$$

Thanks to our competence as speakers of English we understand quite well what kind of relation R is. The extension of R will be a set of ordered pairs $\langle P, y \rangle$ such y is an old entity and P is the property of being identical to that entity. From our understanding of the meaning of R , we know that whenever P and y stand in the relation R , y must have the property P , since P is the property of being identical to y , and y is of course selfidentical (aren't we all?). R is a trivial example of a relation between a property P and an object x such that in order for R to be satisfied by P and x , x must have P . In natural languages, there are relations between properties and (one or more) individuals. Some of them are just like R . That is to say, whenever they are satisfied by a property P and

one or more individuals x, y, \dots it follows that one of these individuals must have P . Other relations in natural languages are not like R . I would like to claim that this very simple fact is the central mechanism of control phenomena. Consider the verb enjoy in:

(32) John enjoys playing tennis

Enjoy can be analyzed as a relation between individuals and activities. Such a relation might be paraphrased roughly as follows: whenever x does P , x is happy. From the meaning of enjoy it follows that if x is in the enjoy relation with a (nominalized) property, say playing tennis, then x is happy when x instantiates that property (when he plays tennis). Of course, things might have been different. It is perfectly possible to imagine a relation, call it enjoy*, such that if x and P are in the enjoy* relation, x would be happy for some other individual y to have P . But it so happens that enjoy* is not the right meaning for enjoy, and we might think of various plausible reasons why this is so.

Consider now the verb talk about, as in:

(33) John talks about playing tennis

Talk about, like enjoy, can be analyzed as a relation between an individual and an activity. But contrary to what happens for enjoy, an individual can talk about playing tennis without ever actually playing. So, an

individual x and a property P can stand in the talk-about relation while x does not have P in any circumstance.

This way of looking at control phenomena is very close in spirit to Montague's original view of the problem. In PTQ, verbs like try are analyzed semantically as, essentially, relations between properties and individuals. Nothing else is specified, thereby entailing that where the 'missing subject' is to be located is an internal semantic property of the verb try. Just as John has killed Bill entails Bill is dead, so John tries to run away entails if John's attempt succeeds, John runs away. On this view, therefore, control phenomena are part of the entailments of natural language predicates, in particular of predicates that among their arguments have nominalized properties (like enjoy or try).

What is the general tool for expressing entailments? The answer is: meaning postulates. Meaning postulates represent of course a very powerful mechanism that certainly needs to be somehow constrained; they per se do not represent a substantive linguistic theory, but merely a notation within which such a theory can be formulated. I know of no semantic theory (model theoretic or non model theoretic) that doesn't have something like meaning postulates. Within the present approach, they are regarded as conditions on admissible models for a language like English. They specify what form a possible model

should have for it to be suitable to interpret English, given the conceptual endowment of the speakers of that language. In this sense, if control theory is to be stated as a set of meaning postulates, then it could be legitimately regarded as a part of what Emmon Bach calls 'English metaphysics'. We have of course yet to show that all the relevant empirical generalizations about control can be optimally captured by means of meaning postulates, which is what we are going to try to do in what follows. If our attempt is successful to some degree, it would provide us with a very interesting example of how model theoretic semantics can offer linguistically significant and explanatory models.

Consider a verb like force. According to our assumptions, force will be a relation between a property and two individuals. The meaning of force is such that the following entailment holds true:

$$(34) \text{force}' (\wedge P)(x)(y) \rightarrow P(x)$$

For instance, if John is in the force-relation with Mary and the property of going to the movies (i.e. John forces Mary to go to the movies), it necessarily must also be the case that Mary goes to the movies. This is what force means: to act on somebody in such a way that he does something as a result of our pressure. So, (34) makes explicit one of the structural characteristics of the meaning of force. This much appears to be completely

straightforward. We might then be led to think that all obligatory control verbs are such that they have an entailment similar to (34). In fact, we could easily express the form of all such entailments in the following way:

$$(35) \quad \alpha(x_1) \dots (P) \dots (x_n) \rightarrow P(x_i) \quad \text{where } i \leq n$$

Of course, as it is, (35) is too simple to be true and in fact is meant to be just suggestive of what will be the leading idea of the present approach. We will now gradually work our way towards more adequate formulations of the principle that (35) is intended to exemplify.

A clear inadequacy of (35) has to do with the intensional nature of most, if not all, obligatory control predicates. It so happens that force has a factive force. But most control verbs are not like that. Consider try, for instance. If I am in the try-relation with the task of writing a paper (i.e. if I am trying to write a paper), it does not follow that I actually write the paper. However, in the possibly counterfactual situations where all my attempts are successfully carried on, the paper will be written. This is part of what try means. Similarly, if I promise to convince you that model theory is the answer to your problems, it does not follow that you actually are convinced. But there are situations where you are, namely those where all promises are fulfilled (no matter how great the metaphysical distance

between these situations and the actual one may be). In other words, I will follow the by now well established theory according to which intensional verbs or, as they are also called, 'verbs of attitude' involve a consideration of alternative states of affairs. The entailment in (35) must, therefore be subject to a modal qualification.

2.2. On the modal nature of control predicates.

There are two ways to go, if we want to build a modal qualification into a principle of control, as far as I can see. The first one can be described along the following lines. We could state a meaning postulate on a verb like, say, try which has the following form:

$$(36) \text{ try}'(P)(j) \leftrightarrow \text{try}'(^{\sim}P(j))(j)$$

What (36) says is that John is in the try-relation with an action (state, etc.) P iff he is in the try-relation with a state of affairs in which he has P. E.g. John tries to jog iff he tries to bring about a situation where he actually jogs. Notice that in our framework actions (i.e. nominalized properties) and states of affairs (i.e. nominalized propositions) are two different sorts of individuals. So the very same propositional function (try' in (36)) can apply truthfully to both actions and states of affairs. We don't need a try₁ and a try₂ to state a condition like (36) (as it would be the case in Montague's IL). We can have the very same try to the

right and to the left of the biconditional in (36). (36) is a closure condition on the meaning of try. It says that whenever an ordered pair of the form $\langle [x], [P] \rangle$ is in the extension of try', so it is also the ordered pair $\langle [x], [P(x)] \rangle$. According to this approach, then, try would be both a relation between individuals and actions and between individuals and states of affairs. Such a relation would be restricted in such a way that x bears it to an action P iff x bears the same relation to a state of affairs where x has P . It is fairly evident that such an approach could be generalized to all control predicates and would provide us with a straightforward way of accommodating their modal nature.

A second way to go in dealing with the modal character of control predicates would be to resort to a possible world approach to modalities. Even if such an approach is known to be defective in several respects, it still provides the most comprehensive and explicitly worked out framework to deal with modal relations of various sorts. Our framework is not a possible world framework, but it is compatible with it and allows us to incorporate into it all the important results that possible worlds semantics has given us in the last twenty years or so. To give a hint of how we might proceed, consider:

$$(37) \text{ try}(P)(j) \leftrightarrow \Box_{\text{try}, j} P(j)$$

(37) might be read as saying that John tries to do P iff in all the situations compatible with John's attempt he does P. The necessity operator on the right hand side of the biconditional in (37) must be a context sensitive one: the set of relevant situations we want to look at are determined by contextual factors such as what is the kind of relation involved and which individuals are involved in it. This context sensitivity is represented by indexing the necessity operator to the relevant parameters (try and j in (37)). What we would need is a way of generalizing (37) to all control predicates. Kratzer (1981) has provided a highly interesting account of the German modal system and its relation to the context of use. Her framework can be fruitfully extended, I believe, to deal with the problems that concern us here.

It is unclear to me whether the two approaches just sketched are different in principle as far as our goals (i.e. providing some representation of the modal nature of control predicates) are concerned. In what follows I will show how to extend Kratzer's approach to cope with the facts we are dealing with, for such an approach is very well worked out and has been shown to give good results in connection with other aspects of natural language semantics. But I cannot detect any reason of principle why the first approach outlined above could not give results

analogous to the ones we will get in developing a theory of control based on Kratzer's theory of modals and context.

According to Kratzer, there are at least two components to modalities in natural languages. One is the 'force' of a given modal statement (e.g. whether it is a necessity or a possibility), which Kratzer calls modal relation. The other is the conversational background, which is responsible for the scenario that licenses a modal statement. So, to repropose one of Kratzer's examples, consider a situation where somebody called Smith has been murdered. The inspector on the basis of the relevant circumstances might utter:

(38) the murderer must be Pierre

The inspector is justified in uttering (38) iff that Pierre is the murderer follows from what he (the inspector) knows about the circumstances of the crime. What the inspector knows about the crime represents the conversational background relevant for the evaluation of (38); the modal relation of (38) is simple necessity.

Clearly the notion of conversational background is a highly context sensitive one. Kratzer proposes to analyze it as a function that assigns to each world (or context of use) a set of propositions. There are various types of conversational backgrounds. Among others, Kratzer points out the following, which are directly relevant for us:

- (39) a. deontic conversational background: functions that assign to each context of use a set of propositions p such that p is commanded by somebody to somebody else.
- b. teleological conversational backgrounds: functions that assign to each context of use a set of propositions p such that p is somebody's aim.
- c. buletic conversational backgrounds: functions that assign to each context of use a set of propositions p such that p is what somebody wants or desires.

In our theory we will represent contexts of use as world-times pairs. This is admittedly inadequate, but to get into issues of contexts, indexicals, etc. would take us too far afield. On the basis of the notion of conversational background it is possible to define relevant modal relations as follows:

- (40) a. a proposition p is a simple necessity in a world-time $\langle w, j \rangle$ with respect to a conversational background f iff p follows from $f(\langle w, j \rangle)$ (i.e. if, for all $q \in f(\langle w, j \rangle)$, $w(q, j) = 1$, then $w(p, j) = 1$).
- b. a proposition p is a simple possibility in a world-time $\langle w, j \rangle$ with respect to a conversational background f iff p is compatible with $f(\langle w, j \rangle)$ (i.e. there exist a w' and j' such that for all $q \in f(\langle w, j \rangle)$, $w'(q, j') = 1$ and $w'(p, j') = 1$).

Let M be a variable ranging over modal relations such as those just defined and let a, b, \dots range over conversational backgrounds. We can then stipulate the following:

- (41) a. if $\varphi \in ME$, then so is $M_a \varphi$.
- b. a formula of the form $M_a \varphi$ is true in a world w at a time j iff $\llbracket \varphi \rrbracket$ is a M (i.e. either a simple necessity or a simple possibility) in $\langle w, j \rangle$ with respect to the conversational background a .

(41) is not fully formalized but is sufficiently explicit for our purposes and a more formal implementation of it in the framework of chapter I sec. 3.4. is purely a matter of routine. Now we may say that each intensional verb (and, in particular, each control verb) selects a certain type of conversational background and a modal relation. This is a simple and straightforward extension of what Kratzer proposes for 'pure' modals (i.e. can, must, etc.) which, nota bene, can be regarded as control verbs in our sense. So for instance, we might have the following:

- (42) a. force: conversational background: deontic
 modal relation: necessity
 b. allow: conversational background: deontic
 modal relation: possibility
 c. try: conversational background: buletic
 modal relation: necessity
 etc.

Let us call the ordered pair constituted by the type of conversational background and by the modal relation that a verb selects the modal frame of that verb. We have now sufficient apparatus to formulate the modal qualification that we need for our control principle. In its general form, it can be formulated as follows:

- (43) Control Principle (CP) -- first version;
 $\Box \alpha(x_1) \dots (P) \dots (x_n) \leftrightarrow M_a P(x_i)$
 where $1 \leq i \leq n$ and $\langle M, a \rangle$ is the relevant modal frame associated with α .

Let us give some instances of CP in order to exemplify how it works:

- (44) a. $\Box \text{force}'(\wedge P)(x)(y) \leftrightarrow M_a P(x)$
 b. $\Box \text{allow}'(\wedge P)(x)(y) \leftrightarrow M_a P(x)$
 c. $\Box \text{try}'(\wedge P)(x) \leftrightarrow M_a P(x)$

Consider (44a) first. Force selects deontic conversational backgrounds and simple necessity. Furthermore, for any x and y , if y forces x to do P , then the "relevant" conversational background will select in each context $\langle w, j \rangle$ the set of propositions p such that p is what y imposes on x .² So, (44a) ends up meaning something like: y forces x to do P iff in all the situations compatible with what y imposes on x , x does y . Since force is a factive, the set of situations compatible with what y imposes on x must include the actual world (i.e. the conversational backgrounds selected by force must be 'realistic' in Kratzer's sense). Consider next (44b). The conversational backgrounds selected by allow are also deontic, but the modal relation is simple possibility. Then (44b) will end up meaning something like: y allows x to do P iff there exist a situation compatible with what y imposes on x where x does P . A consequence of this approach is that it allows us to capture the validity of the following sentences (with respect to the same contexts):

- (45) a. If x allows y to P , then it is not the case that x forces y to not P
 b. If x forces y to P , then it is not the case that x allows y to not P

Consider finally (44c). Try selects buletic conversational backgrounds. So, (44c) says something like: x

tries to P iff in all the situations compatible with x's aims (i.e. in all the situations where what x tries will eventually succeed) x does P.

It should be, at this point, pretty clear that the present approach is perfectly general and extends straightforwardly to all control predicates. Moreover it exploits a machinery (i.e. a theory of modalities) that has been argued to be independently needed and, details of execution aside, is truth conditionally fully explicit. CP is thus a strong candidate as a principle capturing the relevant structural properties of control predicates.

It might be worth discussing a little how an approach along the present lines allows us to deal, without further stipulations, with cases where a control predicate occurs in embedded position. Consider first negation. Irrelevant details aside, the following example illustrates how our approach would work.

- (46) a. John doesn't try to jog
 b. $\neg \text{try}'(\wedge \text{jog})(j)$
 c. $\neg \Box \text{try}, j \text{ jog}'(j)$

(46b) would be the translation of (46a); the relevant instance of CP would yield (46c) as a wff equivalent to (46b). So our intuition of there being an understood subject of the infinitive in (46a) would receive a straightforward account. Consider next a modal verb such as believe. A relevant example would be:

- (47) a. Mary believes that John tries to jog
 b. believe'(^try'(^jog')(j))(m)

On one sense of believing, if Mary believes that John tries to jog, then she is disposed to act as if it were actually the case that John tries to jog. But then she will also be disposed to act as if it were the case that in all the situations compatible with John's aims, he jogs, for the former is the case in all and only the situations where the latter is the case. In other words if Mary is committed to the truth of try'(^jog')(j), then at least in one sense of believe she is also committed to the truth of $\square_{\text{try, j}} \text{jog}'(j)$, for in virtue of the relevant meaning postulate, these two formulae have the same truth conditions. So jog in (47a) will have an understood subject, via such a chain of entailments. The same holds, mutatis mutandis, for any modal context no matter how deeply embedded in it the control verb is. What we do in each case is to go into the relevant set of possible circumstances to which the (possibly nested) modalities are semantically related and then 'run' the meaning postulate. Since a meaning postulate holds for all possible circumstances, it will hold within the circumstances that characterize semantically a modal relation, no matter how complicated. Even if a possible worlds treatment of belief and related contexts is inadequate, it is admittedly adequate to capture modalities as

dispositions to act. So this aspect of a possible world approach will have to be maintained by any theory of modalities that seeks adequacy. Such an aspect will also be sufficient to warrant that our meaning postulate will have the desired effects. Our intuitions about control relations are grounded on the way in which such intuitions are established to work in the simple (i.e. unembedded) cases by CP. A similar line of argumentation can also be used in connection with questions, commands, etc.

There are ways in which the modal nature of control verbs is only approximately captured by something like CP. For example, Kratzer argues that modal frames have to be enriched with a further parameter, namely the ordering source. Ordering sources are conversational backgrounds that do not determine the set of accessible circumstances, but rather which members of the set of accessible circumstances are to be regarded as the most accessible. In terms of the notion of ordering source, Kratzer is able to provide an account of an intricate network of modal notions, including counterfactuals. These extensions, however, are beyond the scope of the present work, and we will disregard them in what follows. Similarly, we will neglect other important issues, such as the intrinsic 'vagueness' of modalities in natural languages and the like.

To sum up we have proposed a first general principle of control that is able to capture in an interesting way the often noted modal nature of control predicates. CP is to be regarded as a lexical condition on the meaning of certain predicates. Formally, it is a meaning postulate schema; by setting α and i at specific values we get actual meaning postulates (like those exemplified in (44)). Each control predicate is associated with a modal frame which is, again, part of its lexical meaning. So once the value of α in CP is established, general principles determine automatically what M and a (i.e. the relevant modal frame of α) are going to be.

However, we must hasten to recognize that something like (43) is still in various ways unsatisfactory as a general principle of control. For example, (43) states that the controller of some predicate α is its i^{th} argument. This is achieved by stating the rather brute force restriction that $1 \leq i \leq n$. It would be as easy to state some other condition requiring that the controller be not an argument of α . So something seems to be missing from the way CP has been so far formulated. Furthermore by merely indicating in terms of an index what the controller is going to be we would be intrinsically unable to represent the relation between the way in which the controller is selected and the thematic structure of the relevant control predicate. We now turn to a more tho-

rough consideration of these potential problems that our first formulation of CP (i.e. (43)) seems to face.

3. Control and Thematic Structure

In the tradition of linguistic semantics, thematic relations have been found to be useful notions in describing regularities having to do primarily with lexical properties of various items (Gruber, 1965, Jackendoff, 1972, 1976). Consider, for instance the following sentence:

(48) John gives Mary a book

In a situation where (48) obtains there is something (a book) that in some suitable abstract sense moves from John to Mary. The protagonist of this movement is called the theme of the giving act described by (48); the origin of the movement (i.e. John) is called the source, and its terminal point (i.e. Mary) is called the goal. Furthermore, one might argue that John is also what causes the movement of the book to take place, and hence is the agent of the giving. Goal, source, theme, agent (together with experiencer, beneficiary, place, manner, etc.), thus, seem to characterize or classify possible relations between a predicate (give in (48)) and its arguments (John, Mary and a book in (48)). The exact formal nature of thematic (θ from now on) relations is still an object of debate and these notions have never been granted theoretical status

(in any systematic way) within the tradition of formal semantics, to my knowledge. The status of Θ -relations is especially relevant for our present purposes for it has been argued quite convincingly (e.g. Jackendoff, 1972) that significant empirical generalizations about control can be stated in terms of them. For instance, it seems to be the case that themes tend to play the role of controllers more often than goals, goals more often than sources and so on. So if a control predicate has both, say, a theme and a source, the unmarked case tends to be that the theme is the controller. Consider:

- (49) a. John regards Mary as a friend
 b. Mary strikes John as reliable
 c. John forces Mary to go away
 d. John promises Mary to go away

In (49a-c) Mary is the theme and it controls the verbal argument irrespective of whether it is the subject or object. In (49d) Mary is not the theme and it doesn't act as a controller. There seems to be something right about this generalization. The problem that we face, if we want somehow to make it precise and to implement it in our approach to control, is to determine how Θ -relations might be regarded within a model theoretic setting. From an intuitive point of view, a plausible first guess might be to regard them as classificatory principles on ways in which entities may participate to events, processes or states. The next section represent a first approximation

attempt to make this idea more explicit. We will not be concerned with the specific content of specific θ -relations but solely with their logical structure.

3.1. θ -roles and complexes: an excursus into English metaphysics.

In Chapter III Sec. 2.3. we have pointed out that one must assume the existence of events, states and processes (collectively referred to as 'complexes') and that these entities are the denotata of so called 'perfect nominalizations' (see e.g. Vendler, 1967) of which what we call CN-gerunds (e.g. Brutus' killing of Caesar) are a case. We haven't said anything so far about how complexes might be represented in our semantic framework; now we have to, if we want to make more explicit the idea that θ -roles are ways of classifying how entities group up into complexes.

How complexes are individuated depends on the conceptual apparatus that we use to look at them. We can represent conceptual apparatuses as a certain set of (n-place) properties. Let us call NPr (for 'natural property') such a set. Not any old property that the combinatorial mechanisms of our theory (i.e. IL_*) allows us to individuate is also going to be a member of NPr. Natural properties are probably best regarded as socio-biological capacities to discriminate things (see on this Cocchiarella, 1983, ch. 2).³ For the present

purposes, we can assume the set NPr to be simply given. In terms of it we can define what a complex is. As a rule of thumb, the denotation of any basic English predicate (except for very general ones such as thing) will be a NPr . But members of NPr are not to be regarded as 'simplex' in any metaphysical sense and various inter-relationships may obtain among them. Given the set NPr , let us define its 'complement' \overline{NPr} as follows:

$$(50) \overline{NPr} = \{ r : \exists h \in NPr \text{ such that } r = \lambda u_1, \dots, \lambda u_n \cdot \sim h(u_1) \dots (u_n) \}$$

\overline{NPr} thus contains all and only the negations of members of NPr . Let $NPr^C = NPr \cup \overline{NPr}$. In terms of NPr^C it is possible to define the notion of ur-complex as follows:

$$(51) \text{ an ur-complex is an } n+1 \text{ tuple } \langle p^n, x_1, \dots, x_n \rangle \text{ such that } p^n \in NPr^C \text{ and } x_i (1 \leq i \leq n) \text{ is in } U \text{ (i.e. the domain of individuals).}^4$$

In symbols: $ur-C = NPr_n^C \times U^n$ (where NPr_n^C is the set of n-place members of NPr^C)

So an ur-complex is going to have a property (or its negation) as its first constituent, and n-individuals as its other constituents. This provides us with a simple way of representing individuals-in-relations. It is pretty evident that our notion of ur-complex corresponds closely to the notion of situation-type as currently developed in situation semantics (see e.g. Barwise and Perry 1981). A complex (i.e. an event, state or process) can then be regarded simply as a set of ur-complexes. A θ -role can instead be construed as a partial function from

ur-complexes into individuals. Intuitively, a θ -role maps each ur-complex $\langle P^n, u_1, \dots, u_n \rangle$ into an individual u_i such that u_i is that member of $\langle P^n, u_1, \dots, u_n \rangle$ which bears the role θ , if such an u_i exists; otherwise θ is undefined for that ur-complex. In this way a θ -role such as agent given an ur-complex $\langle P^n, u_1, \dots, u_n \rangle$ would select the individual that plays the role of the agent among the constituents of the ur-complex and otherwise would be undefined. All this can be summarized as follows:

- (52) a. A complex (i.e. an event, state or process) is a set of ur-complexes. In symbols:
 $C \subseteq \mathcal{P}(\text{ur-C})^5$
- b. A θ -role θ is a partial function in $\text{ur-C} \rightarrow U$ such that for any $k \in \text{ur-C}$ if $\theta(k)$ is defined, then $\theta(k) \in k$.

The picture that emerges from all this is then the following. Events and similar creatures can be modelled within our theory as sets of individuals-standing-in-certain-relations. The definition of event is relativized to a certain conceptual apparatus represented by NPr. Since English verbs will take their denotations in NPr, this provides us with a rather substantive hypothesis concerning how the meaning of a verb relates to the meaning of the corresponding (CN-) nominalization. Events are constructed out of verb-meanings, via the definition of ur-complex. Common nouns derived from verbs are going to denote, according to the hypothesis put forth in Chapter

III Section 2.3., sets of events (or, more generally, sets of complexes). So the denotation of deverbal CNs can actually be built from the denotation of the corresponding verb. This is of course sketchy, but, I think sufficiently clear to be promising.

Furthermore, we have proposed to model θ -roles as partial functions on ur-complexes (i.e. individuals-standing-in-certain-relations). Again, this is only a first approximation, but one that seems to do justice to and make explicit the intuition that θ -roles are ways of classifying how things enter into various kinds of situations. Such an approach might have further interesting consequences as it allows us to have a definition of θ -roles that can be applied both to propositions (since ur-complexes are built out of NPr, which provides also the set of values for basic predicative expressions) and, derivatively, to deverbal CNs (since their denotation is also built via (51) and (52a) out of NPr).

The idea is the following. Take a simple sentence like (53a); its translation is going to be as in (53b); the translation of the corresponding that-clause is going to be (53c).

