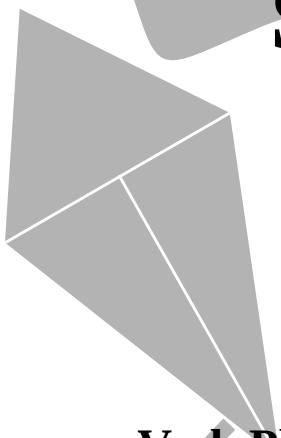
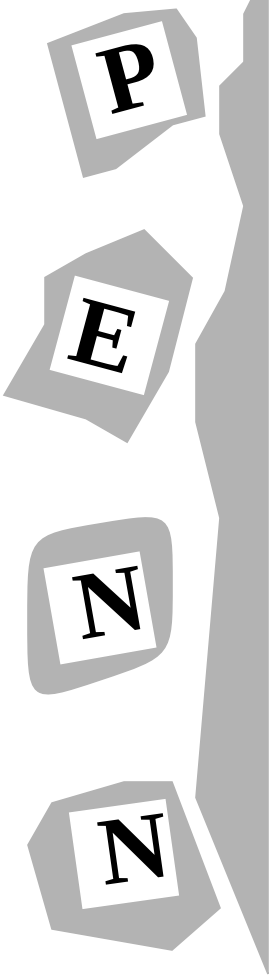
The Institute For Research In Cognitive Science



Verb Phrase Ellipsis: Form, Meaning, and Processing (Ph.D. Dissertation)



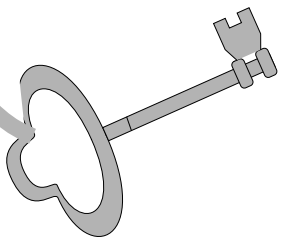
by

Daniel Hardt

University of Pennsylvania
Philadelphia, PA 19104-6228

August 1993

Site of the NSF Science and Technology Center for Research in Cognitive Science



University of Pennsylvania
Founded by Benjamin Franklin in 1740

IRCS Report 93-23

VERB PHRASE ELLIPSIS: FORM, MEANING, AND PROCESSING

Daniel Hardt

A DISSERTATION

in

Computer and Information Science

Presented to the Faculties of the University of Pennsylvania in Partial Fulfillment of the Requirements for the Degree of Doctor of Philosophy

1993

Aravind Joshi and Bonnie Webber

Supervisors of Dissertation

Mark Steedman

Graduate Group Chairperson

Abstract

The central claim of this dissertation is that an elliptical VP is a proform. This claim has two primary consequences: first, the elliptical VP can have no internal syntactic structure. Second, the interpretation of VP ellipsis must be governed by the same general conditions governing other proforms, such as pronouns. The basic condition governing the interpretation of a proform is that it must be semantically identified with its antecedent. A computational model is described in which this identification is mediated by store and retrieve operations defined with respect to a discourse model. Because VP ellipsis is treated on a par with other proforms, the ambiguity arising from "sloppy identity" becomes epiphenomenal, resulting from the fact that the store and retrieve operations are freely ordered.

A primary argument for the proform theory of VP ellipsis concerns syntactic constraints on variables within the antecedent. I examine many different types of variables, including reflexives, reciprocals, negative polarity items, and wh-traces. In all these cases, syntactic constraints are not respected under ellipsis. This indicates that the relation governing VP ellipsis is semantic rather than syntactic. In further support of the proform theory, I show that there is a striking similarity in the antecedence possibilities for VP ellipsis and those for pronouns.

Two computer programs demonstrate the claims of this dissertation. One program implements the semantic copying required to resolve VP ellipsis, demonstrating the correct set of possible readings for the examples of interest. The second program selects the antecedent for a VP ellipsis occurrence. This program has been tested on several hundred examples of VP ellipsis, automatically collected from corpora.

Acknowledgements

The graduate program in Computational Linguistics at the University of Pennsylvania achieves a nearly ideal blend of diversity, intellectual discipline and controlled chaos. I feel very fortunate to have been a part of it. Thanks to everyone, faculty and students, who have made it what it is. In particular, I am indebted to my advisors, Aravind Joshi, who provided invaluable guidance and input through all the stages of my graduate career, and Bonnie Webber, who helped me to develop both an appreciation for the infinite variety in language, and a skeptical attitude about the ability of theory to capture it. My committee members also gave crucial input: Mark Steedman, who provided boundless inspiration and insight about the relation of form and meaning; Ellen Prince, who first showed me how little truth has to do with meaning; and my external member, Shalom Lappin, who gave heroically detailed, constructive criticism and comments on numerous versions of my thesis. Shalom has been consistently supportive of my work, despite the fact that it conflicts in some ways with his own theory.

My fellow graduate students have been a constant source of useful input as well as much-needed diversions: Breck Baldwin, Anuj Dawar, Christy Doran, Barbara Di Eugenio, Bob Frank, Jamie Henderson, Young-Suk Lee, Libby Levison, Mike Moore, Christine Nakatani, Michael Niv, Owen Rambow, Philip Resnik, Matthew Stone, Ramesh Subrahmanyam, Lyn Walker, Mike White, and many others. Other colleagues, both at Penn and elsewhere, have also been great sources of insight and criticism. Thanks in particular to Mary Dalrymple, Michael Hegarty, Arild Hestvik, Sabine Iatridou, Tony Kroch, Mitch Marcus, Dick Oehrle, Mats Rooth, Giorgio Satta, Stuart Shieber, Ivan Sag, and Scott Weinstein.

Although they will surely never read these words, I want to thank my music teachers, Miran Kojian, Ivan Galamian, and especially Jascha Brodsky. Among many other things, they taught me about clarity of purpose. I must also acknowledge the love and unwavering support of my parents

and my brothers. Finally, my wife Birgitte, who understands everything – although she occasionally wonders if there really is any point to studying something like Computational Linguistics...

This work was supported by the following grants: ARO DAAL 03-89-C-0031, DARPA N00014-90-J-1863, NSF IRI 90-16592, and Ben Franklin 91S.3078C-1.