- (53) a. Brutus killed Caesar
 b. kill'(c)(b) (disregarding tense)
 c. ^kill'(c)(b)

The denotation of the corresponding deverbal CN killing (as in the killing of Caesar) is going to be a set of

events. Each event is (modelled as) a set of ur-complexes. So we want each killing event to contain ur-complexes of the form $\langle \text{'kill'}, x, y \rangle$. This is what the semantics associated with deverbals has to guarantee. θ -roles are of course defined on ur-complexes of the form $\langle \text{'kill'}, x, y \rangle$. For example:

- (54) a. $\text{Ag}(\langle \text{'kill'}, x, y \rangle) = y$ (where Ag = agent)
 b. $\text{Pa}(\langle \text{'kill'}, x, y \rangle) = x$ (where Pa = patient)

Hence, θ -roles can be employed to classify function-argument structures in propositions such as (54b) and in deverbal CNs such as the killing.

There are of course many questions that the present approach raises. Here is a sample.

- We have put forth an hypothesis on how propositions might be related to events et similia. However we know how this relation goes only in simple cases. Can we define, in general, what it means for a proposition to identify an event, and, conversely, for an event to support the truth of a proposition?
- Clearly we don't want any old set of ur-complexes to be an event. Can we find a plausible set of constraints concerning how two different ur-complexes can be part of the same event?

Answering these and related questions requires presumably several other dissertations, and I am

having enough trouble with this one. At any rate, what we have done so far accomplishes what we needed (i.e. a model-theoretically acceptable definition of θ -roles) and points toward many other promising avenues of research.

3.2. Further conditions on θ -roles.

So far we have put forth an hypothesis concerning the logical category of θ -roles. This hypothesis has immediate consequences of interest built into it. For instance, given our definition it follows that there can't be two different individuals bearing the same θ -role in an argument complex, since θ -roles are functions. So, in particular, there can't be any predicate P that has, say, two different sources among its arguments. For otherwise the following would have to be the case:

- (55) a. So ($\langle \wedge P, x_1, \dots, x_n \rangle$) = x_i
 b. So ($\langle \wedge P, x_1, \dots, x_n \rangle$) = x_j
 c. $x_i \neq x_j$

But if (55a-c) are simultaneously the case, then θ -roles wouldn't be functions as our definition requires. So one half of what is known in the literature as the ' θ -criterion' is an immediate consequence of our definition of θ -roles. The other half of the θ -criterion, however, not only does not follow but it can't be true on the present approach to θ -roles. For consider the following:

- (56) a. John kills himself
 b. kill (j)(j)
 c. Ag (<'kill', j, j>) = j
 d. Pa (<'kill', j, j>) = j

In (56) b is the translation of a, and John is both the agent and patient of the relevant killing event, which is in full correspondence with what agent and patient are usually taken to mean. If θ -roles are regarded as ways of classifying individuals-standing-in-relations then the idea that one individual cannot carry two distinct θ -roles would amount to claiming that no individual can stand in any relation with him-or herself: a procustean hypothesis, at best. If θ -roles are not something like what we claim they are, then what are they?

Thus far we have provided only what the logical category of θ -roles should be. Further constraints on the structure of these notions can be provided by means of some suitable set of axioms. Minimally, the following are going to be required:

- (57) $\theta 1$ 'consistency'
 $\theta(\langle ^\wedge P, x_1, \dots, x_n \rangle) =$
 $x_i \rightarrow \Diamond P(x_1) \dots (x_n)$
 $\theta 2$ 'completeness'
 $\Diamond P(x_1) \dots (x_i) \dots (x_n) \ \& \ B(x_i) \rightarrow$
 $\exists \theta[\theta(\langle ^\wedge P, x_1, \dots, x_n \rangle) = x_i]$

The first axiom requires that whenever an individual is the bearer of a θ -role with respect to a certain event, then that event must at least be possible (i.e., in some sense, consistent). The second axiom requires that whenever it is possible that an 'ordinary' individual

(i.e. a member of the base B -- see Chapter II Section 3.2.) stands in a relation P (assuming that $P \in NPr^C$) with other individuals, then there has to be some θ -role (though not necessarily a 'named' one) that selects that individual. In other words there have to be enough θ -roles for us to be able to classify all ur-complexes involving 'ordinary individuals' (whence the label 'completeness'). The intuitive content of θ_1 and θ_2 appears to be quite straightforward. Further conditions might be added imposing additional structure on θ -roles. For example, one might require that if John is the patient of the (event corresponding to the) proposition Mary kicks John then he is also the patient of the proposition John is kicked (by Mary) (uttered in the same context). I.e.:

$$(58) Pa (\langle \text{'kick'}, j, m \rangle) = Pa (\langle \text{PAS'('kick')}, j \rangle)$$

In other words, we might build into the definition of θ -roles that they are 'preserved' under manipulations of arguments. However, we will not engage in giving an explicit axiomatization of such a property. It might also be possible to give axioms for specific θ -roles, but, again, we will refrain from doing so here. θ_1 and θ_2 provide us with a minimal theory which is sufficient for our purposes (namely the investigation of relations between control and θ -structure). We can, furthermore, define a 'degenerate' θ -role as follows:

$$(59) \bar{\theta}(\langle \text{P}, x_1, \dots, x_i, \dots, x_n \rangle) = x_i \leftrightarrow \\ \neg \exists \theta [\theta(\langle \text{P}, x_1, \dots, x_i, \dots, x_n \rangle) = x_i \ \& \\ \diamond \text{P}(x_1) \dots (x_n)]$$

Hence an individual x is said to bear the role $\bar{\theta}$ with respect to a given event iff it does not bear any θ -role with respect to that event (and such an event is possible). Notice that an individual bearing $\bar{\theta}$ with respect to an event cannot be an ordinary individual (i.e. a member of the base). This is a logical consequence of $\theta 2$.

So θ -roles are defined as partial functions on ordered n -tuples whose first member is a $(n-1)$ relation (i.e. what we call ur-complexes). These functions are further constrained by a set of axioms which include at least $\theta 1$ and $\theta 2$.⁶

We are now in condition to give a precise formulation to the claim made at the beginning of Section 3, namely that control phenomena are partly predictable in terms of θ -structure. Our control principle can be reformulated as follows:

$$(60) \text{CP (2nd version)} \\ \alpha(x_1) \dots (P) \dots (x_n) \leftrightarrow M_a \\ P(\mathcal{J}(\langle \alpha, x_1, \dots, x_n \rangle))$$

Like its predecessor, (60) is a meaning postulate schema that now has two parameters: α (which as before is simply lexically specified) and \mathcal{J} . \mathcal{J} is a (meta) variable ranging over θ -roles. Its value is taken to be filled in in terms of a lexical redundancy rule that employs the following independently established θ -hierarchy:⁷

(61) Th(eme) > Go(al) > So(urce) > Be(neficiary) > ..
 > $\bar{\theta}$

If a verb that obeys CP has a theme, the lexical redundancy rule will automatically fill in Th as the value of \mathcal{V} . Otherwise, the next θ -role on the θ -hierarchy will be filled in, and so on. Notice that we allow for $\bar{\theta}$ to be the most marked value of \mathcal{V} . Specific lexical items may be lexically specified for a value of \mathcal{V} in violation of (61). So a verb may have a theme and yet be, say, source-controlled in which case our lexical redundancy rule is prevented from applying. Instances of (60) are given in what follows:

- (62) a. $\Box \text{force}'(^{\wedge}P)(x)(y) \leftrightarrow$
 $M_aP(\text{Th}(\langle ^{\wedge}\text{force}', ^{\wedge}P, x, y \rangle))$
 b. $\Box \text{strike-as}(x)(^{\wedge}P)(y) \leftrightarrow$
 $M_aP(\text{Th}(\langle ^{\wedge}\text{strike-as}, x, P, y \rangle))$
 c. $\Box \text{tell}'(^{\wedge}P)(x)(y) \leftrightarrow M_aP(\text{Go}(\langle ^{\wedge}\text{tell}', ^{\wedge}P, x, y \rangle))$
 d. $\Box \text{promise}(x)(^{\wedge}P)(y) \leftrightarrow$
 $M_aP(\text{So}(\langle ^{\wedge}\text{promise}', x, ^{\wedge}P, y \rangle))$

Of the examples in (62), (62a-c) are regular in that the controller bears the θ -role predicted by our lexical redundancy rule in terms of the θ -hierarchy in (61).

Promise, however, is irregular since it has a goal-argument but is source-control; this has to be written into the lexical entry of promise, which makes it costly.

Something like (62a) says in plain English that y forces x to P iff in all the relevant deontic alternatives to the actual world, the theme of that forcing act does P.

(60), thus, allows us to relate in precise way the selection of the controller to the θ -structure of a predicate. Once it has been established for a given predicate whether it obeys CP, the values of M_a (i.e. the modal frame associated with the predicate) and of are automatically predictable, modulo some marked cases.⁸ The idea that the thematic hierarchy in (61) provides a markedness scale for suitability to be a controller is made fully precise within our model theoretic framework.

4. Control and Locality

There is a particular notion of locality built into CP, namely that of 'semantic internal argument'. It is so because CP is defined in terms of θ -roles, which in turn are defined only on the internal arguments of a predicate (cf. 1). We may now ask whether such a notion of locality adequately captures the local character of the control relation. I believe that it doesn't. Consider the following hypothetical predicates:

- (63) a. $R(\underset{\uparrow}{P})(x)(y)$
 b. $R'(x)(y)(\underset{\uparrow}{P})$

Given (60) we might expect x to be able to control P (as indicated by the arrow), since (aside for co-argumenthood), (60) does not impose any condition on how many arguments may intervene between the controller and the

controlled item. However, control predicates such as R or R' do not seem to be attested as far as I know. A predicate that *prima facie* looks like R would be ask as in the following example:

- (64) a. John asked Mary to be allowed to wash
himself
b. asked' (^P)(Mary')(John')

In (64a) a subject seems able to control across an object, thus yielding a structure like the one in (63a). However, Bach (1979, 1980a) has argued that Mary in (64a) and related constructions cannot be an object because it doesn't passivize:

- (65) a. *Mary was asked to be allowed to wash
himself (by John)
b. asked'(Mary')(^P)(John')

Hence the function argument structure associated with ask in (64a) cannot be (64b) but has to be (65b). In other words, ask is ambiguous between two categories: TV/IV and (IV/IV)/NP; only when it belongs to the latter category is ask subject controlled; hence the badness of (65a).

So there do not seem to be predicates with control structures such as the ones represented in (63). But this would come as a surprise on the basis of CP as we have formulated it, which shows its inadequacy. The local character of the control relation is not correctly captured by (60). Further constraints seem to be necessary.

What we will do in what follows is to consider some of the proposals that have been made in the literature

concerning locality requirements on control. We will try to point out various problems they seem to run into and propose an alternative.

4.1. Syntactic approaches

Williams (1980) argues that obligatory control is subject to a restriction based on C-command: the controller must C-command the controlled item. This restriction is part of a more general condition on a syntactic rule of predication, of which obligatory control is a particular case, according to Williams. Williams is aware of the following potential counterexamples to his theory, but explains them away by appealing to a reanalysis process that incorporates prepositions into verbs:

(66) John thinks of Bill as a friend

Williams argues that appeal to reanalysis is independently justified by the well known preposition stranding phenomena. However, I don't think that Williams' position can be maintained in general. Italian seems to provide quite clear counterexamples to the C-command restriction.

Consider the following examples:

- (67) a. ho dato a Mario dell'incompetente
 (I) told to Mario incompetent
 (I accused Mario of being incompetent)
- b. ho avuto in Mario un padre
 (I) had in Mario a father

- riesce
- c. correre piace a Maria
 to-run pleases/succeeds to Mary
 (Mary likes/succeeds in running)

In these constructions adjectives (67a), predicate nominals (67b) and infinitives (67c) enter a control relation (i.e. have an understood subject) which has all the characteristic attributed by Williams to the obligatory control relation. For example, a controller has to be present, as illustrated in what follows:

- (68) a. *ho dato dell'incompetente (cf. 67a)
 b. *correre riesce/piace (cf. 67c)

A lexical subject for the predicative complements of the constructions in (67) is impossible; the antecedent is uniquely determined, and so on. However the C-command condition clearly fails, and appeal to preposition stranding phenomena to justify reanalysis is impossible. Notice that Romance languages do have phenomena for which reanalysis processes have been argued to be necessary (see e.g. Rizzi, 1982). But the lack of preposition stranding suggests quite strongly that there can be no process which incorporates prepositions into verbs. That this is so has been argued convincingly in Kayne (1981). We thus are led to conclude that (67) are genuine counterexamples to Williams' theory and that the C-command condition on (obligatory) control cannot be right in general.

Bresnan (1982) has proposed a different characterization of the local nature of obligatory control (or rather,

in her terms, functional control). For theory internal reasons, she claims to be forced to assume that only subjects, objects and second objects can act as functional controllers. Once more, however, Italian seems to provide crucial counterexamples to Bresnan's claim, basically for the same reasons put forth for Williams' theory. The constructions in (67) have all the characteristics Bresnan attributes to functional control. But the controller is clearly neither a subject nor an object (Italian doesn't have second objects). For example, trying to reanalyze the PP in (67a) as an object appears to be hopeless. Objects undergo passive, while the PP in (67a) and related constructions doesn't:

(69) *Mario é dato dell'incompetente
Mario is told incompetent

Furthermore, if the PP in (67a) is cliticized, it shows oblique case marking, unlike objects:

(70) gli
*li ho dato dell-incompetente
(I) to-him-told incompetent
* him-told

Equally hopeless appears the attempt to analyze the PP in (67c) as some sort of 'non-nominative' postposed subject. Italian, like English has subject-verb agreement, and in constructions such as (67c) the verb clearly does not agree with the PP but with the infinitival complement (or with a lexical NP which is also possible with these constructions):

- (71) a. correre piace a noi
 correre pleases to us
- b. gli hamburger non ci piacciono
 the hamburgers do not to-us-please (3rd pl)

Thus, Italian seems to provide genuine counterexamples to Bresnan's claim that only subjects and objects can be functional controllers. Hence the local nature of obligatory control is not correctly characterised in Bresnan's theory.

It would seem to be the case, then, that no purely syntactic account of the notion of locality relevant for control phenomena is forthcoming. We now turn, therefore, to semantic accounts that have been proposed in the literature.

4.2. Semantic accounts.

Bach (1979) and Bach and Partee (1980) have proposed to characterize locality in control phenomena in terms of some sort of semantic minimum distance principle. More recent semantically based approaches to control (e.g. Klein and Sag, 1982) basically incorporate such a proposal without substantial modifications.

One way of stating Bach and Partee's proposal is the following. Consider the function argument structure in (72).

$$(72) R(x_1) \dots (x_{i-1})(P_i)(x_{i+1}) \dots (x_n)$$

According to Bach and Partee, only the argument immediately to the right of P (i.e. the i th+1 argument) can control P . In other terms, once a structure asking for control has combined with a function, the next argument that comes in will be its controller. There is an obvious sense in which the controller can therefore be characterized as the argument closest to the controllee at semantic structure. Another equivalent way to put it is to say that the controller must minimally f -command the controllee, where f -command can be defined as follows:

(73) an item α f -commands β iff the first function that "contains" α also "contains" β .⁹

So in a structure like (72) all the arguments to the right of P_i f -command it. Hence it is natural to state Bach and Partee's proposal as a 'minimal f -command' requirement.

Such a proposal appears to account for the lack of control configurations such as those exemplified in (63) above, and furthermore does not run into the kind of trouble that syntactic approaches run into because the relevant notion of distance is stated in semantic terms. In this regard, Bach and Partee's proposal appears to be an improvement over the theories considered in the preceding section. However, I think that there are other regards in which the f -command condition is inadequate. Crucial counterexamples to it appear to be constructions such as the following:

(74) Mary recommends reading War and Peace to John.
 It is quite plausible to assume that the PP to John in (74) is an internal argument of the matrix verb recommend. From a semantic point of view, one might notice that it seems impossible to recommend something without recommending it to somebody. Furthermore, from a syntactic point of view recommend seems to select the proposition to in much the same way in which 'dative' arguments are selected:

(75) *Mary recommend of/for/on John reading War and Peace

These facts can easily be accounted for on the assumption that the to-PP in (74) is an internal argument (i.e. 'subcategorized for'). Having established this, there are only two options as to what the function argument structure associated with recommend can be, namely:

- (76) a. $\text{recommend}'(\hat{P})(x)(y)$
 i.e. $\text{recommend} \in (\text{IV/PP})/\text{NP}$
 b. $\text{recommend}'(x)(\hat{P})(y)$
 i.e. $\text{recommend} \in (\text{IV/NP})/\text{PP}$

Is there any evidence that allows us to choose between the two alternatives in (76)? The answer is: yes. Passive provides the relevant evidence, which clearly favours (76b):

- a. *John was recommended to reading War and Peace
 b. Reading War and Peace was recommended to John

So recommend provides us with cases of control structures where the controller is thematically uniquely determined

but the f-command condition fails. In (70b) the controller does not f-command the controllee. Other similar cases would be represented by tough-adjectives:

- (77) a. to write papers is tough for John
 b. tough'(j)(write papers')

If, as it seems plausible to maintain, the function-argument structure of tough-predicates is as shown in (78b), then again we have examples of control relations that do not obey to the f-command restriction.

These are all cases of what we call semi-obligatory control. In other words all the counterexamples to the f-command restriction considered so far involve an argument position which appears to be optional. However, this observation appears to be irrelevant for our purposes because if the f-command restriction is meant to explain the lack of control configurations such as those in (63), then one has to assume that it applies to both obligatory and semi-obligatory control. Moreover, I think that there are also cases of strict obligatory control that provide counterexamples to f-command. One such case is represented by Italian constructions such as those in (67c). If, as argued in the preceding section, infinitival phrases are subject in riuscire type constructions, then the function-argument structure associated with them must be:

- (79) riuscire'(x)([^]P)

The oblique internal argument in (79) is obligatory, obligatorily controls P and does not f-command it. We are therefore led to conclude that the f-command restriction as currently stated cannot be quite right.¹⁰

The problem with the f-command restriction seems to be the following. Let us consider again the function argument structure in (63). F-command says that only the argument immediately to the right of P can control it. What the above consideration seems to show is that we have to allow (at least) also for the argument to the left of P to be the controller. So it is not in terms of f-command that the relevant restriction on control can be stated, but rather in terms of f-adjacency. Here is then the relevant definition:

- (80) a. an argument α is f-adjacent to an argument β iff either α immediately f-commands β or it is immediately f-commanded by β .
- b. α can control β iff it is f-adjacent to it.

(80) provides us with the definition of f-adjacency and of the domain condition that characterizes the local domain of control relations. The relevant domain is crucially characterized in terms of f-adjacency. The present definition is sufficient to rule out control configurations such as those in (63) and doesn't run either into the problem purely syntactic accounts run into or into those that an account in terms of pure f-command seems to

face. We thus submit the hypothesis that the minimum distance principle relevant for control is based on the semantic notion of f-adjacency.

It is quite straightforward to implement at this point the f-adjacency restriction into our former version of CP (i.e. 60 above). It can be done as follows:

$$(81) \text{ CP (3rd version)} \\ \alpha(x_1) \dots (P_i) \dots (x_n) \leftrightarrow M_a P(x_j) \ \& \\ \mathcal{J}(\langle \alpha, x_1, \dots, x_n \rangle) = x_j \text{ where either } j = i + 1 \text{ or } j = i - 1$$

This version of CP is just like the former one except that it is now required that the controller (still identified via a Θ -role) be f-adjacent to the controlled item. We have simply built into our principle of control the domain condition in (80b). This allows us to have all the desirable consequences of our former version of CP (e.g. the predictability of controllers in terms of Θ -structure) while accounting correctly for the local nature of the control relation, granted that 'co-argumenthood' turns out to be insufficient.

5. Inheritance.

Before we turn to a consideration of the empirical consequences of our proposed account of control phenomena, we have to discuss somewhat how the control properties of basic lexical items are preserved by its transforms. We know that verbs can undergo a number of argument manipu-

lating operations such as passive, detransitivization, causativization, argument drop and the like. We call all operations of this sort 'argument manipulations'. Argument manipulations are semantically operations that map n -place properties into m -place ones, where $m \neq n$. The problem is how to account for the fact that (82a-b), for example, have the same control properties as (82c-d).

- (82) a. John promised Mary to go
 b. Mary persuaded John to go
 c. John promised to go
 d. John was persuaded to go

Our control principle is a condition on the meaning of lexical items, i.e. promise' and persuade' in (82a-b).

However the logical form associated with (82c-d) will be roughly the one shown in (83a-b) respectively (disregarding tense).

- (83) a. DTRANS'(promise')(\wedge go')(j)
 b. PASS'(persuade')(\wedge go')(j)

Now for any relation R and any argument manipulation g , it will be the case that $g(R) = R$. Since CP is meant to be a condition on meanings and since e.g. PAS'(persuade') = persuade' how do we know that PAS'(persuade') inherits the control properties of persuade' (i.e. that it is theme-controlled)?

This issue seems to be related to general properties of possible argument manipulations in natural languages. It seems that, in general, linguistically admissible argument manipulations have somehow to preserve the

entailments associated with the relations in their domain. Consider, for instance, existential entailments. If John kills Mary, then there is somebody that kills Mary. PAS' applied to kill' deletes the most external argument of this relation (i.e. the one that corresponds to the subject). Yet the new predicate maintains the existential entailments associated with kill: if Mary is killed, then there is someone or something that kills her. As a further example, consider so called selectional restriction phenomena: if John ate a sandwich, then there is something concrete that he ate. This entailment is preserved under detransitivization: if John ate, then there must have been something concrete that he ate.. So it would seem that argument manipulations in natural languages cannot alter the meaning of an item too much: they must be entailment preserving. Such a constraint, if true, would be rather striking, for there are infinite logically conceivable operations that do not obey to it.¹² Furthermore, since control, on the present approach, is regarded as a particular entailment licensed by lexical items, it does not come as a surprise that also such an entailment should be preserved under argument manipulations, given the general nature of these operations in natural languages.

The point is now the following: how do we represent formally, within our theory of control, inheritance of

control properties under argument manipulations? A priori, two approaches come to mind. The first one would be simply to build into our control principle a closure condition for argument manipulations. It is quite straightforward to do that. Let \underline{h} be a variable ranging over argument manipulations. This means that \underline{h} takes its value from a finite set of the form $\{ \text{PAS}', \text{DTRANS}', \text{ADROP}', \text{CAUSE}', \dots \}$. To simplify the formulation of CP we allow \underline{h} to take also the identity map as a possible value. Our control principle can then be reformulated as follows:

$$(84) \text{ CP (final version)} \\ h(\alpha)(x_1) \dots (P_i) \dots (x_n) \leftrightarrow M_a P(x_j) \ \& \\ \mathcal{J}(\langle \wedge h(\alpha), x_1, \dots, x_n \rangle) = x_j \text{ where } j = i + 1 \text{ or } \\ j = i - 1$$

(84) is just like our preceding version of CP, except that we have substituted $h(\alpha)$ for α . This simply means that if a basic item obeys (84), then also its passive, causative, detransitive, etc. transforms must obey it. So for instance, if we assume that the control properties of promise' are characterized by (84), then also DTRANS'(promise') will have to obey (84). In particular, for any P and x, the following must be the case:

$$(85) \text{ DTRANS' (promise') } (\wedge P)(x) \leftrightarrow \\ M_a P(\text{So} \langle \wedge \text{DTRANS' (promise')}, \wedge P, x \rangle)$$

What (85) says in plain English is that the source-argument of DTRANS'(promise') must be the controller of P, whenever something is promised. This would be a straight-

forward way of accommodating the inheritance of control properties by transforms of basic items.

An alternative might be represented by assuming that only basic forms obey CP (i.e. that the correct form of CP is (81)). On this view, something like DTRANS'(promise') would not obey CP directly. However, by definition of DTRANS' we have that for any P and x, DTRANS'(promise')(^P)(x) entails $\exists y[\text{promise}'(y)(^P)(x)]$. The latter formula, via CP, would entail that in all the relevant situations where promises are fulfilled, x (i.e. the source-argument) does P. In this way, our intuitions about control in non basic cases would be captured by means of a chain of entailments, in an indirect way. Notice that since the ϵ entailment relation is transitive, we would seem to obtain just the same results as we did before using (84).

The next question to ask is then: is there any substantive difference between these two ways to go? We will see that the answer is indeed positive: there is a crucial difference. But in order to see actually what that difference is we will proceed in an indirect way. Our theory faces us with a choice. The choice essentially is between (84) and (81) as principles of control, where the former is closed under argument manipulations, while the latter is not. What if some class of predicates 'chose' one and another the other alternative? Would our

theory predict their properties to be different? In other words, we are suggesting that the presence or absence of the variable h over argument manipulations, which is what differentiates (84) from (81), might be a parameter of CP. A class of predicates might be characterized by the adoption of a control principle closed with respect to argument manipulations (i.e. (84)), while another class of predicates might be characterized by a version of CP not closed with respect to argument manipulations (i.e. (81)). At first sight this might strike one as a weird idea, also because at this point it is not clear what difference it makes. Nevertheless, let us assume that such an idea is right and let us see what happens.

6. Capturing Obligatory Control.

Let us consider first a class of control predicates whose properties are characterized by (84). So take a predicate α and assume that it obeys (84). Let θ be the θ -role that identifies the controller. Assume, furthermore, that g is an argument manipulating operation that deletes the i^{th} argument of predicates and assume that the i^{th} argument of α is the θ of α . By hypothesis, $g(\alpha)$ must obey (84); but g applied to α would delete the θ of α i.e. the designated controller. Hence $g(\alpha)$ would have no controller. Hence $g(\alpha)$ could not obey (84)

contrary to our assumptions: a contradiction. Hence the argument manipulating operation g must be undefined for α .

What the general moral appears to be is that if an item satisfies (84), it follows that argument manipulating operations are not allowed to remove the controller: a controller has to be present in the argument structure of predicates that obeys (84). Doesn't this ring a bell? Indeed. The obligatory presence of a controller is the crucial property of what we call obligatory control predicates (sec. 1.1 above). We are thus led to conjecture that CP (84) is what characterizes the class of obligatory control predicates. In order to show that this is actually the case we have to show that all the relevant properties of these predicates follow if we assume that they obey CP in the form given in (84). So let us try to do so.

In section 1.1. we have identified six different properties that characterize obligatory control predicates. That some of them follow from (84) is, by now, straightforward and hardly needs further comment. We have just argued that obligatory presence of a controller is the characterizing property of items that obey (84). Locality also follows from (84), because f-adjacency is built into it. The controller cannot be separated from the controllee by an intervening argument at semantic structure. We have already argued that this hypothesis

seems to provide the correct characterization of the relevant notion of locality. Thematic uniqueness follows from (84) in virtue of the fact that the controller is identified by means of Θ -roles, via a lexical redundancy rule formulated in terms of a thematic hierarchy. Thus, the controller must bear exactly one Θ -role with respect to the matrix predicate and its transforms. The lack of split antecedents is a corollary of this fact. Since controllers are identified by means of a Θ -role, and since Θ -roles are functions (see the discussion in sec. 3.2.), it follows that there can't be two distinct arguments bearing the same Θ -role. Hence 'split control' phenomena are ruled out. The lack of arbitrary readings for the controlled items is a consequence of the obligatory presence of a controller. Since (84) guarantees that there is always something to control the relevant verbal argument, control can never be arbitrary with these constructions. Also Bach's and Visser's generalizations follow from the obligatory presence of a controller. To see this, let us consider a couple of examples, which will allow us also to illustrate in more details what is meant by 'obligatory presence'.

Let us consider Visser's generalization first. Take a subject control verb such as promise. The relevant instance of (84) for promise, irrelevant details aside, is the following:

$$(85) \text{h}(\text{promise}')(\text{x})(\text{^P})(\text{y}) \leftrightarrow \\ \text{M}_a \text{P}(\text{So}(\langle \text{^h}(\text{promise}'), \text{x}, \text{^P}, \text{y} \rangle))$$

Let us now assume, contrary to facts that promise passivizes, i.e. that PAS' is defined for promise'. Then for some x and some P, the following must be the case:

$$(86) \text{PAS}'(\text{promise}')(\text{x})(\text{^P})$$

Recall that, by definition, (84) is a schema for a finite set of meaning postulates of the form:

- (i) $\text{promise}'(\text{x})(\text{^P})(\text{y}) \leftrightarrow \text{M}_a \text{P}(\text{So}(\langle \text{promise}', \text{x}, \text{^P}, \text{y} \rangle))$
- (ii) $\text{PAS}'(\text{promise}')(\text{x})(\text{^P}) \leftrightarrow \text{M}_a \text{P}(\text{So}(\langle \text{PAS}'(\text{promise}'), \text{x}, \text{^P} \rangle))$
- (iii) $\text{DTRANS}'(\text{promise}')(\text{^P})(\text{y}) \leftrightarrow \text{M}_a(\text{So}(\langle \text{DTRANS}'(\text{promise}'), \text{^P}, \text{y} \rangle))$
-

Hence, (86) qualifies as an instance of the left hand side of (85). Hence (86) is logically equivalent to:

$$(87) \text{M}_a \text{P}(\text{So}(\langle \text{PAS}'(\text{promise}'), \text{x}, \text{^P} \rangle))$$

But (87) will be undefined, for neither x nor P is the source-argument of PAS'(promise). PAS' drops the arguments of verbs that correspond to the subject position, which in the case of promise happens to be the source argument, i.e. the controller. Now if (86), i.e. the left hand side of (an instance of) (85) were true and (87), i.e. the right hand side of (85), were undefined, the entire biconditional, i.e. (85) itself would have to be undefined (or false).¹³ But this contradicts our assumption that promise does obey (84) (i.e. that (86) holds). Hence by assuming (86), i.e. that PAS' is defined

for promise', we derive a contradiction, namely that promise' does and does not obey (84). Hence PAS' must be undefined for promise': promise cannot passivize. The reason why this is so is that there always has to be a controller among the arguments of promise and its transforms. But passive would remove that controller from the arguments of promise, and hence is disallowed. Notice that the presence of a by-PP wouldn't make any difference for our argument, because by-PPs are external arguments. We have just argued that it follows from the fact that promise obeys (84) that PAS'(promise') must be undefined. Evidently, a functor that takes an undefined function as argument will yield an undefined value. External arguments are VP modifiers (i.e. functors). Hence, any by-PP applied to PAS'(promise') will also be undefined. Thus the presence of a by-phrase doesn't make any difference in the case at hand. It is a consequence of (84) that external arguments are 'too far' to act as controllers.

Consider now the following:

- (88) a. (?) John promised Mary to be allowed to wash herself
 b. Mary was promised to be allowed to wash herself

(88), in the dialects that allow for it, would be goal-control. Violations of CP (i.e. of (84)) arise if one attempts to remove an obligatory controller. Passive removes subjects, i.e. in the case of promise, the

source-argument. Since (88) is goal-control and passive removes the source-argument we would expect (88a) to be passivizable, for no violation of CP could arise in such a case. This is in fact the case. In other words, promise is associated with two Θ -structures, as pointed out in Bresnan (1982). On the one hand, to promise something to someone means to commit oneself to do a certain action for someone's benefit. This is the Θ -structure usually associated with the infinitival-taking promise (e.g.: John promises Mary to go). To promise something to someone can also mean to commit oneself to transfer an abstract benefit to somebody. This is the Θ -structure usually associated with \bar{S} -taking promise (e.g.: I promised her that she would be invited). What happens in (88) is that this second Θ -structure is superimposed in the infinitival-taking promise which is thereby turned into a goal-control predicate.¹⁴ Whence the apparent violation of Visser's generalization. Thus the present approach not only can offer a principled account for such a generalization, but can also provide an account for the apparent exceptions to it. The arguments we have given are, of course, fully general and apply to all subject control predicates.

Let us now turn to a consideration of Bach's generalization, and again let us illustrate it by means of an example. Consider an object control verb such as force.

Let us assume, contrary to fact, that it can be detransitivized, i.e. that DTRANS' is defined for force'. Then for some P and some x, the following must be the case:

$$(89) \text{DTRANS}'(\text{force}')(\wedge P)(x)$$

By hypothesis, force' obeys (84) and, more specifically, the following instance of it:

$$(90) h(\text{force}')(\wedge P)(x)(y) \leftrightarrow M_a P(\text{Th}(\langle \wedge h(\text{force}'), P, x, y \rangle))$$

(89) is an instance of the left hand side of (90). Hence by modus ponens, we get:

$$(91) M_a P(\text{Th}(\langle \wedge \text{DTRANS}'(\text{force}'), \wedge P, x \rangle))$$

But (91) must be undefined, for DTRANS'(force') has no theme-argument. The theme happens to be the argument that DTRANS' deletes. If (91) is undefined and (89) is true, (90) must be undefined as well, contrary to our assumptions. Hence a contradiction. Hence DTRANS' must be undefined for force'. That objects can't be removed from object control verbs that obey (84) can be derived by logic alone. Again, the argument just given is perfectly general and applies to all object control verbs that are obligatory control.

This illustrates that on the basis of (84) we are able to provide a unified account of both Visser's and Bach's generalization. They are simply particular instances of the fact that (84) sanctions the impossibility

of removing a controller from the argument structure of a predicate that satisfies it.¹⁵

Thus we are able to derive from (84) all the properties that obligatory control predicates seem to have. This would seem to show that (84) does indeed characterize adequately this class of predicates. Obligatory control predicates are those that obey (84), i.e. those whose meaning licenses the entailment that (84) expresses. We will henceforth refer to (84) with the label OCP (for 'obligatory control principle').

At the end of section 5, we put forth the idea that the presence or absence of closure with respect to argument manipulations (i.e. the presence or absence of the variable h in CP) might be a parameter that characterizes properties of different classes of predicates. We have argued so far that closure under argument manipulations does characterize the properties of obligatory control predicates. Let us now assume that there are predicates that instead obey (81) (i.e. the version of CP which is not closed under argument manipulations) and let us see what their properties are expected to be.

7. Capturing Semi-Obligatory Control.

Assume that α is a predicate that obeys (81) above and that the bearer of θ is the designated controller. If the bearer of θ is present among the arguments of

it will have to be the controller: this is what (81) sanctions. Suppose, however, that there is an argument manipulation g that happens to delete the bearer of θ . In this case, nothing happens, since (81) is not closed under argument manipulations, and hence $g(\alpha)$ does not have to obey it. Thus, we would expect that if an item obeys (81), the designated controller can be removed from the argument structure of α . But this seems to be just the crucial characteristic of semi-obligatory control predicates.

Let us consider this in some more detail. Take semi-obligatory control predicates such as signal, tough or recommend. The relevant instances of (81) will be, irrelevant details aside, the following:

- (92) a. $\text{signal}'(^{\wedge}P)(x)(y) \leftrightarrow M_aP(\text{Go}(\langle^{\wedge}\text{signal}',^{\wedge}P, x, y\rangle))$
 b. $\text{recommend}'(x)(^{\wedge}P)(y) \leftrightarrow M_aP(\text{Go}(\langle^{\wedge}\text{recommend}', x, ^{\wedge}P, y\rangle))$
 c. $\text{tough}'(x)(^{\wedge}P) \leftrightarrow M_aP(\text{Ex}(\langle^{\wedge}\text{tough}', x, ^{\wedge}P\rangle))$

Now it is quite easy to see that the relevant properties of this class of predicates follows on the basis of (81). First, the controller is identified as the bearer of a θ -role just as in the case of obligatory control. Hence thematic uniqueness is immediately accounted for. Second, (81) is subject to the same locality conditions as OCP, namely f-adjacency. So locality also follows. These two conditions together account for the badness of:

(93) *John told me that making a fool of himself was tough for Mary.

(92c) requires that the experiencer-argument of tough (Mary in (93)) and nothing else control the relevant verbal argument, and this is why (93) is out. Similarly, if the designated controller is present it will have to be unique, for it will have to be the (necessarily unique) bearer of a specified Θ -role. This accounts for the lack of 'split control' cases illustrated again below:

(94) *John told me that making fools of themselves was tough for Mary

By a parallel reasoning, it also follows that arbitrary readings for the controlled argument are impossible, if the designated controller is present in the argument structure of this class of predicates. Thus the fact that in these regards predicates such as tough, signal or recommend behave like obligatory control predicates receives a natural account on the basis of the hypothesis that they obey (81).

Let us consider now the cases where the designated controller is deleted by some argument manipulation.

Examples are given in what follows:

- (94) a. Tom signalled to go
 a'. DTRANS'(signal)(^go')(t)
 b. Tom recommends reading War and Peace
 b'. ADROP'(recommend')(^read War and Peace')(t)
 c. to write papers is tough
 c'. ADROP'(tough')(^to write papers')

The primed versions of (94) represent the logical forms associated with the corresponding non primed versions. The predicates under consideration, by hypothesis, obey (81) which is not closed under argument manipulations. So, in particular the relevant instances of (81), i.e. those in (92), do not apply to the cases in (94), which are, therefore, allowed. Notice, furthermore, that by definition of ADROP' and DTRANS', (94a'-c') will entail (95a-c).

- (95) a. $\exists y[\text{signal}(\text{'go'})(y)(t)]$
 b. $\exists y[\text{recommend}'(y)(\text{'reading War and Peace'})(t)]$
 c. $\exists y[\text{tough}'(y)(\text{'write papers'})]$

The relevant versions of CP (i.e. (92)) will of course apply to the formulae in (95), guaranteeing that whatever value of y makes them true will be the controller of the verbal arguments. So our intuitions about the control properties of these predicates are still nicely captured in such an indirect way. Furthermore, this accounts for the fact that when the designated controller is absent from the argument structure of these verbs, then they seem to pattern like prominence control predicates. For example an arbitrary, or generic reading becomes possible. The reason why this is so is that in (94) we attribute certain properties to conditions (i.e. nominalized properties) in general, not to particular events. The statements in (94) entail those in (95) and the latter

are about some event (state, etc.) or other. Only in this indirect way, can the statements in (94) be said to be about events. It is precisely this indirectness that makes arbitrary readings possible.¹⁶ By parallel reasoning it is possible to explain why 'long distance' or 'split antecedent' control phenomena seem to become possible, when and only when the designated controller is absent from the function-argument structure of the predicates under consideration. Consider once more the following examples:

- (96) a. John told Mary that moving in together
would be tough
b. tell'(^ADROP'(tough')(move in to-
gether))(m)(j)

John and Mary can be understood as the implicit argument of ADROP'(tough'), i.e. as the experiencers of the relevant state of toughness. To the extent that this is possible they will automatically be guaranteed to be the designated controllers of the action which is said to be tough in (96), namely moving in together.

Finally, it is also straightforward to explain why argument control predicates appear to violate Visser's and Bach's generalizations:

- (97) a. it was decided by the committee to leave
b. Tom signalled to leave

Visser's and Bach's generalizations are a consequence of the impossibility of removing the designated controller in a control structure, as argued in sec. 6. But we have

just seen that verbs obeying (81) allow for the designated controller to be deleted. Hence, a fortiori, they cannot pattern according to Bach's and Visser's generalizations. Nothing in (81) requires that they should. The logical forms associated with (97) will be:

- (98) a. $\text{by}'(\text{the committee}')(\text{PAS}'(\text{decided}))(\text{'leave'})$
 b. $\text{DTRANS}'(\text{signal})(\text{'leave'})(t)$

Since (81) is not closed under argument manipulations, $\text{PAS}'(\text{decided})$ and $\text{DTRANS}'(\text{signal})$ do not have to satisfy it. However, (92a-b) entail respectively:

- (99) a. $\text{decide}'(\text{'leave})(\text{the committee}')$
 b. $\exists y[\text{signal}'(\text{'leave})(y)(t)]$

(81) will then guarantee automatically the right control relations for the examples in (99).

Thus all the properties of argument control predicates seem to receive a principled account on the basis of the hypothesis that they obey (81). All such properties can, in fact, be deduced from (81) by logic alone. We thus conclude that (81) characterizes correctly the control properties of argument control verbs. Henceforth, we will refer to (81) as SOCP (for 'semi-obligatory control principle'). We will use CP to refer ambiguously to SOCP and OCP.

The situation is then the following. We have a unique parametrized principle of control. Such a principle says essentially that the controller is going to be a bearer of a designated θ -role, subject to a condition of

f-adjacency. The principle is parametrized with respect to the property of being closed under argument manipulations (i.e. having or not having the variable h). Predicates may choose to go either way. If they select 'closure under argument manipulations', (i.e. OCP), then they will necessarily have all the observed properties of obligatory control predicates. If they select 'non closure under argument manipulations', then they will necessarily have all the observed properties of semi-obligatory control predicates. So CP seems to provide us with a maximally simple, truth conditionally fully explicit account of control phenomena. No appeal to phrase-structure geometry, to grammatical relations or to non existent subjects appears to be required. On the contrary such an appeal would seem to be harmful (cf. sec. 4.1 and sec. 10).

8. Capturing Prominence Control.

It is fairly straightforward to imagine, at this point, what our proposal about prominence control predicates (i.e. predicates such as bother, be dangerous, denounce, etc.) is going to be. We want to suggest that prominence control verbs are simply characterized by the fact that they do not obey CP at all. This commits us to saying that the logical forms associated with (100) are going to be (101).

- (100) a. to write papers disturbs Sue
 b. to write papers is dangerous
 c. The president discussed killing
 civilians for no reason (from Horn,
 1975)
- (101) a. disturb'(^write papers')(s)
 b. dangerous (^write papers)
 c. discuss'(^kill civilians')(the presi-
 dent')

The semantic representation of sentences such as those in (100) is simply going to involve reference to actions. Something like (100a) says just that Sue is disturbed by an action. Something like (100b) attributes to an action the property of being dangerous. The fact that the semantic representation of sentences such as those in (100) involves nominalized properties and is not subject to CP of course does not mean that prominence control verbs cannot have a controller. It merely means that they do not have to have one. If they don't, such sentences receive the generic or arbitrary reading.

To make the preceding point clearer consider the following example:

(102) We discussed friendship

Let us think about circumstances that could make (102) true. Such circumstances could be of two kinds. On the one hand we might have discussed friendship in abstracto (e.g. how great it is, how difficult it is to find it, etc.). On the other hand we might have talked about specific instances or cases of friendship (e.g. we might

have talked about a particular friend of ours). Talking about particular instances of friendship might well suffice to make (102) true. So a given relation involving abstract individuals might characterize a situation that actually involves instances or realizations of that abstract individual. The same reasoning evidently applies to something like (100c). Discuss in (100c) is associated semantically with a relation between an individual (the president) and an action (killing civilians for no reason). The truth of (100c) would be warranted either if the president talked about the action as such, or if he talked about particular realizations of that action. What could realizations of actions be? Consider for instance the action of writing papers. A realization of such an action is, for example, my writing this paper now. So events can be regarded as realizations of actions. Events are individuals-in-relations; hence in some intuitive sense events (et similia) 'have subjects'. Something like (100c) can thus involve either an action in general or some realization of it, i.e. an event (complex). In the latter case, there will be an understood subject for the nominalized property.

Various factors may influence whether sentences such as (100) are going to be about actions in general or about realizations of actions. Some such factors might be purely pragmatic. Others might be more directly connected

prominence control verbs and to individuate one of the crucial factors responsible for inducing something like a controlled reading of the verbal arguments in their complement structure. The distinction between 'things' or 'individuals' and their realizations (or 'stages' in Carlson's terminology) represents a substantive hypothesis on the nature of some aspectual phenomena. Such a distinction has been developed by Carlson especially in connection with the analysis of bare plurals and has been subsequently elaborated by a number of authors.¹⁷ To illustrate one of the applications of such a distinction, consider the following example:

- (104) a. dogs are widespread
 b. dogs are barking in the courtyard

(104a) is a generic statement and is therefore about some individual in general. In the present case, the subject to which the property of being widespread is attributed is a bare plural NP. According to Carlson, bare plurals denote kinds. So (104a) is about the kind of dogs in general. (104b) is an episodic statement. It is a report on something that is happening in the courtyard. Hence it will be a statement about some particular realization of what the subject of the statement denotes. If the subject of (104b) denotes a kind, then (104b) will be about some realization of that kind, i.e. about instances of the dog-kind. This explains why the bare plural NP dogs'

receives some sort of 'universal' or 'generic' reading in (104a) and some sort of 'existential' reading in (104b). As pairs such as those in (104) show, the same aspectual mechanism that gives rise to the contrast observed in (103) also seems to play a crucial role in determining when terms referring to conditions (such as infinitives and gerunds) in prominence control structures are understood as if they had a subject. We will come back to some of these issues in ch. V. However, to develop the present suggestion in full details would force us to consider in detail problems of tense and aspect, which is impossible within the limits of the present work.

The above discussion illustrates that if we represent the meaning of prominence control verbs like those in (100) as relations between actions and individuals, it does not mean that control is impossible simply because these predicates do not obey CP. It means that when control phenomena do arise, they have to be related to other 'modules' of the theory of grammar. One such module is a theory of tense and aspect. But there are many subtle factors, some of them of a pragmatical nature, that come into play.

The hypothesis that prominence control predicates do not obey CP accounts for the fact that they do not have to have a controlled reading, since they do not license the entailments that characterize control verbs. This pro-

vides us with a basis for explaining why prominence control verbs allow for non local control phenomena and for controllers that are simply salient entities in the context. These were identified as being the characteristic properties of prominence control verbs.

9. Equi vs. Raising.

The present approach to control is meant to apply indifferently to equi and raising verbs. In the present section, I would like to show how some of the differences between these two classes of verbs fall out as an automatic consequence from CP. What we will say is not meant by any means to be an exhaustive account of raising phenomena, which represent a marked and complex set of facts. Rather, we will focus on just one set of properties of raising constructions, namely what might be called the existential entailments associated with these constructions. In order to see what precisely is meant by the latter locution we need to spell out a little bit more certain general properties of raising constructions.

Raising predicates are generally described as predicates that have virtual or non logical arguments anaphorically related to some verbal argument that they control.

Examples are the following:

- (105) a. John seems to be home
 b. John happens to be home
 c. Mary believes John to be home

The diagnostics which are usually employed to determine the 'non logical' nature of arguments of raising verbs are quite rich. Typically it includes expletive constructions (such as it in it rains and there in there is a spot on the floor) and idiom chunks (such as keep tabs on). This is illustrated by contrasts such as the following.

- (106) a. There seems to be a man in the garden
 b. *There tries to be a man in the garden
- (107) a. The cat seems to be out of the bag
 b. *The cat tries to be out of the bag
 (only literal meaning)

Let us call the capacity of having as arguments expletive pronouns or parts of idioms 'logical virtuality'. Raising constructions always display this property.

A second important characteristic of raising is referential opacity. The typical test for referentially opaque positions is represented by existential generalization. So consider the following contrast:

- (108) a. A unicorn seems to be approaching ↗
 there exists a unicorn
 b. A unicorn tries to approach us →
 there exists a unicorn

While the existential entailment in (108b) is valid, the one in (108a) is not, which shows that the subject position of seem is referentially opaque.

So two characteristic properties of raising, which distinguish it from Equi, are that the controller position can host 'virtual arguments' and is referentially opaque. Within semantically based theories of control the way this

phenomenon is usually captured is by positing the following semantic structures:

- (109) a. raising: $\lambda P \lambda \mathcal{G} [\text{seem}'(P(\mathcal{G}))]$
 (= $\lambda P [\text{seem}' \circ P]$)
 b. Equi: $\lambda P \lambda \mathcal{G} \lambda x [\text{try}'(P(x))(x)]$

The approach illustrated in (109) can be generalized to all Equi and raising structures. Essentially, Equi is treated via 'quantifying in' while raising is via 'function composition.' There are various possible implementations of this idea (see e.g. Dowty, 1978, Klein and Sag, 1982) but they all have something like (109) in common. Now, a priori one might conceive at least two logical possibilities aside from those in (109), namely:

- (110) a. $\lambda P \lambda \mathcal{G} \lambda x [M(P(x))] = R$
 b. $\lambda P \lambda \mathcal{G} [M(P(\mathcal{G}))(\mathcal{G})] = R'$

In the hypothetical verb R , both quantifying in and function composition are realized. In the hypothetical verb R' an NP-meaning would control a VP-meaning without being quantified in. These hypothetical control structures are no more and no less far fetched or ad hoc than those in (109). A question to ask, then, is: are the structures in (110) attested? And if they are not, why is this so?