Contents

Abstract **ii**

Acknowledgements **iii**

1 Introduction **1**

1.1 Form and Meaning 1

1.2 The Proverb Theory: An Identity of Meaning 2

1.3 Processing Meanings 4

1.3.1 Storing Meanings 5

1.3.2 Retrieving Meanings 5

1.4 Overview of this Dissertation 7

2 Evidence for the Proverb Theory **8**

2.1 Overview 8

2.2 Sloppy Identity and Lambda Abstraction 9

2.3 Sloppy Readings for Free Pronouns 11

2.4 Wh-traces in the Antecedent 14

2.5 Apparent wh-traces in the Ellipsis Site 17

2.6 Reflexives 19

2.7 Negative Polarity Items 21

2.8 Reciprocals 23

2.9 A Semantic Constraint: Deictics 25

2.10 The Bound-Sloppy Correlation 27

2.11 More Complicated Antecedents 29

2.11.1 Split Antecedents . . . 29
2.11.2 Long Distance Antecedents . . . 32
2.11.3 Missing Antecedents . . . 33
2.12 Conclusions . . . 38

3 Formal System **40**
3.1 Overview . . . 40
3.2 Two Assumption Types: Definite and Indefinite . . . 42
3.3 Rules for VP Ellipsis . . . 44
3.3.1 Syntactic Background . . . 44
3.3.2 Semantic Rules for VP Ellipsis . . . 45
3.4 Sloppy Identity for NP Anaphora . . . 47
3.5 Derivation System . . . 48
3.5.1 A Function Application Rule . . . 48
3.5.2 Assumption Discharge Rules . . . 49
3.5.3 A Discourse Rule . . . 50
3.5.4 Lexical Items . . . 51
3.6 Sample Derivations . . . 52
3.7 Extending the System: Split Antecedents . . . 57
3.8 Conclusions . . . 59

4 Previous Accounts **60**
4.1 Overview . . . 60
4.2 Syntactic Reconstruction . . . 61
4.3 Lambda Abstraction Approaches . . . 63
4.3.1 The Sag/Williams Account . . . 63
4.3.2 The Equational Approach . . . 64
4.3.3 Two Critiques of Lambda Abstraction Approaches . . . 68
4.3.4 The Lambda Abstraction Approach and Semantic Types . . . 68
4.4 Semantic Approaches . . . 69

5 Computational Implementation **71**

5.1 Introduction . . . 71

5.2 The Identity Question: Retrieving VP Meanings . . . 71

5.2.1 Background: The Pereira and Pollack System . . . 72

5.2.2 Extensions for VP Ellipsis . . . 75

5.3 The Location Question: Determining the Antecedent VP . . . 82

5.3.1 Structure of the Algorithm . . . 82

5.3.2 Impossible Antecedents . . . 84

5.3.3 Preference Rules . . . 85

5.3.4 Selecting the Antecedent . . . 87

5.3.5 Extensions to the Algorithm . . . 87

5.3.6 A Partial Implementation . . . 91

6 Empirical Studies **93**

6.1 Locating VP Ellipsis Occurrences . . . 93

6.1.1 Brown Corpus . . . 93

6.1.2 Treebank . . . 95

6.2 Testing the Antecedent Location Algorithm . . . 97

6.2.1 Brown Corpus . . . 97

6.2.2 Treebank . . . 101

7 Related Issues, Open Problems, and Conclusions **106**

7.1 Related Overt Forms: “do it”, “do so” . . . 106

7.2 VP Relatives: Bound VP Variables . . . 108

7.3 VP Ellipsis and Focus . . . 111

7.4 An Open Problem: Parallelism Constraints on Indexing . . . 114

7.4.1 Simple Case: A Single Pronoun . . . 114

7.4.2 Two Pronoun Cases . . . 115

7.5 Related Elliptical Forms . . . 120

7.6 Conclusions . . . 126

Bibliography **127**

Chapter 1

Introduction

1.1 Form and Meaning

Language is a correlation of syntactic forms with meanings. The goal of Computational Linguistics is to show that this correlation is a computable mapping. It is a puzzling fact about language that meanings are sometimes expressed by a missing, or elided, syntactic object. Elliptical expressions clearly pose a major obstacle for Computational Linguistics; indeed, they might lead one to doubt that there is a systematic mapping from form to meaning in human language. In this dissertation I will attempt to show that such pessimism is unwarranted – that elliptical expressions are compatible with a computable mapping between form and meaning. I will suggest, in fact, that elliptical expressions provide a valuable window onto the nature of this mapping.

This dissertation concerns itself with Verb Phrase ellipsis, in which a Verb Phrase (VP) is left unexpressed, although it is clearly understood from context. Intuitively, it appears that the missing VP is identical to an antecedent VP, i.e., an overt VP appearing in surrounding context. Two major issues arise: First, in what sense is the missing VP identical to its antecedent? Is it identical to the *syntactic form* of the antecedent, or the *meaning* of the antecedent? This question, which I will term the *identity question*, has been a major issue in theoretical linguistics. However, there is significant disagreement on this issue, and I will argue that the solutions put forward are fundamentally inadequate. The second question is: how is the antecedent VP determined? This question, which I call the *location question*, has been largely ignored in the linguistics literature, although its solution is clearly required for any computational model.

The central claim of this dissertation is that an elliptical VP is a proform; I will call this the *proverb theory*. The proverb theory has two primary consequences: first, the elliptical VP can have no internal syntactic structure. Second, the interpretation of VP ellipsis should be governed by the same general conditions governing other proforms, such as pronouns. The basic condition governing the interpretation of a proform is that it must be semantically identified with its antecedent.

The exposition of the proverb theory will be couched in terms of computational processes. This requires explicit consideration of issues normally left unexamined in theoretical linguistics. For example, linguists do not normally define a step in a derivation in which the antecedent for an anaphoric expression is determined. Instead, anaphoric expressions are indexed prior to the derivation, allowing linguists to ignore the problems involved in determining anaphor-antecedent relationships.

A processing system does not permit this luxury; some mechanism must be defined to determine anaphoric relations as part of semantic derivations. I will describe a system of semantic derivation which includes such a mechanism, and I will argue that this does not complicate the picture, but in fact results in considerable simplifications in the theory of VP ellipsis, and the theory of anaphora in general. In particular, many linguists have introduced complex mechanisms to solve the puzzle presented by "sloppy identity" in VP ellipsis. I will show that this is unnecessary; rather, sloppy identity arises from the natural interaction of mechanisms that are independently required to account for the resolution of anaphoric expressions.

1.2 The Proverb Theory: An Identity of Meaning

What is the semantic identity that underlies VP ellipsis? A simple characterization would be identity of reference: this would require that the elliptical VP denotes the same property as the antecedent, just as a pronoun denotes the same individual as its antecedent, and a plural pronoun denotes the same set of individuals as its antecedent. Identity of reference correctly captures the vast majority of cases, such as the following examples:

(1) John **walked**. Harry did too.

(2) John helped **Susan**. Harry helped her too.

In (1), the elliptical VP must be interpreted as denoting the same property as its antecedent "walk". Similarly, in (2), the pronoun "her" must denote the same individual as "Susan" denotes.

This simple picture must be modified, however, because of examples in which identity of reference is violated: [1]

(3) If Tom$_i$ was having trouble in school, I would **help him**$_i$.
On the other hand, if Harry$_j$ was having trouble, I doubt that I would. [help him$_j$]

(4) Smith$_i$ makes **his**$_i$ **children** go to bed at 8 every night.
Jones$_j$ lets them [his$_j$ children] stay up as late as they want.