A related question about the difference between Equi and raising has to do with the way in which logical virtuality and referential opacity interact in control structure. Does any control structure that has one of the

two properties always have the other? I think not. There seem to be control structures where the controller position is referentially opaque but does not allow for virtual arguments. An example would be the following:

- (111) a. I need a friend to talk to
 b. *I need tabs to keep on celebrities

(111a) lacks the existential entailment that characterizes referentially transparent positions. Even if (111) might be a case of argument control (the infinitival being a purpose clause)¹⁸ it still serves to make the point. A control structure might be referentially opaque but disallow virtual arguments. Something like (101a) would be a case where the controller is an 'intensional theme'.¹⁹

The opposite case is, I think, unattested. To my knowledge there are no control structures which allow for virtual arguments to occur in the controller argument-slot but are referentially transparent. In other words there are no control verbs that have the properties illustrated in (112).

- (112) a. V'(VP')(it')
 b. V'(VP')(aCN') → there exists a CN

Verbs that *prima facie* might look as counterexamples to this generalization are happen or continue. These verbs have the property of logical virtuality (113a) and might seem to have also the property of referential transparency (113b).²⁰

- (113) a. it continues to rain
 b. a man continues to knock at my door →
 there exists a man

However, referential transparency for these verbs appears to be parasitic upon their 'factive' or 'semifactive' nature in the following sense. The entailment in (114) appears to be valid:

- (144) it continues to be the case that $p \rightarrow p$

Let us assume that p in (114) is equal to $P(c)$ (i.e. proposition p is gotten by applying the propositional function P to c). Whether $P(c)$ entails $\exists xP(x)$ will depend on whether P is extensional or not. If it is (as in the case of (114b) to knock is), then a continues to P will entail $\exists xP(x)$ since: a continues to p → it continues to be the case that $P(a) \rightarrow P(a) \rightarrow \exists xP(x)$. This chain of entailments will give the feeling that happen, continue and other aspectual verbs are referentially transparent. On the present view, however, we would expect that if P is intensional in, say, a continues to P, then the existential entailment characteristic of referentially transparent positions should disappear, since the last step in the chain of entailments above would be invalid. This indeed seems to be true, as illustrated by the following examples:

- (115) a. A blond tall man continues to be
 sought by the police → there exists a
 blond tall man
 b. A tire continues to be missing from
 the garage → there exists a tire.

So, I think that the generalization illustrated in (112) is a genuine one. We can formulate it as follows:

- (116) **virtuality-opacity correspondence:**
 If a control structure allows virtual arguments to control, then the controller position is referentially opaque.

The next question is then whether it is possible to provide a principled account for the virtuality-opacity correspondence. Before trying to do so, however, we need to say something concerning the treatment of virtual arguments within the present framework.

9.1. On virtual arguments.

There are various proposals concerning treatments of virtual arguments within semantically based theories of grammar. Usually the common line tends to be based of what Karttunen calls the 'ugly object' approach. Virtual arguments (expletive pronouns, NPs in idioms) are assigned as a semantic value some specified abstract entity. Things are then set up in such a way that any normal property applied to such an entity (the ugly object) will be undefined; only the relevant constructions (e.g. rain or to be a man in the park) can apply to the relevant entity (i.e. it and there respectively). The ruling out of constructions such as there eats or there rains can be carried on either purely in the semantics (i.e. by means

of the various ugly objects assigned to virtual NPs) or by some combination of semantic and syntactic techniques.²¹

To illustrate a little bit further, let us consider a possible treatment of expletive it in English. Verbs taking expletive it as subject might be analyzed simply as constant 1-place propositional functions. So rain and seem that Mary is here would be constant functions that map any argument whatsoever into the proposition that it rains and that it seems that Mary is here, respectively. If there are propositional functions, why couldn't natural languages make use of propositional functions which are constant with respect to some of their arguments? On this view something like John rains wouldn't make any sense because we know that rain is insensitive to its arguments, and so what would be the point in applying it to John or to any other object? What we could do, then, is simply to let expletive it be the normal pronoun it, so that the translation of it rains would be simply:

(107) $\text{rain}'(x)$

What is then the value of x going to be? Well, it doesn't really matter, since rain' doesn't care. It certainly will not be any ordinary individual, otherwise (117) would end up meaning something analogous to John rains and would therefore be deviant for the same reasons the latter is. So we can assign to x any arbitrary object we like which lies outside of our referential frame (i.e. $B \cup Q$, cf. ch.

II sec. 3.2). It can play this role because in some intuitive sense it is the NP that contains less positive information about its referent; she wouldn't do because it carries information concerning the gender of its bearer; this wouldn't do because its bearer must either be ostensibly identifiable or somehow contextually prominent, and it is sensible to assume that only entities within our referential frame can be contextually prominent or have feminine gender. So if rain is a constant function, the unmarked pronoun it is really the best choice to play the role of its subject, to the extent that it is interpreted by some arbitrary object outside of our referential frame. The appealing aspect of this approach to expletive it is that it uses very little machinery and allows for a unified treatment of the expletive and the pronominal it. So for instance there is no need to worry about occurrences of expletive it in unwanted contexts. Also straightforward would be to handle cases of expletive it-control such as the one illustrated in what follows:

- (118) a. it seems to rain
 b. seem'(rain')(x)
 c. □_{seem}rain'(x)

The logical form of (118a) would be (118b). seem, being $\bar{\theta}$ -control would license the entailment in 118 which says that in all the situations compatible with what seems, it rains. x in (118c), and hence also in (118b) would have

to be some object that lies outside of our referential frame by the reasoning given above.

This is of course very sketchy but gives, I think, a fair idea concerning the basic line we might take on the issue of virtual arguments. With this in mind let us now turn to a more detailed consideration of the Equi vs. raising distinction.

9.2. The virtuality-opacity correspondence.

A first element to consider is why the controlling NP in Equi constructions always acts as if it were 'quantified in'. This is what most semantically based theories of control capture by assigning to Equi verbs the structure in (109b) above. Our principle of control is a condition on the meaning of lexical items. It follows, therefore, that CP must apply to members of B_A (for some category A) and not to members of P_A , which are derived by means of some syntactic rule. Each grammatical entry is associated with a sortal specification, which defines its sort (ch. II sec. 3.2.). One of the tasks of the sortal specification is to determine whether any given position in a verb is extensional or not. Extensionality is a lexical property of items that has to be determined quite independently of control facts. If, however, a certain control verb is extensional with respect to the position of the controller, the controller will have to be

an ordinary individual (i.e. a member of B, not a member of Q), since CP applies to lexical items. Let us illustrate this with an example. Take a verb like persuade, its category will be TV/IV; its type $\langle e, \langle e, \langle e, p \rangle \rangle \rangle$. Its sortal specification will be: $\langle \langle e, p \rangle, \langle e_B, \langle e_B, p \rangle \rangle \rangle$. This sortal specification determines that the first argument of persuade is a condition and that such a verb is extensional with respect to its object and subject positions. The relevant instance of CP is:

$$(119) \quad h(\text{persuade}')(\wedge P)(x^B)(y^B) \leftrightarrow \\ M_a P(\text{Th}(\langle \wedge h(\text{persuade}') \rangle, \wedge P, x, y))$$

The fact that (119) is restricted to elements of the base (i.e. x^B, y^B , etc.: ordinary individuals) follows from the fact that persuade is extensional and CP a condition on lexical items. A consequence of this is that (120a) entails (something like) (120b) and does not entail (120c).

- (120)
- a. John persuades every boy to go
 - b. John persuades every boy that he should go
 - c. John persuades every boy that every boy should go

The reason why this is so is that the logical form associated with (120a) will be (121a) which is equivalent to (121b) and it is only to the latter that CP applies.

- (121)
- a. $[\text{persuade}'(\text{ go})]_Q(\text{every}'(\text{boy}'))(j)$
 - b. $\forall x[\text{ boy}'(x) \rightarrow \text{persuade}'(\wedge \text{go}')(x)(j)]$

So, on the present approach the 'quantifying in' property typical of Equi doesn't have to be lexically stipulated but is reduced to a lexical stipulation which is independently needed, namely the extensional or intensional nature of predicates.²² A control verb with extensional controllers will always behave like persuade in example (120).

Our theory allows for cases of $\bar{\theta}$ -control as the most marked option. What happens if an item is lexically specified as being $\bar{\theta}$ -control? Let us consider, for example, seem. The relevant instance of CP will be:

$$(122) \quad h(\text{seem}')(\wedge P)(x) \leftrightarrow M_a P(\bar{\theta}(\langle \wedge h(\text{seem}'), \wedge P, x \rangle))$$

If (122) is the case, by the contrapositive of $\theta 2$ of our theory of θ -relations (in (57) above), it immediately follows that the controller cannot be a member of the base, i.e. an ordinary individual. This is precisely as it should be, because a $\bar{\theta}$ -control item, by definition of $\bar{\theta}$, will allow virtual arguments to be controllers and, as we have seen in 9.1., virtual arguments are going to be represented as objects which lie outside our referential frame and hence, a fortiori, outside of \underline{B} . The following is thus a consequence of $\theta 2$ and (122) (i.e. the relevant instance of CP).

$$(123) \quad \neg \exists x^B [\text{seem}'(\wedge P)(x^B)]$$

But (123) is precisely the entailment that characterizes referential opacity. If an item is $\bar{\Theta}$ -control (i.e. allows for virtual arguments to be controllers) it follows that it cannot be referentially transparent. Hence the virtuality-opacity correspondence receives a principled account on the basis of our approach to control.

In other words being $\bar{\Theta}$ -control disallows ordinary individuals from being controllers. What can be, then, a controller in a $\bar{\Theta}$ -control structure? Either some virtual object outside of our referential frame or some member of Q. Two examples are given in what follows.

- (124)
- a. it seems to rain
 - b. $\text{seem}'(\wedge \text{rain}')(x) \leftrightarrow \square_{\text{seem}} \text{rain}'(x) \leftrightarrow \square_{\text{seem}} \text{rain}'$
 - c. a man seems to be approaching
 - d. $\text{seem}([\text{to be approaching}]_Q)(\text{a man}') \leftrightarrow \square_{\text{seem}} [\text{be approaching}]_Q(\text{a man}') \leftrightarrow \square_{\text{seem}} \exists x[\text{man}'(x) \ \& \ \text{be approaching}'(x)]$

Notice that for (124c) to be true, the verbal argument of seem (i.e. be approaching) must have been raised by the sort lifting operator $[\]_Q$, for only properties lifted via $[\]_Q$ can apply to members of Q, such as a man', as required by the relevant instance of CP. This is totally unproblematic, given that we know on independent grounds that $[\]_Q$ applies freely to all verbs (ch. II sec. 3.2.).

Hence, from the way control is defined it follows automatically that there can't be a class of control

predicates that has logical virtuality but lacks referential opacity. Notice that from a purely logical point of view, things could have been otherwise. It is quite easy to conceive a class of verbs that allows for virtual arguments to be controllers but is referentially transparent with respect to the position of the controller. For instance, within the present framework on the basis of the treatment of virtual arguments sketched in 9.1., such a class of verbs might be accommodated by assigning to it the structure shown in (110b). Verbs of such a class would allow expletive it (on the basis of the theory in 9.1) and would be referentially transparent in the controller position. Our principle of control interacts with various other aspects of the present theory (i.e. the theory of θ -relations, the theory of intensionality) to rule these cases out.

In other semantically based approaches to control the differences between Equi and raising constructions have been correctly factored out in terms of a set of stipulations concerning the lexical structure of those verbs, namely sets of stipulations that employ the schemata reported in (109) above. We have seen, however, that such an approach leaves unanswered several questions, such as the exact nature of what we call the virtuality-opacity correspondence. The present approach enables us to move a few steps towards an explanation of this phenomenon, by

allowing us to derive it in terms of the very same mechanism that determines the general properties of the control relation (i.e. CP).

10. Some Comparisons.

In building our theory of control we have drawn from various sources. Our starting point was the approach to control phenomena developed by a number of authors within the tradition of extended MG on the basis of the research program first explicitly worked out (as far as I know) in Bach (1979). I have tried to incorporate within such program some of the insights contained within other frameworks like the one of Williams (1980) and, especially, Bresnan (1982). Occasionally, some comparisons have been drawn between these theories and ours. At this point a more systematic consideration of similarities and differences between the theory developed here and other current approaches to control might be useful. Given the broad spectrum of basic assumptions that differentiates current theories in formal linguistics, the comparisons that we will be able to draw will necessarily turn out to be rather sketchy. We will try to consider both syntactic and semantic theories of control and point out some of the empirical problems they seem to run into.

10.1. Syntactic approaches.

The characteristic of syntactic approaches is represented most prominently by the appeal to either phrase structure geometry (GB) or grammatical relations regarded as a primitive of syntax (LFG). All syntactic theories at some point or other do appeal, as we will see, to semantics (or pragmatics) to account for facts that seem to lack any principled purely syntactic account. However, what typically happens is that this appeal to semantics is, in general, not worked out and is elusive. The present approach, on the other hand, constitutes an attempt to provide a fully explicit semantic account of control phenomena. As soon as some such account is developed, not only do the 'left overs' of syntactic theories seem to fall into place, but furthermore the need for an independent purely syntactic theory of control seems to vanish. This is the general structure of the main conceptual objection that we will raise against all purely syntactic theories of control.

Let us begin by considering Bresnan (1982). Bresnan distinguishes two types of control: functional control (that corresponds roughly to what we call obligatory control) and anaphoric control (that covers both our argument and prominence control). Functional control requires 'identity of all grammatical features' while anaphoric control requires simply 'identity of refer-

ence'. Functionally controlled items (e.g. VPs) are associated with 'open functions', i.e. functional structures with something like a 'gap.' A set of control equations (in the case of lexical induced control) guarantees that these 'gaps' in functional structures are properly filled. Anaphorically controlled structures are functional structures that contain a phonetically unrealized pronominal element ('PRO'). Various factors constrain both the occurrence and the interpretation of PRO. Two such factors are an obviation principle (the referent of PRO must be distinct from the subject of the minimal functional structure containing the anaphoric control structure) and an f-command restriction (the antecedent of PRO must f-command it).

The appealing aspects of Bresnan theory are that she provides an integrated account of Equi and raising and that, furthermore, a uniform explanation of Visser's and Bach's generalization falls out from the requirement, built into Bresnan's formalism, that functionally controlled arguments have to have a controller. We have seen that Bresnan's idea can be reproduced quite straightforwardly within a model theoretic framework as a condition on models (i.e. our OCP), without having to go through (as far as control is concerned) an autonomous level of functional structure labelled by grammatical relations.

For theory internal reasons, Bresnan argues that only subjects and objects can be functional controllers, which is what provides the locality restriction on functional control in her theory. We have already tried to argue in Section 4.1. on the basis of Italian data that there are cases of control by an oblique NP which have all the properties Bresnan attributes to functional control. This seems to be a problem for Bresnan's approach.²³

Aside from the restrictions mentioned above, anaphoric control in Bresnan's theory is pretty free. For example, anaphoric control allows for split antecedents. Without further qualifications, however, this leaves unexplained the ungrammaticality of sentences such as (21) above, reported below:

- (125) a. *Mary signalled Tom to entertain each other
 b. *Mary said that to enjoy themselves was tough for John

Both signal and be tough would be anaphoric control structures on Bresnan's theory. A related difficulty is to explain why, if anaphoric control is free, things like the following are impossible:

- (126) a. *Mary signalled Tom to kiss herself
 b. *Mary recommended to Tom not making a fool of herself

In order to account for these phenomena, Bresnan has to appeal to θ -restrictions (to be stated in the semantics or the pragmatics) on anaphoric control. But once such

' θ -restrictions' have been precisely stated (and adequately parametrized), as on our approach, they can take care of facts like the ones in (125) and (126) (which are left unsolved on Bresnan's theory) as well as of all the rest of the control phenomenology (including what Bresnan calls functional control).

Let us now turn to a consideration of Williams (1980). Williams' theory of control is based essentially on three rules. First there is a rule of predication which coindexes a c-commanding NP with a predicate. \bar{S} s with a PRO subject are regarded as predicates. This rule is obligatory and takes care of what we call obligatory control, but also has, according to Williams, a much broader range of application. When predication takes place, control structures (i.e. PRO headed \bar{S} s on Williams' theory) that do not undergo it are marked arb. Arb- \bar{S} s cover both our argument and prominence control cases. Williams proposes, then, two further rules of arb-rewriting that provide the necessary structure to interpret control phenomena characteristic of arb- \bar{S} s. They are reproduced in what follows:

- (127)
- a. Arb-rewriting I. Coindex arb and NP if they command each other.
 - b. Arb-rewriting II. Coindex arb with an NP that commands it.

These rules are disjunctively ordered (I precedes II) and I is obligatory while II is optional.

In sec. 4.1. we have already pointed out some empirical problems with the c-command restriction on predication seems to face. Those cases are serious difficulties for Williams's theory because the role of c-command is absolutely central to it. For instance, it is c-command that Williams relies on to provide an account of Visser's generalization:

(128) *Bill was promised by Mary to go

Promise in (128) licenses obligatory control. However Mary does not c-command PRO to go and hence predication can't apply. If the c-command restriction does not work, in general for all obligatory control structures, then this explanation of Visser's generalization cannot be maintained.

A further empirical problem for Williams' theory is constituted by the following prominence control structures:

(129) Mary told me that to make a fool of herself
bothered John

In the embedded clause John and to make a fool of herself command each other, hence they should be coindexed by ArbI, which is obligatory. Hence (129) is predicted to be ungrammatical.

A related difficulty arises with gerunds. Williams is not explicit as to whether gerunds are supposed to be treated on a par with infinitives. If they aren't, then the fact that some gerund-taking verbs (such as like or

accuse of) display all the relevant properties of obligatory control would remain mysterious. If they are, then only subject control is incorrectly predicted for cases like the following:

(130) John recommended reading War and Peace.
ArbI should coindex John and the gerund in (130), thus predicting the wrong control relation.

A further difficulty has to do with the fact that according to Williams the Arb-rewriting rules should allow or somehow be compatible with split antecedents. Exactly how this compatibility is to be granted is not spelled out. However, if both Arb-rewriting rules are compatible with split control phenomena, then the ungrammaticality of (125) and related examples would seem to lack any principled explanation. A way out would be to stipulate that only ArbII is compatible with split control phenomena. But aside from the stipulative nature of this move (which would leave us not overwhelmed with insight), one of its consequences would be to duplicate into ArbI virtually all the properties of predication (obligatoriness, thematic uniqueness, lack of split antecedents) except for c-command. This duplication would seem to constitute a significant loss of generality. Nothing in the logic of the system explains why two different rules are allowed to diverge just with respect to the c-command restriction.

Consider next Koster (1981). Koster distinguishes two fundamental grammatical relations: binding and control.

The binding relation has the following properties:

- (131)
- a. it obligatorily connects two items and
 - b. there is only one binder
 - c. the binder must c-command the bound element
 - d. the binding relation is local, where locality is defined in terms of the notion of 'minimal governing category'

The control relation is defined in terms of the notion of function (or predicate) argument structure. In particular the control relation has the following properties:

- (132)
- a. it is obligatory, although the controller may be left implicit
 - b. it is local: the designated controller and the controlled element are contained in the same minimal argument structure (NP,S).
 - c. the controller must be 'superior' (sic) at argument structure
 - d. the controller does not have to be unique

Koster distinguishes then three cases. Some verbs relate to an empty subject in a subcategorized \bar{S} only in terms of the binding relation. These are raising structures. Some verbs relate to empty subjects in a subcategorized \bar{S} via both binding and control; these correspond to our obligatory control structures. Some verbs relate to the empty

subject of an \bar{S} in their argument structure only via control; these cover both implicit and prominence control structures. In terms of this classification, Koster argues that all the relevant properties of these constructions can be accounted for. Koster also argues that verbs where the controller-controllee relation is also characterized in terms of binding are all and only the verbs that trigger \bar{S} -erasure. This accounts for the fact that obligatory control verb (like try) do not take a lexical subject or the complementizer for: the presence of a complementizer blocks \bar{S} -erasure.

A first empirical difficulty for the latter claim (pointed out in Manzini 1982) has to do with the following constructions:

- (133) a. *Mary decided (for) Bill to leave
 b. it was decided to have dinner at 6 (by
 John)

decide in English does not take a complementizer for as (133a) illustrates. Hence, according to Koster it should trigger \bar{S} -erasure and be both a binding and a control structure. But it is not, as shown by the grammaticality of (133b). If it were, control by an implicit or non c-commanding NP would be ruled out by (131).

Further difficulties for Koster's claim that \bar{S} -erasure verbs (i.e. binding plus control) correlate systematically with the lack of a complementizer comes

from Italian.²⁴ The following example serves to make the point:

- (134) a. amo, (*di) dover pagare tasse
 (I) love have to pay taxes
 b. dover pagare tasse é amato da molti
 have to pay taxes is loved by many

As shown in (134a) amare does not take a complementizer. However, contrary to what Koster would predict it allows for an oblique to control the infinitival complement in the passive.

Besides these kinds of problems, which are specific to Koster's approach, most of the empirical difficulties that are problematic for Williams are also problematic for Koster. This applies, in particular, to the c-command condition on obligatory control, that in Koster's theory comes from the fact that obligatory control verbs are also subject to the binding conditions (cf. 131c). Furthermore, given that verbs not subject to binding are all argument control for Koster, cases such as (129) should be ungrammatical, because only arguments in the minimal argument structure of a verb are allowed to control (cf. 132b). But in (129) the controller is not part of the minimal argument structure, and verbs of the bother class (unlike verbs like be interesting) lack implicit arguments that could play the role of local controllers.

Let us now turn to a brief consideration of Manzini (1982). Manzini's theory is based on the following claim:

- (135) a. A PRO subject of a subcategorized sentence in a sentence S is bound in S
- b. A PRO subject of a subject sentence (co)refers freely

So with verbs like bother and be tough the PRO subject of an \bar{S} in subject position is allowed to have its antecedent either in an higher \bar{S} or from some contextually prominent entity. \bar{S} s in a subcategorized position must get an antecedent for their PRO subject in the minimal S that contains them. Manzini tries then to collapse (135a-b) by setting up a notion of domain-governing category (based on the definition of 'accessible antecedent') and requiring PRO to be bound within its domain-governing category.

Like other configurational theories of control, Manzini's approach suffers from the by now familiar difficulties that derive from the attempt to define the notion of 'accessible antecedent' in terms of C-command. We will not further comment on this aspect. A problem which appears to be specific to Manzini's theory is represented by the ungrammaticality of sentences such as (125b). According to Manzini, an \bar{S} in subject position should always allow its PRO subject to corefer freely, for it never has an accessible antecedent in the minimal S containing it (cf. 135b). However, this clearly doesn't happen in sentences like (125b). These kinds of construc-

tions have subject \bar{S} s whose PRO subject are not allowed to corefer freely.

A symmetric difficulty arises with Italian examples such as the following:

(136) L'assessore ha proposto coltivare orti come
attività ricreativa

the assessor proposed take-care-of gardens
as a recreational activity

The PRO subject of the infinitival coltivare orti in (136) has an accessible antecedent (i.e. the matrix subject); hence according to Manzini it should be bound in its domain-governing category, and it should be unable to pick a contextually salient antecedent or to get a generic interpretation. This prediction, however, is not borne out. The PRO subject of the infinitival in question can easily refer to a contextually understood subject or have a generic interpretation.

A further point to notice about Manzini's theory is that it fails to provide any account whatsoever for Visser's and Bach's generalizations. Manzini is aware of this and hints at 'semantic or pragmatic' accounts. But then, again, once such 'semantic or pragmatic' accounts are made available they seem to be able of taking care also of the generalizations Manzini's theory is designed to capture.

Even if in a very sketchy form, we have considered a significant sample of purely syntactic approaches to

control and tried to show that they run into both conceptual and empirical difficulties.²⁵ Our semantically based approach seems able to provide a solution to such difficulties.

10.2. Semantic approaches.

We have already mentioned more than once that the theory developed here is essentially an outgrowth of various other semantically based approaches to control, and we have occasionally pointed at various ways in which the present theory departs from them. Here we would like to give a more general overview of similarities and differences between the present theory and the ones developed in Bach and Partee (1980) and Klein and Sag (1982).