These examples illustrate a phenomenon known as "sloppy identity", which can be defined as cases in which the antecedent contains some variable within it, and that variable receives a different interpretation at the site of the anaphoric expression. In (3), the antecedent for the elliptical VP is "help him", and "him" is interpreted as "Tom" in the antecedent and "Harry" at the ellipsis site. Similarly, in (4), the antecedent for the pronoun "them" is "his children", and the variable "his" is interpreted as "Smith" in the antecedent sentence, but "Jones" in the second sentence, allowing "them" to be interpreted as "Jones' children".

These cases contradict the suggestion that a pronoun denotes the same individual as its antecedent, and an elliptical VP denotes the same property as its antecedent. Of course, in general, the meaning of an NP cannot be treated simply as an individual, but an object that specifies an individual relative to the context. Similarly, the meaning of a VP is a specification of a VP relative to a context. I will term a VP meaning a *dynamic property*, which is defined as a three place relation on an input context, a property, and an output context. Thus the dynamic property denoted by "help him" in (3) is the property of helping some contextually salient male. Similarly, an NP meaning is a *dynamic individual*, which is a relation involving input context, an individual, and an output context.[2] When we recognize that NP and VP meanings are themselves context-dependent in this way, we can see that these examples are quite consistent with the view that proforms have the same *meaning* as their antecedents.

[1] The antecedent for a proform is displayed in **bold**; for convenience, I will often indicate the intended interpretation for a proform in brackets.

[2] The "dynamic" view of meaning is developed in theories such as [Kamp, 1980; Heim, 1982; Barwise, 1987; Rooth, 1987; Groenendijk and Stokhof, 1992].

One might instead argue that the identity here is not one of meaning but rather of uninterpreted syntactic form. For example (1), the elliptical VP can be thought of as identical to the syntactic object "walk", which is then interpreted again as denoting the property of walking. Or in example (3), one could simply copy the syntactic object "help him", allowing it to be reinterpreted in the elliptical VP's context. Although many recent accounts in the linguistics literature have taken this approach, I believe it is untenable. A major problem with this approach can be seen from processing considerations, since, while there appears to be no bound on the amount of material separating the elliptical VP from its antecedent, it is generally believed that memory for syntactic structure is quite short-lived. Similarly, from a theoretical point of view, a syntactic identity theory of VP ellipsis would require a radical extension of the domain of syntactic constraints, which are traditionally restricted to a single sentence. Furthermore, there are many cases where it is simply not possible to syntactically copy the antecedent to the ellipsis site. For example:

(5) This is just the kind of thing that Harris could have **suggested [e]**. And in fact, he did. (Comment during talk at University of Pennsylvania)

Here the antecedent includes a syntactic "trace", but a trace is not possible at the ellipsis site. In fact there are a wide variety of cases in which the antecedent contains a variable under syntactic constraints, where those constraints are not imposed at the ellipsis site. This demonstrates that the ellipsis site does not contain a syntactic copy of the antecedent. Note that this example is quite compatible with semantic identity; the meaning of the elliptical VP is the dynamic property denoted by the antecedent VP "suggested [e]", where the free variable e is fixed by context.

Finally, there are many cases of VP ellipsis where there is no VP antecedent. I will describe a range of antecedence possibilities for VP ellipsis, including split antecedents and several types of missing antecedents. This range of antecedence possibilities is clearly inconsistent with a syntactic identity theory. I will argue that it parallels the range of antecedence possibilities observed for pronouns; this provides strong evidence for the proverb theory.

1.3 Processing Meanings

There are two operations which underly the interpretation of proforms. The first is that of *storing meanings*, as potential antecedents, and the second is *retrieving meanings* – selecting among

the stored antecedents. These operations will be defined to apply to both NP meanings and VP meanings, thus permitting a uniform treatment of pronominal anaphora and VP ellipsis. Each of these operations will be defined as an explicit step in the semantic derivation.

1.3.1 Storing Meanings

Meanings are stored in a semantic data structure termed the *discourse model*. When a sentence is processed, the proposition denoted by that sentence is stored in the discourse model, and, in addition, various subparts of the proposition can be stored in the discourse model. In particular, the system will store NP meanings and VP meanings in the discourse model; that is, dynamic individuals and dynamic properties. These objects may be stored with some contextual dependencies as yet unresolved. As discussed below, this is crucial to the treatment of sloppy identity.

For individuals, this is implemented as follows: following [Heim, 1982], I will assume that indefinites add individuals to the discourse model. I will therefore associate an "indefinite assumption" with indefinite NP's which must be "discharged" at some point in the course of a derivation, adding a new individual to the discourse model. Storing the individual associated with the indefinite determiner reflects the fact that this object is available for subsequent intersentential reference.

Expanding on Heim's insight, I will treat VP's as "indefinite properties". This has been suggested by many theorists, most prominently in the work of [Davidson, 1980]. In the context of the current system, this means that the occurrence of a verb phrase introduces a dynamic property which is available for subsequent intersentential reference. Thus, I will associate an "indefinite" assumption with VP's which, when discharged, adds the VP meaning to the discourse model.

1.3.2 Retrieving Meanings

The interpretation of a pronoun or definite description requires the selection of a dynamic individual from the discourse model, and the interpretation of an elliptical VP requires the selection of a dynamic property. In theoretical linguistics, the selection of an antecedent is generally denoted by an index fixed in advance of the derivation. In such approaches, there is no convenient way to consider alternative indexing possibilities for a given pronoun; one must simply consider different derivations. This is the case, for example, in many versions of Government-Binding theory

[Chomsky, 1981]. Various filters are then defined to rule out structures, based on the indexing of pronouns. To build in preferences involving such factors as recency and salience, one might imagine that, in addition to filtering invalid derivations, valid derivations are ranked according to these preference factors. As a processing model, this is unattractive; it is neither efficient nor psychologically plausible to require the computation of several alternative derivations for a given sentence, which are later filtered and ranked.

In contrast to this, I will develop an approach in which the pronoun is initially not indexed in the semantic derivation; at some stage in the derivation the antecedent for each variable must be determined. As a processing model, this is clearly superior to the standard approaches in which pronouns are indexed in advance. This approach also has an important consequence: a proform is represented both before and after its referent is determined. This gives rise to an ambiguity with respect to the rules for storing meanings. Consider a VP containing a pronoun, such as "help him" in example (3). There are two assumptions associated with the VP: one requiring that the VP meaning be added to the discourse model, and another requiring that the pronoun be evaluated with respect to the discourse model. These assumptions may be discharged in either order. In other words, the referent for the pronoun may be fixed when the VP meaning is stored, or left undetermined. This ambiguity arises quite naturally in a system of processing meanings of the sort described here, and I will argue that it correctly characterizes the range of possible sloppy readings for VP ellipsis.

It is generally believed that sloppy identity in VP ellipsis is possible only for variables that are bound to the subject. No such constraint is naturally statable in the approach I am pursuing, and, based on examples such as (3), I will dispense with this constraint. Instead, I will permit any variable within the antecedent VP to be reinterpreted at the ellipsis site. In addition, sloppy identity for VP ellipsis has generally been explained by reference to the presence of the lambda operator; in this way, these accounts have relied on the semantic type of the VP to explain sloppy identity. This is not the case in my approach; both for NP anaphora and VP ellipsis, any antecedent containing a variable within it can give rise to sloppy identity.