What seems to be common to both Bach and Partee (1980) and Klein and Sag (1982) is the appeal to a semantic minimum distance principle based on f-command and to the notion of 'transitive verb phrase'. The theories differ in that TVP is an actual syntactic category in Bach and Partee (1980) while it is not so in Klein and Sag (1982); what is crucial on the latter theory is actually the semantic type of TVPs. Let us discuss first Bach and Partee (1980). The f-command restriction bans the existence of subject control TVPs. If subject control TVPs existed, then the controller would not (minimally)

f-command the controllee. Furthermore the syntactic category TVP defines the domain of the passive rule. It then follows that subject control verbs cannot undergo passive, because they are not TVPs but IV/IVs. A first difficulty for this approach is represented by subject control verbs that do passivize, such as (133b) above, repeated here for convenience:

(137) It was decided to have dinner at 6 (by John)

A similar problem arises in connection with subject control verbs in Italian, many of which do undergo passive:

(138) a. dover pagare le tasse é amato da molti
(=134b)

b. di essere eletto é stato garantito a
Carlo da Maria

"di" be elected (mas.) was guaranteed
to Carlo (mas.) by Mary

it was guaranteed to Carlo by Mary
that he would be elected

There are two ways I can think of to accommodate these cases within the framework of Bach and Partee (1980). One way to go would be to assume that discuss (and amare or garantire in Italian) are not IV/IVs but, actually, TVs, and this is why they passivize. This solution would entail that at least some infinitival VPs are to be regarded as NPs, as has occasionally been proposed in the literature. Another way to go would be to argue that what goes on in (137) and (138) is not the 'real' passive but some distinct process, say 'impersonal passive'. Both

solutions appear to have serious drawbacks, however. On the first solution, one has essentially to categorize a subject control verb as an IV/IV or as a TV according to whether it undergoes passive or not. While this correctly describes the facts, it would deprive us of any insight concerning what is at the base of Visser's generalization. On the second solution, one might want some independent criteria to differentiate passive from impersonal passive. This appears to be particularly problematic for two reasons. First, I can't think of any property that would differentiate in Italian cases such as (138) from regular passive cases. Second, Bresnan (1982) points out that an analogue of Visser's generalization arises even with cases of 'intransitive passivization' in Icelandic, which do look a lot alike what one might want to call 'impersonal passive'. So while Bach and Partee (1980) do provide the beginning of an account for Visser's generalization such an account seems to suffer from some drawbacks.

The account of Visser's generalization provided in Klein and Sag (1982) is very similar to that of Bach and Partee (1980). What Klein and Sag do is to provide an interpretive procedure for GPSG that makes use of various semantic operations one of which is passive. Such a procedure blocks whenever one tries to apply passive to an item which is of the wrong logical type, i.e. which is not of the type $\langle ty(NP), \langle ty(NP), ty(S) \rangle \rangle$. So it is crucial

for this approach that the logical type associated with VPs be different from the one associated with NPs.²⁶ If one considers a broader range of predicative expression this tenet is going to be very difficult to maintain, as argued at some length in ch.I. This might be not so devastating for Klein and Sag because one could make the interpretive procedure they develop sensitive to semantic sorts rather than to semantic types, and the semantic sort of (nominalized) VPs is different from the one of other individuals. However, the cases where subject control verbs undergo passive do seem to be as problematic for Klein and Sag as they are for Bach and Partee, and for the very same reasons. The syntactic category or the semantic type (sort) fo VPs is not sufficient to predict which subject control VP-taking verbs will be allowed to passive.

The account we have proposed for Visser's generalization is independent of whether passive can apply to IV-taking verbs (i.e. IV/IV) or not. Passive in our theory is an operation that deletes the last argument of a predicate (the one that corresponds to the subject of the active form). If, however, a verb is lexically specified as obeying OCP, it automatically follows that passive cannot apply to it. We have a control principle that is parametrized with respect to closure under argument manipulations; a verb positively specified for this parameter will not be able to undergo passive, as argued

in sec. 6, as a consequence of an independently established property of OCP-verbs, namely that they do not allow for a controller to be removed from their argument structure (which is one of the basic insights of Bresnan, 1982). So cases like (137) or (138) are unproblematic for the present theory; more than that they are predicted to exist, because there are verbs that are negatively specified for closure under argument manipulations (i.e. verbs that obey SOCP).

A further advantage of the present approach is that it allows for a unified account of both Visser's and Bach's generalizations. This is desirable in view of the fact that there is a clear intuitive analogy between these two generalizations: in both cases a grammatical process is prevented from applying just in case it would 'destroy' an obligatory control relation. No such unified account seems to be possible on the basis of the theories developed in Bach and Partee (1980) or Klein and Sag (1982).

Moreover, we have argued at some length in sec. 4.2. that the semantic minimum distance principle relevant for control theory cannot be properly stated in terms of f-command alone. It has to be generalized to a principle of f-adjacency. Such a principle, built into CP, allows us to treat many cases of control that could not be straightforwardly handled in terms of f-command restrictions.

Also to be pointed out is the fact that by stating CP in terms of Θ -roles (for which a model theoretic characterization has been provided) we can formulate precisely generalizations concerning the way the Θ -structure of a predicate may influence its control properties. Those generalizations cannot be stated on the basis of the approaches to control we are considering.

A final difference which is worth point out, is that on our theory, unlike what happens in Bach and Partee (1980) or Klein and Sag (1982) we do not need to posit a phantom subject for verbal arguments at semantic structure. This would be per se a desirable result, by Occam's razor. But we will see that it has also important empirical consequences, once we consider the way anaphora interacts with the behavior of verbal arguments (ch. V).

10.3. Concluding remarks.

The theory we have developed is based on principle CP. Verbs that take verbal arguments and are negatively specified for such a principle (i.e. do not obey it) are prominence control verbs. Verbs that are positively specified for it are the other control verbs. Such a principle is parametrized with respect to closure under argument manipulations. A positive or negative specification for such a parameter yields the properties of obligatory and argument control verbs respectively. Thus all

the properties that characterize the three classes of verbs individuated at the beginning of this chapter receive a principled and integrated account,

CP is a meaning postulate schema, which on the present theory are conditions on the structure of models. A minimum distance principle, based on the semantic notion of f-adjacency, is built into it. Such a principle is also subject to modal qualifications based on the theory of Kratzer 1981) and the controller is identified in terms of its Θ -role. All these notions (f-adjacency, modal qualifications, Θ -roles) are independently required and have a broader range of application in the grammar. No appeal to phrase structure geometry, to primitive grammatical relations or to null subjects appears to be necessary. This conceptual simplicity is quite appealing, or so I would think. But the present theory is not only appealing on purely conceptual grounds. There various control phenomena that are problematic for other current theories of control (both syntactically or semantically oriented) and that our theory seems to be able to begin to solve. Thus, solutions that are more appealing on meta-theoretical grounds also turn out to have some empirical advantages.

Notes

¹When examples are necessary, I will illustrate the relevant property by picking one sample for each of the set of constructions in (4)-(6).

²Recall that it is to a theory of context to specify the details of how this actually takes place. The mechanism informally described in the text (no matter how executed) will achieve the effect of 'indexing' the modal operator to a given control verb and to the relevant tuple of individuals.

³Our 'natural properties' correspond closely to what Cocchiarella calls 'predicative properties'. Cocchiarella (1983) shows how to state within HST* a logic for predicative properties.

⁴This and the following definitions can be carried on entirely with IL*.

⁵For any set A, $\mathcal{P}(A)$ is the power set of A.

⁶It might be possible that axioms such as 1 and 2 will turn out to be derivable from deeper principles of English metaphysics (e.g. principles concerning the internal organization of events).

⁷The Θ -hierarchy in (61) is said to be independently established because its range of application in the grammar is potentially wider than the one represented by (argument) control theory. See Jackendoff (1972, 1976) for discussion.

⁸For an approach that employs similar ideas to deal with what we call modifier control, see Nishigauchi (forthcoming).

⁹In this version, the notion of f-command makes crucial use of the syntax of logical form. However, if propositions (which we treat as primitives for convenience) are assumed to have enough algebraic structure, it might well be possible to define f-command without appealing to the syntax of IL*.

¹⁰The very same arguments can be raised against GPSG implementations of the f-command restriction such as the one in Klein and Sag (1982).

¹¹It might be possible to give a non disjunctive definition of f-adjacency along the following lines: α f-adjacent to β iff there is no intervening argument between α and β . This definition brings about more clearly in what sense the notion of f-adjacency might capture an intuitive notion of 'closeness'.

¹²An example of an operation that alters the adicity of a predicate and is not entailment preserving is the following:

(i) $\lambda P \lambda x \lambda y [P(x) \vee y = \text{Luciano Pavorotti}]$

Also Johan van Benthem has pointed out to me (p.c.) that PAS' and DTRANS' (together with reflexivization, i.e. $\lambda R \lambda x [R(x)(x)]$) satisfy what he calls quantity, restriction, and continuity (see van Benthem (forthcoming) for definitions and discussion) and that modulo some limiting cases, they are the only argument manipulations that satisfy these independently established semantic constraints.

¹³Recall our discussion of semantic deviance in ch. II sec. 3.3.

¹⁴See Bresnan (1982) for further discussion and arguments in favour of the present hypothesis.

¹⁵This strikes me as the fundamental insight of Bresnan (1982). The present theory can be regarded as a way of implementing Bresnan's insight directly with a model-theoretic setting, without the medium of grammatical relations and of Bresnan's 'syntactic' functional structure.

¹⁶The possibility of 'generic readings' are of course related to issues of tense and aspect. For more discussion on this topic see sec. 8 below and ch. V sec. 1.2.

¹⁷E.g. Cooper (1982), Dowty (1978). De Mey (1982), Enç (1981), Chierchia (1982, 1983).

¹⁸According to Bach (1982) purpose clauses are sometime arguments of verbs and sometimes TVP-modifiers.

¹⁹Given Montague's way of representing intensional objects (reconstructed here as members of Q, see ch. II sec.3.2.) examples of 'intensional theme' control would probably have to be analyzed along the lines of (110b). In this case, the answer to whether there are control

structures of this sort (i.e. the one illustrated in (110b)) would be affirmative.

²⁰Thanks to Anna Szabolczi for bringing this fact (back) to my attention.

²¹See e.g. Dowty (1982), Sag (1982), Klein and Sag (1982) for relevant discussion and detailed proposals.

²²Our approach does not exclude the possibility of having intensional controllers. See example (111a) in the text and fn. 19 above.

²³Similar phenomena seem to arise in Irish, as pointed out by McCloskey during a talk given at U Mass (fall 1982).

²⁴Koster uses data concerning the distribution of the complementizer om in Dutch to back up his claim.

²⁵We haven't considered the theories developed in Chomsky (1980, 1982). For criticisms of these approaches, we refer to Williams (1980), Koster (1981) and Manzini (1982).

²⁶Klein and Sag wish to maintain, however, that \bar{S} s can be NPs and hence that the type of \bar{S} s is a subtype of the type of NPs.

CHAPTER V

INFINITIVES, GERUNDS AND ANAPHORA

In considering how anaphora interacts with verbal argument constructions, it is useful to separate off two kinds of anaphoric phenomena. For purposes of exposition, we will label them external and internal anaphora.

By external anaphora we mean roughly anaphoric dependencies displayed by verbal arguments in discourse. In particular, we will consider the use of pronouns to refer back to a verbal argument, as in:

(1) Bill likes playing violin and John likes it too

Also certain valid inferences involving verbal arguments will be considered. We will try to argue that several important aspects of external anaphora lend support to the analysis of infinitives and gerunds put forth in the present work.

By internal anaphora, we refer to the behavior of reflexive and non reflexive pronouns inside verbal arguments. This will lead us to discuss various aspects of the theory of binding, especially with respect to non coindexing constraints (disjoint reference and non co-reference, see Lasnik 1976). The outcome of our analysis will differ from most current approaches in formal semantics in various ways. We will try to show that the

semantic notion of predication has a syntactic counterpart that cannot be analyzed simply as a special case of function-application. A more powerful device (say, some form of coindexation between a predicate and its "subject") will be needed. This is of course one of the central insights of Williams' theory of predication, and we will try to provide a way to incorporate it within the present categorial framework.

1. External Anaphora

In the present section, first the generic reading of verbal arguments (i.e. those analyzed within the tradition of transformational grammar as involving a PRO_{arb}) and its interaction with it-pronominalization will be considered. Then we will turn to an analysis of under what circumstances an anaphoric relationship between it and a verbal argument can give rise to "sloppy" readings. Finally, we will consider some valid inferential processes involving verbal arguments.

1.1. PRO_{arb} : its truth conditional import.

Typical generic occurrences of verbal arguments might be taken to be the following:

- (2) dancing is fun
 to dance

In Chapter IV we have adopted the view that predicates like be fun have an (oblique) argument, that can be deleted by a general (lexical) process of argument drop. In the case of be fun, an application of argument drop would yield say:

$$(3) \text{ ADROP (fun) = } \lambda x \bar{\lambda} y [\text{fun}(y)(x)]$$

Now, in the simple past a sentence like (2) seems to be ambiguous between a generic and an episodic reading:

(3) dancing the polka
to dance was fun

Within the present theory, the episodic reading of (3) (i.e. the one according to which we are characterizing a particular event in the past as being an occasion of fun) would be represented more or less as in (4).

- (4) a. $H[\text{ADROP}(\text{fun})(\text{'dance'})]$
b. $H[\bar{\lambda} y \text{ fun}'(y)(\text{'dance'})]$

By definition of ADROP, (4a) and (4b) turn out to be the same proposition within IL_* . So, since fun is a semi-obligatory control predicate, CP will apply to it ensuring that it was y's dancing (whoever y is) that was fun (for y). This explains why on the episodic reading, (3) is understood as being about a particular dancing event. This point could be made more forcefully by bringing in the analysis of aspect originally developed by Carlson (1977). According to Carlson, episodic sentences are about "stages", i.e. spatiotemporally located indivi-

duals. So (5) would be construed as being about a spatio-temporally located state of affairs. Developing this hint more fully would lead us to consider issues of tense and aspect that cannot properly be dealt with here.

On the assumption that there exist a generic tense operator¹, the generic reading of (3) would be represented as shown in (5).

(5) H[Gn(ADROP(fun'))(^dance')]

Since the simple present tense is usually understood as generic (whatever the reasons for this may be), (2) will be simply translated as (5) without the past tense operator, i.e.:

(6) Gn(ADROP(fun'))(^dance')

Notice that (6) is not directly about dancing events, but about the activity of dancing in and of itself. It is true that the basis for claiming something about an activity is provided generally speaking by considering possibly hypothetical occasions where such activity is actually carried on. This might be built into the logic of Gn by some suitable set of axioms. For instance, one might try to capture the dispositional nature of generic predicates by using standard possible world techniques, e.g. saying that whenever I have some property "generically", then there is a "sufficient" number of possibly counterfactual situations where I have that property non generically.² In this indirect way, (6) might be

regarded as being about particular dancing events. But this effect would be obtained without construing (2) as having as its direct subject matter a state of affairs or an event.

According to this view, then, there are abstract individuals (such as activities). They are part of a special kind of nominalized properties (that we call conditions). Infinitives and gerunds are a way to refer to conditions. Generic statements such as (2) are paradigmatic cases when we say something about "subject-less" activities and not about events, states of affairs, situations and the like.

This would suggest that PRO_{arb} (if there were such a thing) ought to be interpreted as some sort of λ -abstractor, so that we would end up interpreting (2) as something like $fun'(\lambda x[dance'(x)])$ rather than resorting to some sort of "contextually specified subject".

I will now to provide further support for this view by arguing that there is no way to assign a proper "propositional" meaning to generic occurrences of verbal arguments. The thing to ask is what would be a plausible propositional meaning for something like (2). Let us consider the various possibilities that come to mind. Evidently, a representation along the following lines would be inadequate:

$$(7) \exists x[\text{fun}'(\text{dance}'(x))]^3$$

(7) would be true just in case there is someone such that it is fun that he dances; but this is clearly not what (2) means. What about the universal quantifier?

$$(8) \forall x[\text{fun}'(\text{dance}'(x))]$$

If (8) were to represent (2), (2) would be true just in case everybody is such that his/her dancing is fun. This, however, appears to be quite clearly too strong. (2) might well be true in a situation where some, possibly even most dancing events constitute sad occasions. One way to weaken (8), might be by restricting the universal quantifier by using, say, a context variable as follows:

$$(9) \forall x[P(x) \rightarrow \text{fun}'(\text{dance}'(x))], \text{ where } P \text{ is understood as being specified by the context}$$

It should be noted that for an analysis like the one in (9) to be plausible it is crucial for the universal quantifier to have the scope that it does in (9). In other words, something has to prevent the following as a representation of the meaning of (2):

$$(10) \text{fun}'(\forall x[P(x) \rightarrow \text{dance}(x)])$$

According to (10) what constitutes fun is an occasion where everybody (in a certain context) dances. But again, this is simply not what (2) intuitively means. So (10) has to be disallowed. On a theory where the representation of (2) involves a PRO it is not clear, however, from what principle the impossibility of (10) should

follow, for the syntactic representation of (2) would be something like:

(11) $s[s[\text{PRO}_{arb} \text{ to dance}] \text{ is fun}]$

If S is a possible domain for scope assignment, then it is not clear why something like (10) should be ruled out on such a theory.

Be that as it may, (9) appears to be the closest one can get to assigning a truth conditionally plausible meaning to the generic verbal arguments in (2). Still, (9) appears to be unsatisfactory in many respects. (9) is about a contextually specified class of particular dancing events (or states of affairs where a certain group of people dances). But (2) doesn't have to be about particular events. This is especially evident for cases like the following:

- (11) a. writing dissertations is difficult
 b. $\forall x[\text{difficult}'(x\text{'s writing dissertations})]^4$
- (12) a. we talked about writing dissertations
 b. $\forall x[\text{we talked about}'(x\text{'s writing dissertations})]$

It seems intuitively clear that (11) and (12) might obtain in situations where no specific process of dissertation writing is difficult or is the topic of our conversations. Activities may be difficult or may be talked about just as abstract entities (e.g. friendship or goodness) are, i.e. without bringing into consideration any of their specific instances.

A possible problem for this view might be represented by sentences like the following⁵:

(13) smoking pot is dangerous for kids

(13) has two readings: one according to which what is dangerous is the kids' smoking, the other according to which what is dangerous for kids is that someone smokes in their presence. If we assume that the gerund in (13) contains a PRO_{arb}, one might attribute this difference to the fact that in one case PRO_{arb} is controlled by the matrix clause 'dative' argument, while in the other case it is interpreted as a variable bound by our universal, context sensitive quantifier. Within our approach there are various ways to go. The simplest one is presumably to assume that there are two predicates be dangerous: one is semi-obligatory control (i.e. obeys SOCP) the other prominence control (i.e. does not obey CP at all). So there would be dangerous₁ and dangerous₂ both in ADJ/PP, associated with different meanings. The two meanings associated with (13) would simply derive from the choice of one or the other dangerous'. It should be noted that the following sentence does not display an ambiguity parallel to the one of (13):

(14) smoking is fun for kids

Note also that the analysis just sketched predicts on the basis of our theory of control that dangerous should pattern with bother with respect to reflexives (on one

reading). Such a prediction seems to be born out as shown by the following contrast:

- (15) *John told me that getting drunk and making a fool of himself was fun for his wife
- (16) John told me that getting drunk and making a fool of himself bothered his wife
- (17) John told me that getting drunk and making a fool of himself was dangerous for his wife.

The dangerous that allows for (15) is the prominence control one. On the present approach, there would be nothing structural in the ambiguity that (13) displays. There several conceivable variants of this analysis that would yield analogous results.

So, postulating that PRO_{arb} is something like a variable bound by some generic quantifiers (a) does not do for us any work that could not be done on a theory where verbal arguments denote properties and (b) brings in a rather obscure notion of generic quantification, that in cases such as (11) and (12) runs quite clearly against our intuitions of what those sentences mean.

One might argue that there are languages that have overt generic subjects, such as the Italian si:

- (18) si corre troppo da queste parti
one runs too much around here

The interpretation of PRO_{arb} could be analogous to that of si in (18), whatever the latter may be. However, aside from the interpretation, si and PRO_{arb} have virtually nothing in common. According to the view just sketched,

the only thing they really have in common is the obscurity of their interpretation. This strikes me as a very weak defense of the role of PRO_{arb} unless it is coupled with a more substantive analysis of both construction. On the other hand, the analysis we have offered seems to do the job and is quite precise and general; whether it has anything to say about overt generic subjects is a separate issue that we will not try to address here.⁶

Further support for our analysis can be found in the way PRO_{arb} interacts with pronominalization. It turns out that generic verbal arguments can be antecedents of non-generic ones and vice versa⁷:

(19) Flying the Atlantic is perfectly safe today but it scares John.

(20) John loves chasing rabbits even if it is illegal

In (19) a gerund whose 'subject' would be PRO_{arb} seems to act as an antecedent for a gerund whose 'subject' would be PRO_{John}; in (20) it is the other way around. Let us assume, for sake of discussion, that pronouns are simply translated as variables that can either be bound by some antecedent or get their value from the context. Even on the basis of such a skeletal approach to pronouns, it is obvious that cases such as (20) and (19) are precisely what one would expect given our approach to verbal arguments. The it in both cases refers back to the underlined property-denoting expression (i.e. it refers to a pro-

perty). Then CP determines the right control relations automatically. So (19)-(20) on the present view are fully parallel to simple sentences such as:

(21) John is a nice guy but Mary hates him

On the other hand on a propositional analysis of verbal arguments the story about (19) and (20) cannot be so simple. The antecedent of it in both cases will have to be reconstructed in some way. It is true that there are examples showing that 'reconstructions' of some sort will be needed anyway for cases like the following:

(22) If one sleeps on a bed of nails, it hurts. But I do it anyway.

The antecedent of the underlined it seems to be something like sleeping on a bed of nails, which is not overtly present. It has to be reconstructed from the context. So on this view, (19) and (20) would be more similar to (22) than to (21). Our approach makes the opposite claim. So, at least in conception, the two theories are quite distinct and eventually it should be possible to find empirical evidence to choose among them.

What would help us to find relevant evidence would be cases of syntactic binding involving verbal arguments. If there were cases where one could claim that infinitives and gerunds are "quantified in" (or any analogon to that), they would be totally unproblematic for our theory, while they would seem to create problems for a propositional

approach (for the item to be 'quantified in' would have to have PRO_{arb} and Pro_i according to where it 'landed'). Of course, it is not easy to claim that examples such as (19) or (20) may involve syntactic binding, because verbal arguments, on our approach, are simply names of activities. Quantifying in proper names does not give rise to detectable scope ambiguities. However a construction that could help us to make the point could be represented by a clefting⁸; consider:

- (23) It is writing papers that Mary likes and John hates

The PRO in the clefted gerund in (23) has to be both PRO_{John} and PRO_{Mary} . Similar examples can also be constructed with generic PROs. For many speakers the following is grammatical:

- (24) It is hunting rabbits_i that ____i pleases me and ____i is not illegal.

The PRO in the underlined gerund has to be somehow arb and I. Nothing special needs to be said about these cases within our theory. Take your favorite theory of cleft constructions and couple it with the present approach to infinitives and gerunds. Presumably, the logical form of (24) would look roughly as follows:

- (25) $\lambda x[x \text{ pleases me and } x \text{ is not illegal}] (^{\text{hunting rabbits}})$

Then CP will automatically guarantee the right results as to the interpretation of the 'understood subjects' of the

gerund. On the other hand, things do not seem to be so smooth for a propositional theory and in general for a theory that has PROs. I am not sure of what one would have to do to get things right. Prima facie, it would seem necessary first to undo movement at LF (i.e. to "put back" the clause [PRO hunting rabbits]_S in the ____i slots in (24)) and second to make sure that control applies after this "reconstruction rule". Something like this seems to be necessary. However, both these moves appear totally stipulative. Nothing like that is needed on our approach.