1.4 Overview of this Dissertation

There are many evident similarities between VP ellipsis and pronominal anaphora: in both cases, the antecedent can be across utterance boundaries and speaker boundaries, and both appear to obey fundamental constraints on anaphora such as the "Backwards Anaphora Constraint"[3]. The analogy can be extended by treating a VP as an indefinite property expression. Thus, just as indefinite individual expressions are available for intersentential reference, VP's, which are indefinite property expressions, are also available for intersentential reference.

To explain VP ellipsis in terms of syntactic identity, it would be necessary to abandon the traditional restriction of syntactic rules to a single sentence. I will argue that this is no more attractive than the notion that a pronoun is syntactically identified with its antecedent. I will argue that VP ellipsis does not in fact involve ellipsis at all – it is simply a proform that is semantically interpreted as other proforms are.

This dissertation is organized as follows: In Chapter Two, I give evidence for the proverb theory: I show that VP ellipsis cannot be explained as syntactic identity, and that the antecedence possibilities for VP ellipsis are parallel to those of pronouns. In Chapter Three, I give a formal characterization of the semantic identity underlying VP ellipsis. I describe a system of semantic derivation which extends Discourse Representation Theory in two ways: first, anaphoric terms are not preindexed. Second, VP's are treated as indefinite properties, and VP ellipsis is treated as a proverb; this makes it possible to assimilate VP ellipsis to the DRT treatment of intersentential pronominal anaphora. In Chapter Four, I discuss other approaches to VP ellipsis in the literature. Chapter Five concerns the computational implementation of the system described here. In Chapter Six I describe empirical studies based on data collected from corpora. In Chapter Seven, I discuss related issues and open problems, and draw conclusions.

[3] This constraint, described in [Sag, 1976], rules out cases in which an anaphor precedes and commands its antecedent.

Chapter 2

Evidence for the Proverb Theory

2.1 Overview

In this chapter, I present two forms of evidence for the proverb theory of VP ellipsis. First, I give evidence that VP ellipsis is governed by semantic identity with the antecedent. Second, I show that the range of antecedence possibilities parallels that of other proforms.

I will begin with the issue of "sloppy identity": cases in which a pronoun (or other variable) within the antecedent VP is interpreted differently at the ellipsis site. The Sag/Williams account [Sag, 1976; Williams, 1977] is based on lambda-abstraction. This account allows a variable to "switch" from antecedent to ellipsis only if it is λ-bound; in other words, only a subject-coreferent variable can have a sloppy reading. I will argue that this must be generalized to allow a sloppy reading for any variable in the antecedent. I present evidence for this with cases involving sloppy readings for pronouns in a variety of configurations. This generalization follows from the notion that VP ellipsis is an identity of "dynamic properties"; since a variable within a dynamic property is dependent on the input context, the interpretation of that variable can change as the input context changes.

Next, I will turn to syntactic constraints on variables within the antecedent. I will show that, when a variable appears in the antecedent under a particular syntactic constraint, this constraint is not imposed under ellipsis. For example, when a trace occurs in the antecedent VP, there is a syntactic requirement that the trace be bound by a wh-operator. Under ellipsis, this syntactic requirement is not imposed, and the variable corresponding to the trace need not be bound by

a wh-operator. A similar phenomenon is found with negative polarity items, where there is a requirement for a negation operator in a particular syntactic configuration. This requirement is not imposed under ellipsis. Similar argumentation is applied to examples involving reflexives and reciprocals.

Next I examine cases where a variable appears under a semantic constraint. It has been argued by Kaplan and others that so-called *indexical* or *deictic* expressions are not permitted to vary across contexts. I show that this semantic constraint holds under ellipsis.

The Sag/Williams account has resulting in an influential generalization, namely, that the strict/sloppy dichotomy corresponds to the dichotomy between two types of variables: bound and free variables. The standard ways of defining this correspondence cannot be maintained in the face of the data given in this chapter, in particular, one cannot restrict sloppy readings to a configurationally-defined notion of bound variable. However, I will suggest that the semantic distinction between deictic and non-deictic variables is properly correlated with the strict/sloppy dichotomy. This characterization suggests a further prediction: that an obligatory strict reading correlates with the lack of interaction with modal operators.

Finally, I examine some more complicated types of antecedents for VP ellipsis. The range of possibilities, I argue, parallels that of other proforms, such as pronouns. Typically, the antecedent for a pronoun is an NP antecedent. However, this is not required; "split antecedents" and "missing antecedents" are also possible. The same is true of VP ellipsis: while the antecedent is typically a VP, I show that there are also cases of split antecedents and missing antecedents. This parallelism in antecedent possibilities between VP ellipsis and other proforms provides striking support for the proverb theory.

2.2 Sloppy Identity and Lambda Abstraction

Sloppy identity can be defined as cases in which a pronoun or other variable within the antecedent VP is allowed to change its referent in the ellipsis site. Examples of the following type are discussed as early as [Ross, 1967]:

(1) John **thinks he is smart**. Harry does too.

There is a reading in which "John thinks John is smart" and "Harry thinks Harry is smart". This presents a dilemma: it is generally assumed that the elided VP should be identical in some sense to the antecedent VP. But here, the antecedent VP contains a pronoun "he" which refers to John in the antecedent, but Harry in the ellipsis site. It seems that either we are forced to give up the idea of an identity condition, or we need some looser, or "sloppier" identity relation.

The Sag/Williams account provides an elegant solution to this problem by representing the antecedent VP as the following lambda expression:

λ x.x thinks x is smart

This representation can be applied to both John and Harry, allowing the sloppy reading while retaining an identity condition on VP ellipsis. This approach relies on two rules: first, the Derived VP rule, which allows VP's to be represented as lambda abstracts. Second, the Pronoun Rule, allowing a subject-coreferent pronoun to be replaced with a lambda-bound variables. These rules are defined as follows:

Derived VP Rule: VP $\Rightarrow$ λ x.x VP

Pronoun Rule: λ x.x . . . he . . . $\Rightarrow$ λ x.x . . . x . . .

With these rules, together with a **VP Copy** rule which copies a VP to an ellipsis site, the derivation of (1) proceeds as follows:

John [thinks he is smart]$_{VP}$. Harry does too.

DVP

John, [λ x.x thinks he is smart]$_{VP}$. Harry does too.

Pronoun Rule

John, [λ x.x thinks x is smart]$_{VP}$. Harry does too.

VP Copy

John, [λ x.x thinks x is smart]$_{VP}$. Harry does [λ x.x thinks x is smart]$_{VP}$ too.

A basic constraint in the Sag/Williams account is that only lambda-bound variables can give rise to sloppy readings; variables that are bound outside the VP, or "discourse bound" variables

cannot give rise to sloppy readings. Together, the operation of DVP rule and the Pronoun Rule result in the following schema for sloppy identity:

Schema 1 λ x[. . . x . . .]$_{VP}$ . . . λ y [. . . y . . .]$_{VPE}$

A sloppy reading results when a variable x that is lambda-bound in the antecedent is "captured" by the lambda operator at ellipsis site. Thus, since a lambda-bound variable in the antecedent must corefer with the subject, this approach predicts that only variables that corefer with the subject can have sloppy readings.

In the next section, I show that this prediction is false, motivating a generalization of Schema 1.

2.3 Sloppy Readings for Free Pronouns

In this section, we see that sloppy readings are possible for free variables, i.e., variables in the antecedent VP that do not corefer with the subject.[1] [Sag, 1976] points to the following contrast:

(2) John said Mary **hit him**, and Bill did, too.

(3) John said Mary **hit him**, and Bill said she did, too.

Sag argues that, while example (2) has both a strict and sloppy reading, example (3) permits only the strict reading. That is, (2) has a reading in which "Bill said Mary hit Bill", whereas (3) lacks this reading.

This is predicted by the Sag/Williams account, since sloppy readings are only possible for variables that corefer with the subject of the elided VP[2]. In fact, subject coreference cannot be the basis for this contrast, since it would also rule out a sloppy reading in the following example:

(4) A: Who was hit by Mary?
B: John said Mary **hit him**, Bill said she did, and Harry said she did.

Here, the sloppy reading is available – in fact it seems to be preferred. The following example pragmatically requires the sloppy reading for the free pronoun.

[1] Of course, there can also be variables that are bound within the antecedent VP that do not corefer with the subject, if they are bound by some operator other than the lambda operator of the VP. I will not consider such cases.

[2] The same prediction is made by the equational approach of [Dalrymple *et al.*, 1991], as pointed out in [Dalrymple, 1992].

(5) John$_i$ admitted that Mary had **bribed him**$_i$.

(6) Bill$_j$ admitted that she had too. [bribed him$_j$]

Similarly the Sag/Williams account would rule out the following discourse:

(7) Did anyone$_i$ admit that Mary had **bribed him**$_i$?

(8) JOHN$_j$ admitted that she had.

According to the Sag/Williams account, no reading should be possible here, since the pronoun "him" would have to be bound by "anyone" in the ellipsis site, although it is outside of its scope.

It may be felt that these examples have a slightly artificial quality. This can perhaps be ascribed to the availability of a more concise form, in which the matrix VP is elided. For example, instead of (8) a more natural response to (7) is simply "John did."

In the following examples, the matrix VP cannot be elided, because contrastive stress is required within the matrix VP[3]:

(9) a. John admitted that Mary had **bribed him**.

b. Bill didn't ADMIT that she had. He implied it though.

(10) a. John admitted that Mary had **bribed him**.

b. Bill didn't admit that MARY had. But he admitted that SOMEBODY had.

In these examples, only the sloppy reading is possible, and (at least to my ear) the artificiality is removed.

A naturally occurring example of this is the following:

(11) If women are often frustrated because men **do not respond to their troubles by offering matching troubles,** men are often frustrated because women do. (*That's Not What I Meant*, Tannen (1991))

Here the pronoun "their" does not corefer with the subject of the antecedent VP, but it gets a sloppy reading at the ellipsis site, switching from "women" to "men".

I have shown, then, that a sloppy reading is possible for a pronoun free within the antecedent VP when it is bound by some other lambda operator, i.e., when the pronoun corefers with the subject of a "higher" VP [4]. I will consider now two other possibilities for variables free within the

[3]This point was suggested to me by Bonnie Webber.

[4]This possibility is also observed in [Kitagawa, 1991].

antecedent VP. First, a variable may be bound by a quantifier outside the VP (QBound). Second, it may be "discourse bound" (Dbound); that is, it may refer to some salient object in the current discourse context. In all these cases, I will show that the variable can be reinterpreted at the ellipsis site.

Consider again example (3) from Chapter One, repeated below:

(12) If Tom$_i$ was having trouble in school, I would **help him**$_i$.
On the other hand, if Harry$_j$ was having trouble, I doubt that I would. [help him$_j$]
DBound-DBound

Here the variable is Dbound in both the antecedent and ellipsis site; the antecedent VP is "help him"; "him" refers to "Tom" in the antecedent but "Harry" at the ellipsis site. This sloppy reading cannot be derived by the Sag/Williams account, since the antecedent is represented as λ x.help(x,him$_i$). The pronoun "him" cannot be represented as a lambda-bound variable, and therefore cannot give rise to a sloppy reading on that approach.

(13) Every boy$_i$ in Bill's class hoped Mary would **ask him**$_i$ **out**, but every boy$_j$ in John's class hoped that she wouldn't. [ask him$_j$ out]
QBound-QBound

In this case, the sloppy variable is bound by two distinct quantifiers.[5] In the antecedent, the pronoun "him" is bound by "Every boy in Bill's class". However, at the ellipsis site, it becomes bound by another quantifier – "every boy in John's class".

In the following example, the pronoun is QBound outside the antecedent, but it becomes DBound in the ellipsis site.

(14) Every boy$_i$ hopes Professor Davidson will **like his**$_i$ **work**, but in Bill$_j$'s case, I think she actually will. [like his$_j$ work]
QBound-DBound

The converse is also possible; a pronoun that is DBound in the antecedent becomes QBound at the ellipsis site.

[5]This example is a variation on examples presented in[Lappin, 1984].

(15) A: Speaking of Mary$_i$, John **went out with her**$_i$.

B: Really – I'm surprised that any girl$_j$ would want him to. [go out with her$_j$]

DBound-QBound

All of these examples involve sloppy readings for free variables in the antecedent VP, that is variables that are not lambda-bound.[6] This contradicts the prediction of the Sag/Williams account expressed by Schema 1, which I repeat here:

Schema 1 λ x[. . . x . . .]$_{VP}$ . . . λ y [. . . y . . .]$_{VPE}$

I suggest, therefore, the following generalization to Schema 2:

Schema 2 [. . . x . . .]$_{XP}$ . . . [. . . y . . .]$_{XP'}$

Here XP is the antecedent for XP'. This removes two constraints from Schema 1: first, it no longer confines sloppy identity to VP anaphora – an anaphoric expression of any category permits sloppy identity. Second, any variable x embedded within an antecedent can now receive a sloppy reading – not only lambda-bound variables.[7] In the above examples, we have seen that x and y can be bound by quantifiers or "discourse-binders". In the next section, we examine cases involving "wh-operators", i.e., operators that bind variables in relative clauses.