It might perhaps be argued that this is not so devastating. For instance, on Williams' (1980) theory of control, both to please and to be illegal are non obligatory control verbs. This means that their clausal subjects would both have a PRO_{arb} and would be subject to his Arb-rewriting rule I. So, in particular the aforementioned strategy (some sort of "reconstruction" rule plus a stipulation, ordering PRO-interpretation after it) would merely affect this Arb-rewriting rule, and not the rule of predication that applies prior to that and is subject to stricter conditions, such as C-commands (see Williams, 1980 for details). However, there are languages, like Italian, where infinitives can be clefted, even infinitives occurring as complements of obligatory

control verbs such as the equivalent of English try, as the following example shows:

- (25) E' di uscire dall'URSS; che Dimitri sta
 it is to get out from USSR that Dimitri is
 cercando ___i e che il Comitato Centrale
 trying and that the Central Committee
 sta discutendo ___i
 is discussing

cercare ('to try') has all the properties of obligatory control structures: an antecedent has to be present, passive is impossible, no lexical subject can appear in its complement structure. This shows that it is impossible, in general, to circumscribe the effects of examples like (24) to free control cases, at least in the case of Italian. Now, examples like (25) appear to be especially problematic for Williams, because according to him clefting and obligatory control are one and the same thing: predication. But any theory that handles control via PRO would seem to run into trouble.

I should emphasize once more that these constructions posit no special problem for our theory; on the contrary, infinitives are non clausal and denote properties; control is an entailment licensed by certain verbs. This will automatically guarantee the correct interpretation of the understood subjects of the verbal arguments in the constructions under discussion.

To sum up, I think that a consideration of the truth conditional import of the so called PRO_{arb} constructions

and of their interaction with various kind of anaphoric processes lends considerable support to the view that verbal arguments are syntactically VPs and semantically associated with properties. PROs seem systematically to get in the way of an otherwise quite natural and general account of the phenomena we have considered.

1.2. Sloppy and non sloppy anaphora.

Predicates can apply only to properties, only to propositions or to both. This is a simple reflex of the fact that there are verbs that can take only verbal arguments or only sentential arguments or both. How does this simple fact interact with it-pronominalization? According to the very rough view of it-pronominalization we have sketched above, a verb having a pronoun it in one of its argument positions would be analyzed semantically as a function of the form $V'(x)$ where x is either bound by something or gets a value from the context. Now, x will be able to range over properties, over propositions or over both according to what sort of arguments V' can take. This makes an interesting prediction, on the basis of our approach to verbal arguments, concerning the distribution of "sloppy" readings. If a verb takes only properties as arguments (i.e. it subcategorizes for verbal arguments but not for \bar{S} s) and is not prominence control, then only a "sloppy" reading of the pronoun will be

possible; the understood subject of the property to which it refers will have to be different from the understood subject of the antecedent. So the relevant structure would be of the form $V_1'(P) \dots V_2'(\underline{it}')$, where V_2 does not take sentential complements and hence is semantically a function that applies solely to properties. This is a totally straightforward consequence of our theory of syntax-semantics encoding and of our theory of control. If, on the other hand, V_2' also takes sentential complements (and hence V_2' can apply to propositions), then we should expect both the sloppy and non sloppy reading to be possible, in principle at least. So in $V_1'(P) \dots V_2'(\underline{it}')$ where V_2 can take also sentential complements it might either refer back to the property P directly (in which case as in the preceding example we will have the sloppy reading, i.e. P will have different understood subjects in V_1 and V_2) or to some proposition $P(y)$ available in the context, in which case we will have a non sloppy reading. Let us illustrate all this by means of examples, which will also allow us to test such a prediction.

Let us consider the second case first. The verb type that we need is well exemplified by want. Want takes S[-TNS] (I want (for) Mary to do the dishes) and what on our theory are VPs (I want to do the dishes). Want, in many dialects, also licenses it-pronominalization.

According to the preceding prediction, then, an occurrence of it as an argument of want in the appropriate context should allow for both a sloppy and a non sloppy reading. This seems to be correct for many dialects. Consider:

- (26) John wants to become chairman of the department and Mary wants it too.

The underlined it in (26) can mean either for John to become chairman (strict reading) or for Mary to become chairman (sloppy reading). On our account, the first reading is obtained by letting the it refer to the proposition for John to become chairman made available in the context by applying CP to the first conjunct. The second reading would be obtained by letting it refer to the infinitival to become chairman. Then, CP would yield that the understood subject of this property has to be Mary. So, want behaves just as expected with respect to it-pronominalization.

It should be possible to make the same point for gerunds as well. Gerunds without possessives denote nominalized properties, while gerunds with possessive subjects denote, roughly, propositions. So, take a verb like like:

- (27) a. John likes playing tennis
b. I like his always being punctual

In (27a) like occurs with a subjectless gerund and it displays all the characteristics of obligatory control. In our terms, like is positively specified for CP. In

(27b) like takes a gerund with a subject. Gerunds with subjects in our theory denote states of affairs; hence CP cannot apply to (27b). Given this, we should expect like to behave similarly to want with respect to it pronominalization. Indeed, this seems to be so. Consider:

(28) John likes fooling around, but I would bet that his wife doesn't like it at all

Again, it would seem that the underlined it in (28) can be interpreted as either John's fooling around or as his wife's fooling around. The reasons why this is precisely as expected are just the same as in the case of want. Notice, that I am not claiming that in all the cases of this sort the sloppy and non sloppy readings should be both equally preferred. There are many other complicated factors that may come into play in selecting pronoun antecedents in discourse. What we are concerned with here is simply the possibility of certain readings occurring.

In that regard, as pointed out above, there are cases where our theory would predict strict readings to be absolutely impossible. Those are the cases of verbs that take only verbal arguments and license it-pronominalization. Now, in English it turns out to be quite difficult to test this prediction. If we look at infinitives, the only verbs subcategorized for them but not for \bar{S} 's appear to be verbs like try and verbs like force; unfortunately these verbs do not license it-pronominali-

zation. It should be possible, however, to find languages where such a prediction could be tested. In fact, there seems to be plenty. Italian, for example, is one such language. The Italian counterparts of verbs like try and force behave like the English ones in all relevant respects. In particular they take infinitival complements but not \bar{S} -complements. And, more interestingly, they allow for an analogon of it-pronominalization. Hence, we would expect the strict reading of pronouns to be totally out with those verbs. Consider:

(29) Paolo ha tentato di fuggire di prigione ieri, e
 Marco lo tenterá domani
 Paolo tried to run away from prison yesterday
 and Marco will try it tomorrow

(30) Prima Paolo ha costretto Maria ad andarsene, poi
vi ha costretto anche Angela
 First, Paolo forced Maria to go away, then (he)
 forced to it (obl.) also Angela.

The strict readings of the underlined pronouns in (29)-(30) (whereby lo would stand for Paolo's running away and vi for Maria's going away) are impossible. Only the sloppy interpretations are allowed. The reason why this is exactly as it should be according to our theory is simply that try' and force' take only properties (and not states of affairs) as arguments. So a pronoun in their argument structure will be able to refer back only to properties not to states of affairs. Since these are control verbs, CP will then apply to them putting the understood subjects in place. Notice that it would be

quite implausible to blame the lack of strict readings on some idiosyncratic property of try or force. It is perfectly possible to see what, say, the second conjunct of (29) should mean for such a situation to obtain. It should mean something like "Marco will try to bring it about that Paolo runs away". Why is such a meaning disallowed precisely with verbs that do not take \bar{S} -complements? Our theory is able to offer a plausible semantic explanation of such a situation.

It would seem quite difficult to account for the behavior of pronouns referring to verbal complements for theories where these constructions are always clausal. The relevant distinction (between property denoting and proposition denoting constructions) is simply lacking in such a theory. Even if we assume (with Williams) that verbal arguments are syntactically clausal but semantically properties, still the situation wouldn't improve much. It would seem that such a theory would have to stipulate that in the cases where both the strict and the non strict readings of pronouns are allowed, the interpretation of PRO (i.e. control) is freely ordered with respect to the interpretation of pronouns. In the cases where, however, only the sloppy reading is possible something has to ensure that the interpretation of PRO takes place only after the interpretation of pronouns, so that PRO is not bound when pronominalization takes place.

Now why should precisely this situation obtain? On our approach no stipulation of this sort is necessary, and the lack of strict readings with certain verbs receives a rather straightforward and principled account. It may be that even in English the above prediction might be tested. Gerunds in general allow for it pronominalization much more freely than infinitives. So if we could find some control verb that can take only subjectless gerunds, it should provide the relevant case: such a verb should allow only sloppy readings of pronouns. A verb of this sort should be practise. Consider:

- (31) a. I practised playing sonatas
 b. *I practised his playing sonatas
 my

According to our theory, the pattern in (31) shows that practise is a verb that can only take properties as arguments and that furthermore obeys CP. Hence we should expect an it in the argument structure of practise to license solely sloppy readings for the reasons given above in connection with the Italian examples. This seems to be exactly right:

- (32) John practised swimming regularly for two years
 and Bill practised it for 3

There is no way for the underlined it to mean John's swimming, as expected. I am told that at least in some dialects the same case can be made even for gerunds which

are not NPs, such as those that verbs like begin subcategorize for. Consider:

- (33) (?) John began swimming two years ago, while
his brother began it 3 years ago.

In those dialect where (33) is possible, only the sloppy reading of it is allowed, as our theory would predict.

So our theory of verbal arguments leads us to predict in a totally straightforward way which class of verbs licenses both sloppy and strict readings of pronouns anaphorically related to verbal arguments, and which class of verbs can only license sloppy readings. Not only does a similar prediction not seem to fall out from theories that have PROs, but it appears to be difficult to extend such theories in a natural way so as to accommodate these facts.

It should perhaps also be recalled that there are other ways in which PROs do not behave like other "empty nodes" commonly assumed to exist (such as wh-traces). For instance they never block contractions and they do not display weak cross-over effects. Consider the following:

- (34) Which boy_i does his_i mother like t_i?

(34) is said to be bad. Whatever rules (34) out must not rule out the following:

- (35) PRO_i voting for himself damaged every candi-
date_i

(35) is fine on the reading indicated by the subscripts; such a reading would give rise to weak cross-over con-

figuration, which for both wh-traces and variables 'left behind' by quantifier raising are said to yield ungrammatical outputs. For our theory, these are non problems because there are no PROs.

A final point which is worth raising has to do with equative and quasi-equative sentences, such as the following:

(36) John is the chairman of the department

The is in (36) is commonly analysed as being semantically the identity relation. Just as terms referring to individuals can flank the be of identity, we should expect terms referring to properties to be able to. In other words, we should expect phrases that have the form $P = Q$ to be just as possible as (36) (where P and Q range over properties). Of course, by the very meaning of '=', the 'understood subjects' of P and Q will have to be the same. This prediction is borne out:

(37) to know him is to love him

The logical form of (37) would simply be:

(38) love him' = know him'

(38) obviously entails that anybody who has the first property has the second and vice versa. The same holds for quasi-equative sentences such as the following:

(39) winning a game involves training hard.

On the present view, the logical form of (39) would be something roughly like the following:

(40) winning a game' \subseteq training hard'

(40) will obviously entail that anybody who has the first property has the second, but not vice versa. The fact that the 'understood subjects' in (40) are felt to be anaphorically connected (i.e. equal one-by-one) is a straightforward consequence of our theory.

What would a theory that has PRO have to do to ensure a similar result? It is not clear. One possibility would be to assume that there is some operator in the matrix COMP position binding the two PROs roughly as illustrated below:

(41) O_i [PRO_i to know him is PRO_i love him]¹⁰

However in view of what we have said so far (especially in connection with the analysis of PRO_{arb} in the preceding section) this move would give rise to a number of serious problems. The operator in (41) whatever it may be would come in the way of any attempt to state properly the behaviour of it referring back to verbal arguments.

In conclusion, there seem to be a number of anaphoric processes involving verbal arguments that fall quite naturally into place within our theory, and, at the very best, are not so naturally accounted for within theories that postulate a clausal character for infinitives and gerunds. This seems to provide further support to the idea that verbal arguments are neither clausal in the syntax nor semantically associated with propositions.

1.3. A few more sound arguments.

In the present section we are going to consider some purely semantic data that we think militate against propositional analyses of verbal arguments. In particular, we are going to consider some valid entailments (i.e. sound arguments) that our theory predicts to be valid, while their intuitive validity appears to be at best mysterious on a propositional analysis. Consider first the following.

(42) Anybody that John likes, Mary likes
 John likes Bill
 Mary likes Bill

The validity of (42) is unquestionable. It is also evident why (42) is valid. The first premise says that any person (i.e. a real thing out in the world) such that John stands in the like-relation to him is such that Mary stands in the like relation to him too. So the logical form of (42) would simply be:

(43) $\forall x$ [John likes $x \rightarrow$ Mary likes x]
 John likes Bill
 Mary likes Bill

Consider now the following:

(44) Anything that John likes (whatever John likes),
 Mary likes
 John likes jogging
 Mary likes jogging

(44) is just as valid as (42). Also the reasons for their validity seem to be just the same. The first premise of (44) simply says that any entity (out in the real world)

such that John stands in the like-relation to, Mary stands in the like relation to. The second premise says that the activity of jogging is one such thing. Hence the conclusion follows. On our theory, gerunds are analysed as (nominalized) properties. So the validity of (44) is predicted straightforwardly. The logical form of the first premise of (44) is just like the first premise in (43) (recalling that nominalized properties are legitimate values of individual variables in IL_*). Nothing has to be said concerning why (44) is recognized as a valid entailment by speakers of English.

If, however, verbal arguments are analyzed as being semantically propositional, then the validity of (44) immediately becomes a mystery. Clearly the propositions or state of affairs to which John and Mary stand in the like-relation to are not, in general, the same. If John likes jogging, then the state of affair to which he stands in the like-relation to is one when he (John) jogs. How come, then, we can infer from the second premise in (44) the conclusion that Mary likes jogging too? Notice that this problem seems to be independent of the particular choice of items that one might want to assign as values to for-to clauses. Either such an item is a "closed" entity as in general propositions are taken to be, and then the validity of (44) is a puzzle. Or the value of verbal

arguments is an "open" entity (in our terms, a nominalized property), and then our point is established.

Arguments of a form parallel to the one of (44) can be built ad libitum for constructions involving both infinitives and gerunds with all garden varieties of control predicates. Let us give a few more examples.

- (45) a. Anything that is fun (easy, etc.) for John,
is fun for Mary
to jog is fun (easy, etc.) for John
to jog is fun (easy, etc.) for Mary
- b. Whatever John tried, Mary tried
John tried to jog
Mary tried to jog
- c. Whatever John promised to Mary last night,
Bill promised last night
John promised to marry her
Bill promised to marry her

On our approach, all of (45a-c) are predicted to be valid for the very same reasons that (44) is predicted to be valid. Infinitives denote (individual projections of) "open structures" (i.e. properties) just like gerunds do, even though they might differ semantically in other regards. It is precisely this feature that accounts for the validity of the arguments under consideration. The examples in (45a-c) also show that the point that I am making is totally independent from the kind of control predicate (strict or implicit control) involved, from the distribution of PRO_{arb} and from considerations of tense and aspect. It applies to verbal arguments in virtually all their manifestations.

Of course one might engage oneself in the elaboration of some fancy interpretative procedure that could somehow rescue propositional theories of verbal arguments. I am sure that something that works might eventually be found. Otherwise, there is always pragmatics to appeal to. Allow me, however, not to try to engage myself in trying to imagine what such an account might be. The point I am making is quite simple. The validity of a certain number of arguments is predicted by our theory and is not predicted by propositional theories as currently conceived. I take this as supporting evidence for the approach I am trying to defend.

I should perhaps conclude by emphasizing that the argument developed in the present section militates against any propositional theory of verbal arguments, i.e. against any theory that associates verbal arguments with propositions, states of affairs and the like as their semantic value. So in particular, it provides disconfirming evidence for GB approaches, LFG (on the basis of the interpretation procedure developed in Halvorsen, 1982), and some version of MG (such as the one of Bach and Partee 1980). I don't know whether the present argument represents a problem for theories where verbal arguments are analyzed as being syntactically clausal but semantically associated with properties (as in Williams' theory of predication).

1.4. Summary.

In preceding chapters we have argued for a VP = P analysis of verbal arguments mostly on the basis of distributional characteristics of verbal arguments and on the basis of control phenomena. Here we have considered a number of anaphoric phenomena involving verbal arguments. In particular we have to consider the behavior of so called PRO_{arb} constructions, the sloppy and non sloppy readings in it-pronominalization phenomena and some valid entailments that constructions involving verbal arguments license. All these phenomena receive a straightforward and simple account on the basis of the present approach. Even if the same range of data can be accommodated in various alternatives to the theory developed here, it appears difficult to believe that they can be dealt with as straightforwardly.

2. Internal Anaphora

In the present section we will focus on the behavior of verbal arguments with respect to opacity, i.e. the well known complementary distribution of reflexive and non reflexive pronouns within certain local domains. This will lead us to consider various aspects of a theory of binding. Since, however, we are interested here solely in the behavior of verbal arguments with respect to binding, we will not be able to consider many of the intriguing

complexities of binding phenomena. In particular, we will not discuss opacity phenomena within NPs and how they relate to opacity within clausal domains. We will also be unable to consider conditions on wide scope quantification, island phenomena, conditions such as "leftness" (Jacobson 1977) and the like. The range of facts we will consider is limited to various non coindexing constraints within Ss (Lasnik, 1976, Reinhart, 1976, 1979) and to the interpretation of reflexives. Our starting point will be the theory developed in Bach and Partee (1980) -- BP from now on. We will first outline the basic ideas put forth in BP. Then we will consider several aspects in which such a theory turns out not to generalize smoothly to all the constructions considered in the past chapters, and try to provide an alternative in the same general spirit.

2.1. Non reflexive pronouns.

The central task of any theory of binding is to characterize the domain in which two pronouns cannot "corefer". Such domains are the very same as those within which reflexives have to have an antecedent. Achieving this task implies individuating the relevant notion of "distance" and "locality" in terms of which the domains in question can be optimally characterized. Within the present framework, this has to be accomplished within the straitjacket of the rule-by-rule hypothesis, which also

means that we will have to spell out explicitly what the truth conditional import of claims about binding actually is.

What appears to be the basic insight of BP is that opacity domains cannot be properly individuated in terms of syntactic configurations. The "configurations" that can do the job are those that we need to do semantics, namely function argument structure. With respect to pronouns, a rough but good first approximation seems to be that they cannot 'corefer' (or cannot be 'coindexed') when they are arguments of the same function of some specified type (essentially, predicative functions -- see below).

The way BP implement this idea is by assigning an index to (the translation of) each pronoun and keeping track of these indices by storing them in a local pronoun store (LPS). In combining a function with an argument a general condition will require that the LPSs associated with the two to be disjoint. LPSs are emptied at each S and NP node. In this way a sentence such as he loves him cannot be associated with the IL translation in (46a), but can be associated with the one in (1b):

- (46) a. love'(x₁)(x₁)
 b. love'(x₁)(x₂)

Nothing prevents, it should be noted, x₁ and x₂ from being mapped into the same real-world individual. But this is probably as it should be. Setting up an appro-

priate context, it is possible to use sentences like the one in question where two 'locally free' pronouns turn out to refer to the same individual. Of course, such contexts are rather far fetched and one might want to build that into the grammar by either making (46b), say presuppositionally deviant whenever x_1 and x_2 are mapped onto the same individual, or by developing a "dynamic" theory of pronoun interpretation a la Kamp (1981). However, we will not try to carry on either within the limits of the present work.

The reason why it is necessary to keep track of pronoun indices in the syntax is to avoid generating (47a) with the reading in (47b):

- (47) a. everyone loves him
 b. $\forall x[\text{love}'(x)(x)]$

Recall that the way quantification (and wh-movement) is treated in the present framework is by using a Cooper-store device, whereby an NP can be associated with a variable and its meaning is stored along with the variable index (to be, later on, pulled out of store at the relevant "binding node"). In storing an NP meaning, one might by accident catch a variable index associated with a pronoun in the same local context, thereby generating something like (47b) as meaning for (47a). This can be easily avoided on BP's approach by assuming that the pronoun meaning of a stored NP works like any other

pronoun meaning, and that therefore its index is put in LPS. Then the usual disjointness requirement on LPSs will guarantee that it is impossible to associate with sentences like (47a) something like (47b).

Let us state all this in slightly more general form. We know that predicates have adicity; by internal argument of a predicate P^n , we mean something that saturates one position in the predicate, turning it into P^{n-1} . Natural languages have also devices to add arguments to predicates, typically PPs. An argument which is "added" by a modifying PP to a predicate will be called, as usual, an external argument. It should be recalled that certain PPs are really NPs with some sort of oblique marking. In the present framework, this means simply that they have an NP-meaning and saturate an internal position in a predicate. With the label "argument" we will refer to either an internal or an external argument of a verb. Finally, let us define the predicate domain of some predicate P as the sets of its arguments. BP's approach can now be summarized simply as follows:¹¹

- (B^f) Two pronouns cannot have the same index (i.e. must be free) in the same predicate domain (i.e. when their meanings occur within the predicate domain of the same predicate).

B^f is nothing but another way of stating what BP do. Now what exactly is principle B^f a principle on? On the one hand it exploits a semantic characterization of a

certain domain, on the other it requires indices to be different. Where are these indices? Quite clearly, they are not model theoretic entities. They could be turned into such, but at the cost of unbearable artificiality. They are not in the surface structure of phrases either, at least as we perceive it. In fact, Landman and Moerdijk (1983) raise the question whether a principle like B^f in BP's implementation would not make the syntax of logical form a crucial level of linguistic theory.

This is so because the indices stored in LPS are neither model theoretic entities nor syntactic items (because of the well-formedness constraint). So they must be properties of the translation of pronouns into IL, which makes IL more than a simple device for picking out model theoretic entities. The point is well taken. BP's approach indeed requires either assuming as crucial a level of logical form or a weakening of the well-formedness constraint. Both ways to go appear a priori possible. Landman and Moerdijk choose the second alternative, as we will also. However, what they do is build a configurational approach to anaphora (essentially Reinhart 1976) into a Montague-like grammar. We believe, though, that the basic principles of the binding theory cannot be properly stated configurationally and hence we will not follow Landman and Moerdijk in their attempt.

What we propose to do, essentially, is to enrich the syntactic labelled bracketing by means of an indexing device. We will assume simply that each NP carries a numerical index, in pretty much the way in which it carries a number of feature specifications.

For pronouns, we follow Landman and Moerdijk in assuming that a pronoun he_i is translated, by general convention, as x_i . For other NPs, we take their index to be the index of the pronoun meaning that would have to be used if they were stored (i.e. quantified in). Then we will be able to maintain essentially BP's implementation of principle B^f , and assume that the grammar keeps track of NP indices via LPS and keeps them disjoint in the relevant local domain (i.e. in their predicate domain). So principle B^f is a syntactic condition on binding, that uses a semantic characterization of locality. The requirement that LPS be emptied at S and NP level is now understood as a convention that deletes indices from NPs. So, no indexed phrase will show up eventually. We regard numerical indices of NPs as a special sort of syntactic feature. It is pretty obvious, though, that these numerical "features" will not obey the quite tight principles that constrain feature matching and percolation in the present framework. For morphological features, island constraints are automatically determined once and for all by functional application. Statements that have the

effects of the following "delete feature γ at node A" are disallowed. This is not so for numerical indices. What this means, essentially, is that a theory of binding cannot be reduced to a theory of government and percolation. Different sets of principles (and a different use of function argument structure), illustrated by principle B^f , is called for by the former. This is reflected in the different mechanism that will be used to deal with binding.

It would be hard to deny that the present approach has something in common with various others, such as Reinhart (1976), Williams (1980) or Chomsky (1982). The evident and important difference lies mostly in the fact that function argument "geometry" rather than constituent structure geometry is taken to be the crucial factor in establishing which anaphoric linkings are possible and which aren't. There are a few cases in English where the two approaches would seem to make different predictions. Such cases might be exemplified by the following sentences:

(48) I talked to him about him

(49) I counted on him to keep him quiet

The underlined pronouns in (48) and (49) have to be 'locally free', i.e. non coindexed. Let us consider briefly what each of the above mentioned theories has to say about this fact. Consider first Reinhart (1979):

- (50) An NP which is not a pronoun must be interpreted as non coreferential with any pronoun that C-commands it.¹²

This principle works for most cases but leaves unexplained why the two occurrences of him in (48) have to be disjoint.