2.4 Wh-traces in the Antecedent

Next, I look at cases in which there is a variable in the antecedent that is bound by a "wh-operator". I will give examples that exhibit the same phenomenon discussed above, namely, a variable bound by one operator in the antecedent becomes bound by another operator (or "discourse bound") at the ellipsis site. Cases involving a wh-operator are particularly interesting with respect to the question of syntactic vs. semantic identity, since wh-variables are governed by well-known syntactic constraints. I will show that these syntactic constraints are not respected under ellipsis. Consider the following example:

[6] Examples violating Schema 1 due to [Wescoat, 1989] are discussed in[Dalrymple *et al.*, 1991], such as: "The policeman who arrested John failed to read him his rights, and so did the one who arrested Bill."

[7] This generalization is consistent with the results of a psycholinguistic study described by[Hirschberg and Ward, 1991], where it is argued that the structural constraints of the sort described in Schema 1 are merely tendencies that can be overridden by intonational and contextual factors.

(16) China is a country that Joe wants to **visit e**, and he will too, if he gets enough money. [Webber, 1978]

The antecedent VP is "visit e", where e denotes a wh-trace. The ellipsis site is not within the relative clause. Thus the above example would be ungrammatical if a syntactic representation of the antecedent were copied to the ellipsis site, since the trace would not be bound. Semantically, the antecedent simply denotes a property with a free variable. As usual, then, I treat the antecedent simply as a dynamic property, which will denote a property based on the current input context. That is, any free variables will be determined by the current context. In the above example, the variable is syntactically constrained to denote "China" in the antecedent. It is not so constrained in the ellipsis site; it continues to denote "China" simply because no alternative has been added to the context. This suggests that, if the context changes substantially between antecedent and ellipsis site, a trace might switch its referent at the ellipsis site. The following examples shows that this is indeed the case.

(17) China is one of many countries that Joe doesn't **want to visit e**. In the case of India, he does.

(18) There are many Asian countries that Joe doesn't want to **give money to e**.
In the case of India, I KNOW he won't.

(19) There are many Asian countries that Joe doesn't want to **give money to e**.
Some countries already know that he won't.

These examples are difficult to explain on a syntactic copying approach. One might perhaps argue that it is not the entire VP that is copied, but merely the verb. For example, in (16), "visit" is copied, resulting in "he will visit, too". That is, the verb becomes "detransitivized" at the ellipsis site. This analysis is violated by the following examples, where there is additional material within the VP.

(20) Mary is the girl Tom wanted to **invite e to the party**. But he didn't, because he's so shy.

(21) Harry is someone they would like to **send e to the Olympics**. And they will too, if they can finance it.

Furthermore, there are some verbs that cannot be detransitivized, as in the following examples.

(22) This is just the kind of thing that Harris could have **suggested e**. And, in fact, he did. (comment at Sloan Talk, UPenn)

(23) We thought that was what do-support came in to **save e**. It doesn't. (comment at UPenn Talk, 12/3/92)

(24) It was something, Warren said, choosing his words carefully, he wouldn't **do e**. Yes he could, Miriam said. (J.C.Oates, *You Must Remember This*, p391)

In none of these examples could the verb be copied. For example, in (24), simply copying the verb "do" results in the ungrammatical "Yes he could do".

The following are additional naturally-occurring examples:

(25) A: What interesting lives you all **lead e**!

(26) B: No we don't. (from the film: *Howard's End*)

(27) Republicans claimed that a Presidential veto was what the Democrats **wanted e**.

(28) And some Democrats did. (NPR 6/18/92)

(29) If we had another witness of sterling character to **bring forward**, we would. ("Reasonable Doubts", 12/15/92)

In these examples, we have seen that a wh-trace in the antecedent is no longer bound by a wh-operator in the ellipsis site. If VP ellipsis involved syntactic copying, this would violate the requirement that a wh-trace be bound by a wh-operator. These examples are therefore strong evidence against a syntactic copy approach.

To summarize the argument up to this point: the Sag/Williams account predicts that a variable bound by an operator outside the antecedent VP must remain bound by the same operator at the ellipsis site. I have argued that such a variable can in fact become bound by a different operator at the ellipsis site. First, I showed this with examples involving lambda binding, quantifier binding and discourse binding. In this section, I presented similar examples with wh-traces, ie., variables bound by a wh-operator in the antecedent. Although a wh-trace is syntactically required to be bound by a wh-operator, we have seen that it need not be bound by a wh-operator at the ellipsis site.

In the next section I will examine another configuration involving wh-traces; one which has been taken to indicate that VP ellipsis does involve syntactic reconstruction. I will show that such a conclusion is, in fact, unwarranted.

2.5 Apparent wh-traces in the Ellipsis Site

In the previous section, I considered cases where the antecedent takes this form:

$$\text{wh}_x \ldots [\ldots \text{x} \ldots]_{VP}$$

I will now consider the converse of this configuration, where the ellipsis site appears to have this form:

$$\text{wh}_y \ldots [\ldots \text{y} \ldots]_{VPE}$$

It has been argued [Chao, 1987; Tancredi, 1992] that this configuration is evidence for syntactic reconstruction, since there appears to be a wh-trace, y, within the ellipsis site. I will show that this argument is based on a faulty analysis of the relevant examples.

Consider the following examples:

(30) John knows who Bill **criticized**, and Mary knows who Sue did.

(31) Almost every girl said that Bill **criticized her**. Who didn't he?

These examples, on the above analysis, seem to show that the elided VP can contain a wh-trace and therefore must have internal syntactic structure. However, this is not consistent with the proverb theory, which requires that the elided VP contain no internal syntactic structure. [Chao, 1987] and others have argued that, in examples of this type, VP ellipsis *must* be syntactically reconstructed to allow wh-binding within the ellipsis site. In addition, it has been pointed out [Haik, 1987] that such wh-traces are governed by syntactic subjacency conditions, based on examples such as the following:

(32) * John met everyone that Peter wondered when he could.

(33) * John read everything which Bill believes the claim that he did.

Both examples are (correctly) ruled out by subjacency, if one assumes the existence of a wh-trace, as follows:

(34) * John met everyone Op_i that Peter wondered when he could . . . e_i.

(35) * John read everything $which_i$ Bill believes the claim that he did . . . e_i.

However, on closer examination, it will emerge that these are not in fact cases of VP ellipsis at all. [Lappin and McCord, 1990] have shown that such examples can be analyzed as "pseudo-gapping". Pseudo-gapping involves an elided verb, but a non-elided argument to the verb. For example:

(36) Tom writes books, and Harry does magazines.

(37) Tom kissed Mary, and Harry did Susan.

At first glance, it would appear that, when the argument to the verb is a trace, the difference between VP ellipsis and pseudo-gapping is unobservable. However, pseudo-gapping has a much more constrained distribution than VP ellipsis, appearing most naturally in constructions such as comparative clauses, subordinate clauses, or directly adjacent clauses. (This is pointed out in [Levin, 1985].) I will show that the examples of the sort mentioned above pattern with pseudo-gapping rather than VP ellipsis.

While "backward" VP ellipsis is quite acceptable, backward pseudo-gapping is difficult if not impossible, as shown by the following contrast between pseudo-gapping and VP ellipsis [8]:

(38) *Although I don't know if Tom does books, I know Harry writes magazines.