Let us consider next Chomsky (1982). The principle that should take care of cases like (48) is principle B of the binding theory (see e.g. Chomsky, 1982, pp. 180 ff):

- (51) A pronoun must be free within its governing category.¹³

If we assume that C-command is a condition for binding (as Chomsky, 1982, p. 184 does), then even if the two pronouns in (48) would be coindexed they would not count as 'bound' on the basis of (51) and their non coreferentiality would remain unexplained.

Let us consider finally Williams (1980). According to Williams, the infinitival in (49) can have both the subject and the on-PP as its controller. Hence this construction must be a case of non obligatory control in his sense, which is subject to his Arb-rewriting rule I. If, however, Arb-rewriting selects the subject as the controller, we should expect the two occurrences of him to be possibly coreferential according to Williams' definition of opacity. But they never seem to be.

It is evident what the "cheating points", as BP put it, for both configurational and semantically based theories of anaphora are going to be. Configurational

theories have to provide syntactic analysis that can make their principles work; semantically based theories have to analyze function-argument structure to make things come out right.

So, one might argue that after all both theories have to say something special about the PPs in (48) and (49) and that is all. However, there does seem to be a difference. In BP's theory, the crucial principle is stated in terms of function argument structure; since certain departures of function argument structure (a semantic universal notion) and constituent structure are expected, cases like (48) and (49) where this occurs will not hit "the heart" of the principle. On the other hand, configurational theories such as those we have considered state their principle in terms of phrase structure configurations. So cases like (48) and (49) will hit the heart of their principles, and some abstract level of syntactic structure will have to be postulated. In a sense BP's theory amounts to saying that the only level of "abstract" structure that one needs is the one that is needed anyway to do semantics.

There are, furthermore, a number of empirical problems that get in the way of any appeal to reanalysis for cases such as (48) and (49). For instance, if reanalysis is what triggers non coreference in (48), then it should

make a difference whether the PP is preposed or not. But as BP point out, it doesn't:

(52) To him I talked about him

The situation is similar, only worse, for languages like Italian where cases analogous to (48) and (49) also arise:

(53) Ho chiesto di lui con lui
(I) asked about him with him

(54) A lui piace parlargli
to him it likes speaking to him

The underlined pronouns have to be disjoint in reference in (53) and (54). Appeal to reanalysis in Romance languages cannot be justified by preposition stranding phenomena (see on the impossibility of prepositions undergoing reanalysis in Romance Kayne, 1981).

There is a more important respect, however, in which the two classes of theories we are discussing diverge. It has to do with languages that have little or no configurationality. For these languages the theories under consideration would seem to make wildly different predictions. Function argument structure is universal. So on BP's theory non coindexing constraints are expected to work roughly as in English in those languages. If however these constraints (and in general the binding theory) look at phrase structural configurations so crucially, as GB and related approaches claim, then either they are expected to work in a radically different way in non configurational languages, or some English-like abstract

syntactic structure will have to be created. We hit here basically the same problem that we found in connection with a configurational theory of government. We therefore conclude with BP that the domain of binding principles (such as non coindexing constraints) is best captured in terms of function argument structure.

2.2. Reflexives.

The way BP propose to treat reflexives is by building in the grammar a mechanism that forces them to be 'bound' locally, i.e. in their predicate domain. The set of potential 'binders' of reflexives has to be somehow constrained by some "superiority" condition as to avoid things like *himself_i loves John_i. Again, the crucial difference between BP's and GB-like approaches seems to be that superiority conditions on reflexive-binding are stated in terms of function-argument structure (in BP's approach) rather than in terms of C-command. In particular, BP seem to claim that only the last two positions in a predicate (namely those that correspond to objects and subjects) can bear an antecedent for reflexives. This last claim, however, appears to be problematic in view of examples like the following:

(55) John sent to Mary a letter about herself

In (55) clearly an oblique NP acts as antecedent for a reflexive pronoun. This would seem to show that internal

- (A^f) reflexives must be bound in their predicate domain. (An NP α is bound in its predicate domain β iff either it is an external argument in β coindexed with an internal argument in β or it is an internal argument coindexed with an f-commanding argument).

The rest of this section will be devoted to providing an implementation of principle A^f within the recursive and truth conditional setting of our grammar-format. We will elaborate on BP's apparatus by introducing into it one of the central insights of Williams (1980). What we will get in the end is nothing more and nothing less than one way of spelling out principle A^f. So, the reader satisfied with such principle in its present form may skip the rest of this section.

The difficulty in implementing A^f is that we now recognize the possibility that the antecedent of a reflexive doesn't have to come in after the reflexive itself. The most plausible function-argument structure associated with (55) and (56) above appears to be:

(55') send'(Mary')(a letter about herself')(John')

(56') (about'(himself'))(talk-to'(John'))(I')

Within our framework, function-argument structure induces the order in which expressions are combined. This means that in the cases of (55) and (56) the antecedent of the reflexive is going to be already "there" when the reflexive comes in. This means that we will have to keep track

of potential reflexive antecedents (i.e. internal arguments) till they cease to be such.

A related difficulty is that one potential antecedent for reflexives is the argument that corresponds to the subject, encoded in the present framework as the last argument of a predicate (which could be regarded as the most external among the internal arguments). However one of the things that we have been arguing all along is that infinitives and gerunds have no subjects because they are verbal constituents. It follows that reflexives within verbal arguments may have to be bound by arguments that are not there.

So reflexives seem capable of having antecedents which are already "buried" inside the predicate in which they occur as well as antecedents that will never be part of the predicate domain to which they belong: a rather appalling situation. However, I think that there might be a quite elegant way out.

So far we have assumed that there isn't any special syntactic counterpart to a predication relation. From a semantic point of view such a relation is regarded here as a special case of functional application: the one that obtains when a 1-place property is applied to an argument. We have acted as if nothing special needs to be said in the syntax about predication. Williams (1980) argues instead that there is a special syntactic relation

of predication, represented on phrase markers by coindexing predicates and their subjects. If he is right, we can provide a rather simple implementation of the principles of binding A^f and B^f along the following lines.

We, first of all, have to assume that if k is the index associated with an NP α , then whenever α combines with an NP-taking function β , k is marked as k_i (i.e. internal) or k_e (i.e. external) according to whether β is an 'argument-adder' or not. To exemplify, we will have:

- (59) a. $\langle \text{love him, IV, love}'(x_3), \{3_i\}, \emptyset \rangle$
 $\quad \langle \text{love, TV, love}', \emptyset, \emptyset \rangle \quad \langle \text{him, NP, } x_3, \{3\}, \emptyset \rangle$
 b. $\langle \text{about him, IV//IV, about}'(x_3), \{3_e\}, \emptyset \rangle$
 $\quad \langle \text{about, (IV//IV)//NP, about}', \emptyset, \emptyset \rangle \quad \langle \text{him, NP, } x_3, \{3\}, \emptyset \rangle$

This gives us a way of keeping track of which pronouns in LPS are external and which are internal arguments.

Furthermore, we assume that every predicate carries an index that must be matched by the index of its subject. This is what we take from Williams' proposals. Formally, we realize this by assuming that each item α that belongs to a category $\frac{A}{X}_{+Pred}$ (where X might possibly be null) carries an index α_i that is also in the LPS of α . So we have phrases of the following type:

- (60) a. $\langle \text{run}_j, \text{IV, run}', \{j_i\}, \emptyset \rangle$
 b. $\langle \text{love}_j, \text{IV/NP, } \{j_i\}, \emptyset \rangle$
 c. $\langle \text{be tall}_j, \text{IV, } \lambda x [\text{'tall}'(x)], \{j_i\}, \emptyset \rangle$

Finally we assume that reflexives translate simply as pronouns, i.e. as x_i for some index i ; however they

trigger the introduction in LPS of superscripted indices i^s (s for self). So in LPS we will have indices of the form k^s (reflexives),¹⁵ k_i (internal arguments) and k_e (external arguments).

We then assume the following pair of binding rule schemata:

- (61) a. $IV \rightarrow IV$
 1 2 conditions:
 $k^s, j_i \in LPS(1)$
 $1 \notin LPS(1)$
 if k^s is internal, j
 f-commands k
 $LPS(2) = LPS(1) - \{k^s\}$
 b. $2' = \lambda x_1 [1'(x_1) \ \& \ x_k = x_j]$

This is the reflexivization rule; it "binds" a reflexive pronoun to an (f-commanding) internal argument and takes its index out of LPS.

- (62) a. $IV \rightarrow IV$
 1 2 conditions:
 there is no k^s such that
 $k^s \in LPS(1)$
 $LPS(2) = \emptyset$.¹⁶
 b. $2' = \lambda x_j [1'(x_j)]$
 where j is the predicative index of 1.

(62) is the rule that empties the local pronoun store of a VP right before it combines with a subject; this rule also puts into place (i.e. last argument slot) the predicative index of the VP. Probably it is easier to understand how this mechanism works (and yields the desired results) by means of a few examples.

- (63) a. John loves himself
 b. $\langle \text{John loves himself}, S, \lambda x_1 [\text{love}'(x_2)(x_1) \ \& \ x_2 = x_1](j), \emptyset, \emptyset \rangle$
 $\langle \text{John}, NP, j, \{1\}, \emptyset \rangle \langle [\text{love himself}]_1, \lambda x_1 [\text{love}'(x_2)(x_1) \ \& \ x_1 = x_2], \emptyset, \emptyset \rangle$

rule (62) $\langle [\text{love himself}]_1, \text{IV}, \lambda x_1[\lambda y[\text{love}'(x_2)(y) \ \& \ x_2 = x_1](x_1)], \emptyset, \emptyset \rangle$

rule (61) $\langle [\text{love himself}]_1, \text{IV}, \lambda y[\text{love}'(x_2)(y) \ \& \ x_1 = x_2], \{1_i\}, \emptyset \rangle$

$\langle [\text{love himself}]_1, \text{IV}, \text{love}'(x_2), \{1_i, 2^s\}, \emptyset \rangle$
 $\langle \text{love}_1, \text{TV}, \text{love}', \{1_i\}, \emptyset \rangle \langle \text{himself}, \text{NP}, \{x_2, 2^s\}, \emptyset \rangle$

As is clear from the example, we require that whenever we apply an $\text{IV}\alpha$ to an $\text{NP}\beta$ (a) their indices match and 'cancel out' and (b) the LPS of α be empty. We also require that indices of all NPs in the same predicate domain are kept disjoint, till LPS is emptied by rule (62). Rules (61) and (62) are meant to replace BP's REFL_1 and REFL_2 and the conditions on the storage and retrieval of these items. It should be clear how (61) and (62) enable us, in principle, to deal with any case of the sort illustrated by (55) and (56). Let us give a sample (partial) derivation of (55) to illustrate this.

(64) $\langle \text{John sent to Mary a letter about herself, S, send}'(m)(\text{a letter about } m)(j), \emptyset, \emptyset \rangle$

$\langle \text{John sent to Mary a letter about herself, S, "Mary" } \lambda x_1[\text{send}'(x_1)(\text{a letter about } x_3)(j) \ \& \ x_3 = x_1], \emptyset, \emptyset \rangle$
 $\langle \text{John sent to Mary a letter about herself, S, send}'(x_1)(\text{a letter about } x_3)(j) \ \& \ x_1 = x_3, \emptyset, \langle \text{Mary}', 1 \rangle \rangle$
 $\langle \text{John, NP, } j, \{2\}, 0 \rangle \langle [\text{sent to Mary a letter about herself}]_2, \text{IV}, \lambda x_2[\text{send}'(x_1)(\text{a letter about } x_3)(x_2) \ \& \ x_3 = x_1], \emptyset, \langle \text{Mary}', 1 \rangle \rangle$

<[sent to Mary a letter about herself]₂,
 λy[send (x₁)(a letter about x₃)(y) &
 x₃ = x₁], {2_i, 1_i}, <Mary', 1>>

|

<[sent to Mary a letter about herself]₂, send
 (x₁)(a letter about x₃), {2_i, 1_i, 3^s},
 <Mary'1>>

/ \

<[sent to Mary]₂, TV, send'(x₁), {2_i,
 1_i}, <Mary', 1>>

|

<a letter about herself, NP, a letter about
 x₃, {3^s}, ∅>

Examples like this could not be generated by the mechanism developed by BP. So, the rules we are proposing appear to give better results than those that can be achieved within the framework of BP, while being in the same spirit as their proposal.

Essentially, rule (61) and rule (62) provide one way of spelling out Williams' dictum: predicates are opaque. What we have done is first to provide the local domain of opacity in terms of function argument structure and second to provide an implementation of the relevant principles of the binding theory that makes essential use of the notion of (1-place) predicate. This has led us to introduce in our framework a novel indexing device that requires predicates to bear the same index as their arguments, in terms of which we have reformulated, with some empirical advantage, the theory developed by BP.

We leave it open, for the time being, the way the notion of opacity extends to NPs. Within the present framework this would amount to figuring out the conditions

under which the LPS of NP's should be emptied. In this regard, the proposal of Williams (1980) that "definite" NP's are opaque appears to be intriguing. Barwise and Cooper (1981) have illustrated various ways in which "definiteness" might be defined in semantic terms. It would be quite interesting if the notion of "definiteness" relevant for opacity in NPs could be characterized in the precise and general way Barwise and Cooper propose. However, we have to leave this for further investigation.

3. Opacity and Control

3.1. A "propositional" account.

It is one of the basic insights of early transformational accounts of English complementation that verbal arguments behave with respect to opacity as if they had a subject. This insight has in fact provided one of the strongest arguments for assuming that verbal arguments indeed have a phonetically unrealized syntactic subject. In the present section we wish to investigate how the account of opacity we have developed in the preceding section, together with our theory of control make the appeal to PRO's unnecessary. Given the arguments we have developed throughout this work, this is all to the good.

BP develop an extension of their account of non-co-indexing constraints to control structures which in fact accepts the standard transformational view that verbal

arguments are clausal (or propositional). Their approach can be informally described as follows. They assume that there two kinds of infinitives: subject and object control, just as, according to them, there are two kinds of reflexives. So what they do is to generate infinitives with a semantic subject (i.e. a pronoun meaning is plugged in the subject position slot of infinitival verbs) and to put in store the devices used to deal with reflexive-binding (i.e. REFL₁ and REFL₂). Infinitival taking verbs are marked with a feature +OBJ CON or +SUB CON; their infinitival complements have, obviously, to match such feature.¹⁷ As soon as a verb has combined with an infinitival complement, various conditions require that REFL₁ or REFL₂ are retrieved out of store. The effect of this will be the identification of the "null subject" of infinitives with the next argument that the verbal complex is going to combine with. This approach ensures the right results with respect to opacity and control for verbs like promise or force.

BP's approach, however, runs into serious problems, I believe. Some of them are of a general nature. Either infinitives are syntactically verbal constituents of some sort, in which case BP's proposal represents a substantive weakening of the quite tight category-type relation we have been arguing for, for infinitives are going to be associated with propositions on their approach. Or

infinitives are syntactically S's, in which case BP's approach would appear to be a notational variant of current GB approaches. In either case, such an approach would suffer from all the drawbacks that a propositional theory of verbal arguments suffers from.

Independently of this general point, however, there are more specific reasons for doubting that BP's approach could extend successfully to deal with other cases of control considered here.

A first reason has to do with the fact that just as there does not seem to be only subject and object control reflexives, so there seem to be control predicates that are neither subject nor object control, as we have tried to show in Chapter IV. So consider again:

(65) John recommended to every student_i shaving
him_i

On the reading where every student binds him as indicated by the subscripts, (65) seems to be as bad as (66) (on the analogous reading):

(66) John forces every student_i to shave him_i

Furthermore, (65) and (66) seem to be bad for just the same reasons (i.e. opacity). However, while (66) on the intended reading could not be generated by BP's grammar, there does not seem to be any way to rule out (65), if, as argued in Chapter IV, its function-argument structure is:

(67) recommend'(every student')(shaving him')(John')

In (67) the controller of the gerund is neither a subject nor an object while BP's approach can work only for subject or object control structures. So there does not seem to be any way to extend their approach as to cope with (65) and related cases.

What is possibly even more worrying about BP's approach is the existence of opacity phenomena in so called "small clause" constructions. Consider the following:

(68) John considers every man_i proud of him_i

(69) John thinks of every man_i as a friend of his_i

Again something has to prevent the quantifier every man from binding him (as indicated by the subscripts) in (68) and (69). Again whatever mechanism accomplishes the ruling out of (66) should also be able to rule out (68) and (69), for they seem to be bad for the very same reasons. Notice that replacing every man with a pronoun in (68) and (69) yields the expected "disjoint reference" effects, while it is not so for:

(70) He considers John proud of him

(71) He regards John as a good friend of his

The underlined pronouns in (70) and (71) can be coreferential. What makes these constructions especially problematic is the non constituent behavior of the sequences NP AP and NP ^{NP} +pred', and the passivizability of the object NP. On the basis of BP's approach, the only way to

cope with the opacity effects induced by "small clause" constructions would seem to be regarding them, in effect, as small clauses. So the AP in (68) and the predicative NP in (69) would have to be analysed as being clausal (or semantically associated with propositions) to trigger the right non coreference effects. Within a Montague like grammar, this move has the obviously undesirable consequence of virtually losing completely the simple relation between syntactic categories and semantic types that represents one of the original strengths of the framework. Furthermore, Williams (1983) has provided a number of totally independent reasons for rejecting a small clause analysis of the constructions under consideration.

So, quite apart from the general problems that a clausal or propositional analysis of verbal arguments seems to raise, there appear to be various more specific problems for the theory of opacity and control developed in BP.

In the preceding section we have developed an approach to opacity based on the motto "predicates are opaque." Such a theory should allow us to develop an approach to opacity in control structures that does not give rise to the problems just considered. In particular, our theory was based on the recognition of a syntactic counterpart of the semantic notion of predication built into our framework, represented as a form of coindexation

between a predicate and its subject argument. Now, our theory of control brings about a form of indirect predication. Consider the following example:

- (72) a. John tries to run
 b. try'(^run')(j)
 c. □_{try}run'(j)

(72b) represents the function argument structure associated with (72a). run' is not attributed to John in (72b). However, (72c) represents the entailment brought about by CP and in (72c) run' is predicated of John. So there is a clear sense in which control predicates systematically bring about an (indirect) form of predication involving their verbal arguments. If predication has a syntactic counterpart, then we ought to expect also the kind of predication brought about by control predicates to have analogous effects. Hence, if we represent predication in the syntax by coindexing, we should also represent the indirect form of predication that arises in control structures by coindexing. An immediate consequence of this, will be the extension of the treatment of opacity developed above to control structures. Essentially, verbal arguments in a control environment are going to be coindexed with their controllers just as matrix verbs are coindexed with their subjects. The mechanism that triggers non coreference in matrix clauses will then automatically trigger the same effects in control structures.

3.2. Syntactic predication and control.

In the present section we are going to present a way of spelling out the above idea. What we want to do is build into our grammar a device that coindexes verbal arguments in control structures with their controllers, just as we now have a device that coindexes predicates and subjects in matrix clauses. In our approach to control the selection of a controller is, in general, carried on by means of θ -roles. So we will somehow have to build a way of keeping track of the relevant θ -roles into the syntax of control predicates. One way of doing this is by assuming that control predicates (i.e. predicates positively specified for CP) carry a feature $\langle \theta, i \rangle$ where θ is the θ -role of the controller among the arguments and i is an arbitrary index. θ is selected in the very same way it is selected in CP, i.e. in terms of a redundancy rule based on the θ -hierarchy. The feature $\langle \theta, i \rangle$ will play a role analogous to that of +SUBJ CON and +OBJ CON in BP's approach. One might argue that a feature of this sort is independently needed for one needs to mark on lexical entries whether they give rise to control structures or not. The feature $\langle \theta, i \rangle$ simply says that its bearer obeys CP with respect to θ . Examples of relevant lexical entries will be:

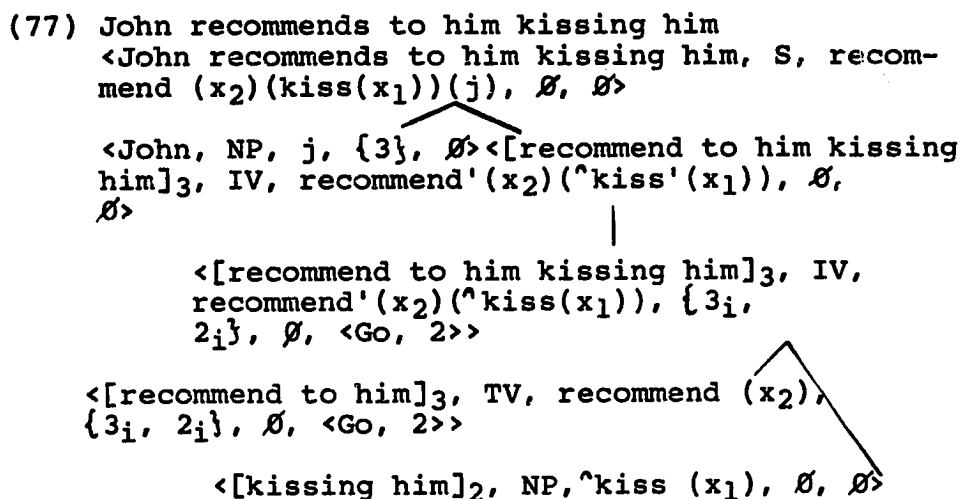
This example illustrates that it is impossible to associate with (74a) a translation where the two occurrences of he are coindexed. Thus the pronoun inside the verbal argument in (74a) will never be caught by a quantifier in subject position, which means that (75b) is not a possible meaning for (75a):

- (75) a. Every man tries to love him
 b. $\forall x[\text{man}'(x) \rightarrow \text{try}'(\text{'love}'(x))(x)]$

The conditions we impose on predication-indices force a subject-control verb like try to have the same indices as its verbal argument and as its subject, which is precisely as it should be. Furthermore, we assume that once predication-indices are checked, they are deleted (i.e. they cancel out). This, however, is not crucial.

Let us consider now cases of object and oblique control:

- (76) John forces him to kiss him.
 <John forces him to kiss him, S, force'
 (^kiss'(x₁))(x₂)(j), ∅, ∅>
- <John, NP, j, {3}, ∅> <[force him to kiss
 him]₃, force' (kiss(x₁))(x₂), {2_i,
 3_i}, ∅, <Th, 2>>
- <he, NP, x₂, {2}, ∅> <[force [to kiss
 him]₂]₃, TV, force (kiss (x₁)), {3_i},
 <Th, 2>>
- <force₃, TV/IV, force', {3_i}, ∅, <Th, 2>>
 <[to kiss him]₂, IV, λx₂[kiss(x₁)(x₂)],
 ∅, ∅>
- <[kiss him]₂, IV, kiss (x₁), {1_i,
 2_i}, ∅>



The same mechanism that takes care of ruling out (75) will also rule out (78) on the reading indicated by the subscripts:

- (78) a. John forces every man_i to kiss him_i
 b. John recommends to every man_i kissing him_i

(78b) is representative of the kind of cases that were problematic for BP's approach.

For sake of explicitness, let us see how the rule for combining verbs with their arguments now looks:

- (79) Let $I(\alpha)$ be the index of α and $I_{\theta}(\alpha)$ be the second member of a feature $\langle \theta, i \rangle$ associated with α .

- a. $X/Y, Y \rightarrow X$
 1 2 3
 b. $3' = 1'(2')$
 c. conditions: if $X = S$: $I(1) = I(2)$,
 $LPS(3) = QST(1) \cup$
 $QST(2)$
 $LPS(1) = \emptyset$
 For no $i, I(2) = j^S$

if $Y = NP$: $LPS(1) \cap LPS(2) = \emptyset$
 $QST(3) = QST(1) \cup$
 $QST(2)$
 $LPS(3) = LPS(1) \cup$
 $LPS(2)$
 if 2 is the Θ of 1,
 then $I_{\Theta}(1) = I(2)$
 if $Y = IV$: $LPS(2) = \emptyset$
 $LPS(3) = LPS(1) \cup$
 $QST(2)$
 $I_{\Theta}(1) = I(2)$

(79) is only a first approximation. It can be improved in various ways and redundancies may be reduced. Also, notice that the conditions in (79c) apply only to the application of a predicate to an argument, not to the application of a functor to an argument (i.e. they are not conditions on rules of the form $X//Y, Y \rightarrow X$). The latter are much simpler (e.g. they do not require emptying of LPS's). However, even in this approximate form, (79) is sufficiently precise to warrant our claims.

3.3. Conclusions.

The main thrust of the present approach to anaphora may be seen in the recognition of a purely syntactic relation of predication, regarded as a special sort of anaphoric relation. This is, of course, one of the central insights of Williams' theory of predication. Predication is represented by coindexing. Our characterization of opacity makes a central use of this syntactic notion of predication together with the semantic notion of predicate domain. In fact, predication-indices represent

a way of marking in phrase structures the semantic notion of predicate domain (i.e. the relation that holds between a propositional function and its internal and external arguments). In one sense, the approach developed in the present chapter could be regarded as an attempt to spell out what it would take to give a truth conditional interpretation to Williams' theory translated into a categorial framework and how such a theory might be related to a semantic notion of predication.

In particular, we have assumed that just like subjects are coindexed with verbs in matrix sentences, verbal arguments are coindexed with their controllers. In this way the definition of opacity given for matrix sentences extends automatically to control structures, without running into the problems that propositional approaches to opacity and control run into. By assuming that verbal arguments are coindexed with their controllers we are able to apply the notion of opacity to nominalized predicates as well as to predicates in their primary role (i.e. to predicates qua propositional functions).

There are obvious differences between the present approach to predication, however, and that of Williams'. In our framework, the syntactic relation of predication is not defined in terms of configurational notions such as C-command. Rather we have a totally independently given semantic notion of predication, viz. the relation that

obtains between a 1-place property and its arguments; this notion is needed to do semantics and may be regarded as the central object of semantic theory. Syntactic predication is the way semantic predication is encoded in syntactic structure. Furthermore, unlike Williams, we do not seek to reduce certain cases of control to syntactic predication. Control is autonomously defined as a set of entailments brought about by certain predicates. These entailments, however, bring into existence predications (in the semantic sense). Thus, if the semantic predications are encoded in the syntax (or give rise to special sort of anaphoric relations encoded in the syntax), those brought about by control predicates will also undergo the same fate, with all the consequences that follow for opacity and binding. So, in our theory the syntactic relation of predication is defined in terms of semantic function argument structure (more specifically, in terms of the notion of predicate domain).

The two approaches, therefore, are quite distinct in conception. These differences give rise to empirical differences. Some of them, those that pertain to opacity, have been discussed in Section 2.1. of the present chapter; others, pertaining to anaphora more in general, have been discussed in Section 1. Finally those that have to do with control have been discussed in Chapter IV. In all these cases we have tried to argue that there are empiri-

cal advantages that descend from our approach. There are theoretical advantages as well, mainly of two sorts. First we do not have to posit a PRO. And why should one want to multiply abstract entities beyond necessity? Second, we can maintain a very tight relation between syntax and semantics; in particular, Ss are always associated with propositions and verbal constituents are always associated with properties. Wouldn't it be great if we could maintain it also against a wider range of constructions?

There are further interesting spin offs that derive from the present approach. In Chapter II Section 3.5. we have seen that our theory of government and percolation cannot handle long distance agreement phenomena. We have tried to argue there that some such phenomena might be purely semantic facts. Recognizing, however, the existence of a syntactic predication relation opens new perspectives in this regard. We can now simply say that an NP agrees with a predicate coindexed with it. There are various ways in which this idea might be formally implemented. The point is quite clear, however: sameness of indices between an NP and a predicate must entail sameness of number, gender and person specification. It might even be that sameness of indices is nothing but sameness of feature specifications. In this way the fact that a predicate agrees with its subject in simple matrix

clauses generalizes automatically to all control cases: controlled items have to agree with their controllers. Nothing special needs to be said about them. This strikes me as a considerable advantage over, for example, current GPSG approaches to agreement in control structures. Klein and Sag (1982) have to stipulate some copying conventions that attach the features of the controller to the controlled verbal argument. In our approach, the device that handles agreement in control structures would be the very same device that handles opacity phenomena.

In fact the potential range of this idea might well go beyond that of (syntactic) predication and extend to all cases of binding. We have seen that any adequate theory of binding seem to require either a weakening of the well formedness constraint or the crucial, ineliminable appeal to the syntax of logical form. The route we are exploring here is the first one. Our grammar still generates directly and recursively well formed surface structures. It still conforms to the rule-by-rule hypothesis. But it does so by going through stages where labelled bracketings are enriched by means of various indexing devices. To constrain these devices is the task of a theory of binding, some principles of which we have considered in the preceding sections. However, even in the form of conventions on stores, what we are doing evidently represents a weakening of the strictest possible

form of the well-formedness constraint. Granted that, we may then be able to exploit our indexing devices to handle all cases of long distance agreement triggered by binding. Consider the following sentence:

(80) Every student_i thinks that Mary loves him_i

To generate (80) on the reading indicated by the subscripts, within our framework we have to assign to the NP every student and him the same index. If sameness of indices entails sameness of feature specification, then the fact that the co-subscripted NPs in (80) have to agree follows. Thus in general we might distinguish between agreement by percolation, whose locality conditions are strictly determined by function argument structure, and anaphoric agreement determined by the theory of binding along the following lines:

(81) Anaphorically related items must agree.

(81) might be implemented as a condition that fills in feature specifications of coindexed items or as a condition that filters out structures with coindexed items with different features. It should be noted that as argued in Chapter II, Section 3.5., there still exist genuine cases of semantic agreement, that would be impossible to treat along the lines of (81). But the burden of handling structures such as (80) would not be on semantic agreement. A further aspect to consider is that while treating (syntactic) predication as governing agreement between

predicates and their subjects appears to be totally straightforward (because predication still is a local relation), this might be not so for quantification and wh-movement, which are potentially unbounded phenomena. However, we will not pursue this issue any further within the limits of the present work.

In conclusion, in the present chapter we have tried to integrate some of the insights developed within the GB tradition (especially aspects of Williams' theory of predication) within a semantically based categorial theory of syntax, that uses crucially and in various ways the notion of function argument structure. Our attempt is in many ways a rough first approximation. But we hope to have shown that there is something to be gained from pursuing the present line of inquiry.

Notes

¹See Carlson (1977) for a totally independent justification and analysis of the generic tense operator. Cf. also Chierchia (1983) for some discussion.

²The issues involved in spelling out more precisely the logic of Gn are quite intricate. See the references quoted in fn 1 for some discussion.

³On the assumption that fun has an extra internal argument, (7) would have to be replaced by:
 (7') $\exists x[\text{fun}'(x)(\text{dance}'(x))]$
 However, for the point I wish to make the presence of an internal argument is immaterial, so I will omit marking it explicitly.

⁴It is only for ease of notation that I am omitting here reference to the contextual restriction over the range of \forall .

⁵Thanks to Irene Heim for bringing these examples to my attention.

⁶Maybe there is such a connection. A very adequate paraphrase of example (18) would be something like:
 (18') running goes on a lot around here.
 So si might correspond semantically to a function from properties to propositions roughly of the following type:
 (18'') si'(VP') = Goes on'(VP')
 (18'') would amount to predicating something of a property. Something like (18'') would certainly not be any worse than appealing to "generic quantification" over a variable.

⁷Exactly the same point has been made by G. Carlson in connection with bare plurals. Also we are limiting ourselves to gerunds because in English they license it-pronominalization more easily infinitives do. Other languages allow for it-pronominalization of infinitives quite freely. Italian is one such language and the translations of (19) and (20) into it would involve infinitival constructions.

⁸Again, in English we have to limit ourselves to gerunds. But see the discussion below.

⁹See fn 5.

¹⁰A proposal along these lines has been developed by D. Lebeaux in a paper delivered at the 1983 GLOW conference. It should also be noted that these examples might not be a problem for Williams' theory, according to which infinitives are supposed to be semantically associated with properties.

¹¹Later on we will also have a principle A_f . The reason we have chosen these labels is to point out more clearly (should that be necessary) the parallelism between the present theory and various versions of theories of binding developed within the GB framework. The superscripted f stands for 'functional'.

¹²We say that a node A C-commands a node B iff the first branching node that (properly) dominates A, (properly) dominates B.

¹³What counts as a governing category varies somewhat from paper to paper within GB. For the point we wish to make it turns out to be quite immaterial which definition to choose. For sake of explicitness let us take the following (from Chomsky, 1982, pp. 184 ff.).

- (i) α is a governing category for β iff α is the minimal NP, S containing β and a governor of β .
- (ii) α governs β iff
 - (a) α is a head
 - (b) no maximal projection intervenes between α and β
 - (c) α C-commands β
- (iii) α is bound iff there is a β such that C-commands α and α and β are coindexed.

¹⁴It might be possible to give a generalized notion of f-command in such a way that all external arguments are f-commanded by internal argument. In this way it might be possible to state the superiority condition on reflexive binding purely in terms of f-command.

¹⁵Reflexive indices are also subscripted as to whether they are external or internal argument (e.g. $k_{\bar{f}}$, $k_{\bar{g}}$). Whenever they turn out to be irrelevant, we will omit mentioning subscripts on reflexives.

¹⁶Conditions on wide scope quantification would require, among other things, keeping in LPS the indices associated with NPs stored in QS. Since, however, these issues are orthogonal to what is our main concern here, we will avoid stating this proviso explicitly and refer to the discussion in BP.

¹⁷There are of course also infinitives marked as +FREE CON to deal with free control. Free control infinitives do not trigger the storage of any binder; their null subject is interpreted according to the context.

REFERENCES

- Abe, Y. (forthcoming) Argument Structure and Derivational Morphology, Ph.D. Dissertation, University of Massachusetts, Amherst.
- Ades, A. and M. Steedman (1982) "On the Order of Words", Linguistics and Philosophy, 4:517-558.
- Akmajian, A. (1977) "The Complement Structure of Perception Verbs in an Autonomous Syntax Framework" in Akmajian et al. (1977).
- Akmajian, A., P. Culicover and T. Wasow (1977) Formal Syntax, New York: Academic Press.
- Bach, E. (1979) "Control in Montague Grammar", Linguistic Inquiry, 10.3:515-531.
- Bach, E. (1980a) "In Defense of Passive", Linguistics and Philosophy, 3:297-341.
- Bach, E. (1980b) "Tenses and Aspects as Functions on Verb-Phrases" in Rohrer, C. (ed.) Time, Tense and Quantifiers, Tübingen: Max Niemeyer.
- Bach, E. (1982) "Purpose Clauses and Control" in Jacobson and Pullum (1982).
- Bach, E. (1983) "On the Relationship between Word-grammar and Phrase-grammar", Natural Language and Linguistic Theory, 1:65-80.
- Bach, E. and B.H. Partee (1980) "Anaphora and Semantic Structure", Chicago Linguistic Society, 10.
- Barwise, J. (1981) "Scenes and Other Situations", The Journal of Philosophy, LXXVIII.7:368-397.
- Barwise, J. and R. Cooper (1981) "Generalized Quantifiers and Natural Languages", Linguistics and Philosophy, 4:159-220.
- Barwise, J. and J. Perry (1981) "The Situation Underground" in French, P., T. Uehling and H. Wettstein (eds.) Midwest Studies in Philosophy VI: The Foundations of Analytic Philosophy, Minneapolis: University of Minnesota Press.
- Bealer, G. (1982) Quality and Concept, Oxford: Clarendon Press.

- Bennett, M. (1976) Some Extensions of a Montague Fragment of English, Bloomington: Indiana University Linguistic Club.
- Bentham, J. van (forthcoming) "Determiners and Logic", Linguistics and Philosophy.
- Bigelow, J. (1978) "Believing in Semantics", Linguistics and Philosophy, 2:101-144.
- Bresnan, J. (1972) Theory of Complementation in English Syntax, Ph.D. Dissertation, M.I.T., published in 1979 by Garland Press.
- Bresnan, J. (1982) "Control and Complementation", Linguistic Inquiry, 13:343-434.
- Carlson, G. (1977) Reference to Kinds in English, Ph.D. Dissertation, University of Massachusetts, Amherst, distributed by Indiana University Linguistic Club.
- Carlson, G. (1982) "Generic Terms and Generic Sentences", Journal of Philosophical Logic, 11:163-182.
- Chierchia, G. (1982) "Nominalization and Montague Grammar: a Semantics without Types for Natural Languages", Linguistics and Philosophy, 5:303-354.
- Chierchia, G. (1983) "On Plural and Mass Nominals and the Structure of the World" in Borowsky, T. and D. Finer (eds.) University of Massachusetts Occasional Papers VIII, Amherst: GLSA.
- Chomsky, N. (1957) Syntactic Structures, Le Hague: Mouton.
- Chomsky, N. (1965) Aspects of a Theory of Syntax, Cambridge, Mass.: M.I.T. Press.
- Chomsky, N. (1980) "On Binding", Linguistic Inquiry, 11:1-46.
- Chomsky, N. (1982) Lectures on Government and Binding, Dordrecht: Foris.
- Cocchiarella, N. (1974) "Fregean Semantics for a Realist Ontology", Notre Dame Journal of Formal Logic, 15:552-568.
- Cocchiarella, N. (1976) "On the Logic of Natural Kinds", Philosophy of Science, 43:202-222.

- Cocchiarella, N. (1978) "On the Logic of Nominalized Predicates and its Philosophical Interpretations", Erkenntniss, 13:339-369.
- Cocchiarella, N. (1979) "The Theory of Homogeneous Simple Types as a Second Order Logic", Notre Dame Journal of Formal Logic, 20:505-524.
- Cocchiarella, N. (1982) "Meinong Reconstructed versus Early Russell Reconstructed: a Critical Commentary on Parsons' Non Existent Objects", Journal of Philosophical Logic, 11:183-214.
- Cochhiarella, N. (1983) Logical Investigations of Predication Theory and the Problem of Universals, Naples: Bibliopolis Press, in press.
- Cooper, R. (1975) Montague's Semantic Theory and Transformational Syntax, Ph.D. Dissertation, University of Massachusetts, Amherst.
- Cooper, R. (1982) "Situation Semantics, Stages and the Present Tense", unpublished manuscript, University of Wisconsin, Madison.
- Cooper, R. (1983) Quantification and Syntactic Theory, Dordrecht: Reidel, in press.
- Cresswell, M. (1973) Logics and Languages, London: Methuen.
- Davidson, D. and G. Harman (1972) Semantics of Natural Languages, Dordrecht: Reidel.
- De Mey, S. (1982) "Aspects of the Interpretation of Bare Plurals", in Linguistics in the Netherlands.
- Dowty, D. (1978) "Governed Transformations as Lexical Rules in a Montague Grammar", Linguistics Inquiry, 9:393-426.
- Dowty, D. (1982a) "Grammatical Relations and Montague Grammar", in Jacobson and Pullum (1982).
- Dowty, D. (1982b) "Tenses, Time Adverbs and Compositional Semantic Theory", Linguistics and Philosophy, 5:23-55.
- Emonols, J. (1976) A Transformational Approach to English Syntax, New York: Academic Press.
- Enc, M. (1981) Tense without Scope, Ph.D. Dissertation, University of Wisconsin, Madison.

- Engdahl, E. (1983) The Syntax and Semantics of Constituent Questions with Special Reference to Swedish, Dordrecht: Reidel, in press.
- Feferman, S. (1982) "Towards some Useful Type-free Theories. Part I", unpublished manuscript, Stanford.
- Flynn, M. (1982) A Categorical Theory of Structure Building, Ph.D. Dissertation, University of Massachusetts, Amherst.
- Gallin, D. (1975) Intensional and Higher Order Logic, Amsterdam: North Holland.
- Gazdar, G. (1980) "A Cross-categorial Semantics for Coordination", Linguistics and Philosophy, 3:407-409.
- Gazdar, G. and J. Pullum (1982) Generalized Phrase Structure Grammar: a Theoretical Synopsis, Bloomington: Indiana University Linguistic Club.
- Gazdar, G. and J. Pullum (1983) "Easy to Solve", Linguistic Analysis, in press.
- Gazdar, G., J. Pullum and I. Sag (1981) "Auxiliaries and Related Phenomena in a Restrictive Theory of Grammar", Language, 58:591-638.
- Gee, J. (1977) "Comments on Akmajian's Paper" in Akmajian et al. (1977).
- Gruber, J. (1965) Studies in Lexical Relations, Ph.D. Dissertation, M.I.T., Cambridge, Mass.
- Halvorsen, P.K. (1982) "An Interpretive Procedure for Lexical-Functional Grammar", unpublished manuscript, M.I.T., Cambridge, Mass.
- Hellan, L. (1980) "Towards an Integrated Theory of Noun Phrases", unpublished manuscript, University of Trondheim.
- Hempel, C. (1966) Philosophy of Natural Science, Prentice Hall.
- Henkin, L. (1950) "Completeness in the Theory of Types", Journal of Symbolic Logic, 15:81-91.
- Horn, G. (1975) "On the Nonsentential Nature of POSS-ING Constructions", Linguistic Analysis, 1:333-388.

- Huntley, M. (1982) "The Semantics of English Imperatives", forthcoming in Linguistics and Philosophy.
- Jackendoff, R. (1972) Semantic Interpretation in Generative Grammar, Cambridge, Mass.: M.I.T. Press.
- Jackendoff, R. (1976) "Toward an Explanatory Semantic Representation", Linguistic Inquiry, 7:89-150.
- Jacobson, P. (1977) The Syntax of Crossing Coreference Sentences, Ph.D. Dissertation, University of California, Berkeley, published by Garland Press.
- Jacobson, P. and J. Pullum (1982) (eds.) The Nature of Syntactic Representation, Dordrecht: Reidel.
- Jensen, R. (1968) "On the Consistency of a Slight (?) Modification of Quine's New Foundations", Synthese, 19:250-263.
- Jespersen, O. (1924) Philosophy of Grammar, London: Allen and Unwin.
- Jubien, M. (1981) "First Order Property Theory", unpublished manuscript, University of Massachusetts, Amherst.
- Kamp, H. (1981) "A Theory of Truth and Semantic Representation" in Groenendijk, J., T. Janssen and M. Stokhof (eds.) Formal Methods in the Study of Language, Amsterdam: Mathematisch Centrum.
- Kayne, R. (1981) "On Certain Differences between French and English" Linguistic Inquiry, 12:349-372.
- Keenan, E. (1974) "The Functional Principle: Generalizing the Notion of 'Subject-of'", Chicago Linguistic Society, 10, Chicago.
- Keenan, E. (1980) "Passive is Phrasal not (Sentential or Lexical)" in Hoekstra, T., H. van der Hulst and M. Moortgat (eds.) Lexical Grammar, Dordrecht: Foris.
- Keenan, E. and L. Faltz (1978) Logical Types for Natural Languages, UCLA Working Papers in Linguistics, 3.
- Klein, E. and I. Sag (1982) "Semantic Type and Control" in Barlow, M., D. Flickinger and I. Sag (eds.) Developments in Generalized Phrase Structure Grammar, Bloomington: Indiana University Linguistic Club.

- Koster, J. (1981) "On Binding and Control", unpublished manuscript, Tilburg University.
- Koster, J. and R. May (1982) "On the Constituency of Infinitives", Language, 58:116-143.
- Kratzer, A. (1981) "The Notional Category of Modality" in Eikmeyer, H. and H. Rieser (eds.) Words, Worlds and Contexts, Berlin: De Gruyter.
- Landman, F. (1982) "Constructive Semantics", unpublished manuscript, University of Amsterdam.
- Landman, F. and I. Moerdijk (1982) "Compositionality and Anaphora", forthcoming in Linguistics and Philosophy.
- Lasnik, H. (1976) "Remarks on Coreference", Linguistic Analysis, 2:1-22.
- Levin, H. (1976) First Order Logic as a Formal Language: an Investigation, Ph.D. Dissertation, M.I.T., Cambridge, Mass.
- Link, G. (1983) "The Logical Analysis of Plural and Mass Terms: a Lattice-theoretic Approach" in Bauerle, R., C. Schwartze and A. von Stechow (eds.) Meaning, Use and Interpretation of Language, Berlin: de Gruyter.
- Manzini, R. (1982) "On Control and Control Theory", forthcoming in Linguistic Inquiry.
- May, R. (1977) The Grammar of Quantification, Ph.D. Dissertation, M.I.T., Cambridge, Mass.
- McCawley, J. (1983) "What's with with", Language, 59:271-287.
- McConnell-Ginet, S. (1982) "Adverbs and Logical Form", Language, 58:144-184.
- Montague, R. (1974) Formal Philosophy, edited by R. Thomason, New Haven: Yale University Press.
- Nishigaucci, T. (forthcoming) "Control and the Thematic Domain", Language.
- Parsons, T. (1979) "The Theory of Types and Ordinary Language" in Davis, S. and M. Mithun (eds.) Linguistics, Philosophy and Montague Grammar, Austin: University of Texas Press.

- Parsons, T. (1981) Non-existent Objects, New Haven: Yale University Press.
- Partee, B.H. (1976) (ed.) Montague Grammar, New York: Academic Press.
- Partee, B.H. (1979a) "Montague Grammar and the Well-formedness Constraint" in Heny, F. and H. Schnelle (eds.) Syntax and Semantics, vol. 10, New York: Academic Press.
- Partee, B.H. (1979b) "Semantics - Mathematics or Psychology?" in Bauerle, R., U. Egli and A. von Stechow (eds.) Semantics from Different Points of View, Berlin: Springer.
- Partee, B.H. (1983) "Compositionality" to appear in the proceedings of the 1982 Amsterdam Colloquium on Formal Linguistics.
- Quine, W.v.O. (1958) "New Foundations for Mathematical Logic" in From a Logical Point of View, Cambridge, Mass.: Harvard University Press.
- Reinhart, T. (1976) The Syntactic Domain of Anaphora, Ph.D. Dissertation, M.I.T., Cambridge, Mass.
- Reinhart, T. (1979) "Syntactic Domains for Semantic Rules" in Guenther, F. and S.J. Schmidt (eds.) Formal Semantics and Pragmatics for Natural Languages, Dordrecht: Reidel.
- Reuland, E. (1983) "Governing ING" in Linguistic Inquiry, 14:101-136.
- Rizzi, L. (1982) "Lexical Subjects in Infinitives: Government, Case and Binding" in Issues in Italian Syntax, Dordrecht: Foris.
- Roeper, T. and T. Wasow (1972) "On the Subject of Gerunds", Foundation of Language, 8:44-61.
- Rooth, M. and Partee, B.H. (1983) "Generalized Conjunction and Type Ambiguity" in Bauerle, R., C. Schwarze and A. von Stechow (eds.) Meaning, Use and Interpretation of Language, Berlin: de Gruyter.
- Sag, I. (1976) Deletion and Logical Form, Ph.D. Dissertation, M.I.T., Cambridge, Mass.

- Sag, I. (1982) "A Semantic Theory of 'NP-movement' Dependencies" in Jacobson and Pullum (1982).
- Szabolcsi, A. (1983) "PRO and Cons -- A Double Dealer Account", Paper presented at the Conference on Sentential Complementation, Brussels.
- Ter Meulen, A. (1980) Substances, Quantities and Individuals, Bloomington: Indiana University Linguistic Club.
- Thomason, R. (1979) On the Semantic Interpretation of the 1972 Fragment, Bloomington: Indiana University Linguistic Club.
- Thomason, R. (1980) "A Model Theory for Propositional Attitudes", Linguistics and Philosophy, 4:47-70.
- Thomason, R. and Stalnaker, R. (1973) "A Semantic Theory of Adverbs", Linguistic Inquiry, 4:195-220.
- Thompson, S. (1973) "On Subjectless Gerunds in English", Foundation of Language, 9:374-383.
- Turner, R. (1982) "Montague Semantics and Scott-Strachey Semantics" in D. Flickinger, M. Macken and N. Wiegand (eds.) Proceedings of WCCFL I, Stanford, California.
- Turner, R. (1983) "Montague Semantics, Nominalization and Scott's Domains", forthcoming in Linguistics and Philosophy.
- Williams, E. (1977) "Discourse and Logical Form", Linguistic Inquiry, 8:101-139.
- Williams, E. (1980) "Predication", Linguistic Inquiry, 11:203-238.
- Williams, E. (1983) "Against Small Clauses", Linguistic Inquiry, 14:101-136.
- Williams, E. (forthcoming) "Semantic vs. Syntactic Categories", Linguistics and Philosophy.