(39) Although I don't know if Tom does, I know Harry writes magazines.

(40) *Although I don't know if Tom did Mary, I know Harry kissed Susan.

(41) Although I don't know if Tom did, I know Harry kissed Susan.

The cases with wh-traces pattern with pseudo-gapping rather than VP ellipsis:

(42) *Although I don't know what Tom does, I know Harry writes magazines.

(43) *Although I don't know who Tom did, I know Harry kissed Susan.

Another example is the following:

(44) *Although it doesn't me, it takes Karen a long time to clean the hamster's cage.

[8] This is observed in [Sag, 1976, page 83].

(45) Although it doesn't always, it sometimes takes a long time to clean the hamster's cage. [Levin, 1985, page 53]

Again the wh-trace example patterns with pseudo-gapping:

(46) *Although I don't know who it doesn't, it takes Karen a long time to clean the hamster's cage.

Finally, Levin observes that pseudo-gapping cannot appear in infinitive clauses, although VP ellipsis can.

(47) Van Gogh's work is beginning to impress me.

(48) *It's starting to me, too.

(49) Well, it's finally starting to. [Levin, 1985, page 54]

The wh-trace example again patterns with pseudo-gapping:

(50) Van Gogh's work is starting to impress many people.

(51) * Who is it starting to?

To sum up: many authors have argued that the occurrence of VP ellipsis with a wh-trace in the ellipsis site is evidence for syntactic reconstruction in VP ellipsis. This argument loses its force when one acknowledges that the pseudo-gapping analysis of these sentences seems equally possible. On closer examination, it emerges that the pseudo-gapping analysis of these sentences is actually superior to the VP ellipsis analysis. Since the proverb theory *requires* this analysis, these examples are in fact strong evidence in favor of the proverb theory.

2.6 Reflexives

It has been long been recognized ([Sag, 1976; Dalrymple, 1991]) that syntactic constraints on reflexive binding can be violated under ellipsis, based on examples such as the following:

(52) John **defended himself** because his lawyer couldn't.

In the antecedent, the reflexive "himself" introduces a variable under a well- known syntactic constraint, namely that it corefer with a "local" NP. If the elliptical VP is syntactically reconstructed, this syntactic requirement is violated under the reading where "himself" in the ellipsis site refers to "John". This reading is not available in the non-elliptical counterpart of this example:

(53) John defended himself because his lawyer couldn't defend himself.

A syntactic copy theory would predict reflexives always receive "sloppy" interpretations. The following are two naturally-occurring examples where a reflexive gets a strict reading under ellipsis:

(54) I wouldn't **be able to respect myself afterwards**. Nor would you, Aurora. (From the film "Impromptu").

(55) I would **hurt myself**, before he could. (*The Secret Diary of Laura Palmer*,p. 48)

With reflexives, just as in the case of wh-traces, we see that syntactic constraints on the interpretation of a variable are not respected under ellipsis. However, in many cases, there does seem to be a tendency to favor a sloppy reading. For many speakers, the following example exhibits a strong preference for a sloppy reading:

(56) John **likes himself**. Bill does too.

Based on the contrast between examples like (53) and (56), [Hestvik, 1992] suggests that, in coordinated sentences, the reflexive shows a strong tendency for a sloppy reading, and that a strict reading must be "forced" pragmatically or otherwise. However, if the VP ellipsis clause is *subordinated* to the antecedent clause, Hestvik argues, the strict reading is generally available. However, even in simple coordinated cases like the previous example, the strict reading still seems to be available. And, as Hestvik acknowledges, there are many other ways in which the strict reading can be made more salient in a coordinated structure, such as the following:

(57) Sam **loves himself**, but nobody else does.

Another example is the following:

(58) If John didn't **defend himself**, then who did?

Again, the strict reading is available. In view of these examples, it is hard to see how the possible readings can be correlated with the syntactic distinction of subordination vs. coordination. In the account of [Sag, 1976], reflexives are treated the same as other variables, such as pronouns, under ellipsis. However, although there may be no categorical syntactic constraint distinguishing reflexives from pronouns under ellipsis, one might still suspect that there is a stronger tendency for reflexives to receive a sloppy reading. Consider the following contrast:

(59) John **likes himself**. Bill does too.

(60) John **likes his cat**. Bill does too.

It appears that the tendency for a sloppy reading is stronger in (59) than it is in (60). However, the difference might be attributed to another factor: Consider (60) in a "null context". It is necessary to "accommodate" the fact that John has a cat, to interpret the left conjunct. Then, the sloppy reading would require an additional accommodation of a cat for Bill, whereas the strict reading requires no further accommodation. If one assumes there is a cost associated with accommodation, this would make the sloppy reading more costly in a null context. What about in a context in which it is known that John and Bill have cats? Here, it seems there would be a marked preference for the sloppy reading. Thus, once we control for "accommodation", the uniform treatment of reflexives and pronouns under ellipsis appears quite reasonable.

2.7 Negative Polarity Items

We have seen that variables associated with wh-traces and reflexives are not governed by syntactic constraints in the ellipsis site. Negative polarity items (NPI's) provide another illustration of this phenomenon. It is generally recognized that an NPI such as "anyone" is subject to syntactic configurational constraints governing its distribution [Ladusaw, 1979; Linebarger, 1987]. For example:

(61) Tom doesn't have any paper.

In this example, the negation operator "licenses" the appearance of the NPI "any" The meaning of example (61) can be represented:

(62) NOT $\exists$ x:paper(x).have(Tom,x)

Thus, NPI's are analyzed as indefinites, together with a syntactic configurational constraint governing their distribution. As pointed out in [Sag, 1976], under ellipsis this constraint is violated. For example:

(63) Tom doesn't **have any paper**. Harry does, though.

The meaning of the elliptical sentence is "Harry does have some paper." The syntactic copy of the antecedent would result in ungrammaticality[9]:

(64) *Harry does have any paper.

This fact is explained by treating NPI "any" semantically as an indefinite, together with some syntactic constraints on its distribution. The situation is parallel to that of reflexives and wh-traces: under ellipsis, only the semantic representation is copied, and syntactic constraints within the antecedent are therefore not imposed under ellipsis.

Idiomatic negative polarity items behave similarly:

(65) A: Nobody **lifted a finger** when the crisis took place!
B: What do you mean? Jones certainly did!

Here are some naturally occurring examples of this:

(66) At the prices we were charged, there should have **been some return for the dollar**. There wasn't. (Wall St. Journal)

(67) A USX spokesman said the company hadn't yet **received any documents from OSHA regarding the penalty or fine**. "Once we do, they will receive very serious evaluation." (Wall St. Journal)

(68) "We haven't decided to **blacklist any firms**. But there's a chance we might," said David Wilson, head of Penn Mutual's $100 million stock portfolio. (Wall St. Journal)

(69) Some people say there **are no real movie stars anymore**. I say there are. (Academy Awards 3/25/91)

[9]This is ignoring the "free choice" reading of "any", under which the example might be acceptable.

2.8 Reciprocals

I turn next to cases in which the antecedent VP contains a reciprocal. Under some natural assumptions about reciprocals, these cases fall under Schema 2 as well. I will argue that a reciprocal expression introduces a variable which is bound by an operator outside the antecedent VP. Thus, we expect to see a similar flexibility in sloppy readings for such variables. Consider the following example:

(70) Martha and Irv wanted to dance with each other.

I will assume that the following predicate ϕ is being applied to the subject "Martha and Irv":

ϕ = λ x.want(x,dance(x,y))

Following recent work on reciprocals [Heim *et al.*, 1991; Roberts, 1991], I will assume that the reciprocal "each other" makes two separate semantic contributions: "each" is a distributive operator (DIST) applied to the VP ϕ;"other" is a variable that is bound outside ϕ. Thus (70) is represented:

[Martha and Irv] [DIST [λ x.want(x,dance(x,y))]$_{VP}$]$_{VP}$

Consider now the following example:

(71) Martha and Irv wanted to **dance with each other**, but Martha couldn't, because her husband was here. [Webber, 1978]

In this example, the syntactic antecedent "dance with each other" is not be possible at the ellipsis site:

(72) * Martha couldn't dance with each other.

However, the example can be readily explained by reference to the semantic representation given above. The VP meaning ϕ:

λ x.want(x,dance(x,y))

is selected as the antecedent for the elliptical VP. Then, as we have seen in a wide variety of examples, the free variable can be reinterpreted at the ellipsis site. Thus, once the representation given above is established, the crucial point is that we allow the meaning of the non-maximal VP to be selected as the antecedent: that is,

[λ x.want(x,dance(x,y))]

rather than

[DIST [λ x.want(x,dance(x,y))]]

While there may be a general preference for selecting the maximal VP as antecedent, the non-maximal VP can clearly be selecting under appropriate circumstances. This is readily seen with another type of VP operator, namely adverbials.

(73) a. John walked quickly,

b. and Bill did too.

c. and Bill did slowly.

Thus while (73)b clearly selects the maximal VP "walked quickly", (73)c selects the non-maximal VP "walked". There is a similar pattern with the distributive operator:

(74) a. The boys [DIST [ate a bagel for breakfast.]$_{VP}$]$_{VP}$

b. The girls didn't.

c. Susan didn't.

Here, (74)b selects the maximal VP, including the DIST operator, while (74)c selects the non-maximal VP, without the DIST operator.

To return to example (70), we have seen that the dynamic property ϕ, λ**x. dance(x,y)** is applied distributively in the antecedent to both Martha and Irv. The same dynamic property is then applied (non-distributively) to Martha in the elliptical sentence, with the context determining the referent of the free variable.

It might appear that the following alternative analysis of (70) is available: only the verb "dance" is copied to the ellipsis site, appearing there in "detransitivized" form. However, the same reading is possible for a verb like "nominate" which cannot be detransitivized:

(75) Martha and Irv had planned to **nominate each other**, but Martha couldn't, because of her political obligations.

Here, the reading is "Martha couldn't nominate Irv".

Just as in the cases involving pronouns and traces, we have an antecedent VP with a variable that is bound outside the VP. Therefore, the ellipsis is resolved by copying the dynamic property denoted by the antecedent VP, and the binding of the free variable is determined independently in the ellipsis site.

2.9 A Semantic Constraint: Deictics

We have seen that syntactic constraints on variables are not imposed under ellipsis. This is consistent with the proverb theory, since it does not treat the elliptical VP as a syntactic copy of the antecedent, but rather, it interprets the proverb as semantically identical to its antecedent. This implies that any semantic constraints on variables in the antecedent should be respected under ellipsis. In this section, I give evidence that this is the case. I will examine what is perhaps the most widely studied semantic constraint on variables: the constraint that indexical or deictic variables cannot vary across contexts; that is, deictics are "rigid designators" (cf. [Kaplan, 1989; Nunberg, 1991].

In general, variables are permitted to vary with context, so that a single variable might designate different individuals. Consider the following example:

(76) The speaker could have been a linguist.

Here, the modal operator "could" requires the consideration of many different possible contexts. There is no requirement that "the speaker" designate the same individual in all these contexts. That is, (76) would be truly uttered by me, a computer scientist, in a situation in which I am the speaker, but a linguist friend of mine had been a possible alternative speaker. If we assume that "could" is a sentential modal operator expressing possibility, we can represent (76) as follows:

$\exists$ w. **The Speaker** is a linguist in w.

Assume that the current world is w_0; in w_0, **The Speaker** denotes me. However, some world w_n could make (76) true, where **The Speaker** denotes some other individual who is a linguist in w_n.

This is not possible if we replace "the speaker" with a rigid designator, such as the deictic "I":

(77) I could have been a linguist.

This does not have the same truth conditions as (76). Again, the modal requires consideration of many possible contexts, but in all these contexts, "I" must designate the same individual.

What is special about "deictics" is that their meaning does not become part of the utterance meaning; a deictic expression is simply used to pick out a referent and then discarded. The distinction between deictic and non-deictic expressions becomes observable when an expression is evaluated with respect to more than one context: the deictic will designate the same individual in all such contexts, while the non-deictic may designate different individuals in different contexts.[10]

This requirement has been investigated mainly in modal contexts, where a single variable is considered with respect to several contexts. It has generally been assumed in this literature that third-person pronouns are ambiguous, having both a deictic and non-deictic use. First and second-person pronouns have generally been held to be "dedicated deictic" expressions, having no non-deictic use.[11]

A similar possibility arises with VP ellipsis; if there is a variable within the antecedent, it will be evaluated with respect to two contexts – the antecedent context, and that of the ellipsis site. We would expect that a deictic variable within the antecedent VP would not switch referent under ellipsis. The following example shows that this is the case:

(78) Speaker A: Do you think they will **like me**?

(79) Speaker B: Yes, I'm sure they will.
[Sag and Hankamer, 1984]

The only possible reading for the elliptical VP is "like (Speaker A)". This follows the requirement that the deictic "me" in the antecedent must designate the same individual in both the antecedent and ellipsis contexts.[12] If one syntactically copied the antecedent "like me", the wrong

[10] See [Nunberg, 1992] for a recent elaboration of this distinction, including some observations on its relevance to ellipsis phenomena.

[11] In fact, there appear to be non-deictic uses of first and second-person pronouns, as well. However, the examples considered here are, I think, clearly deictic uses.

[12] There are apparent contradictions to this claim, as pointed out by Mitch Marcus (p.c.) For example:

(80) Speaker A: My students like me.

(81) Speaker B: Mine do too.

I would claim that the occurrence of "me" in the antecedent is not a deictic expression in this case.

reading would result.

It is interesting to examine another type of variable in this context: it has been suggested [Partee, 1991] that expressions like "locally" involve a "hidden" variable that can be bound by other operators. On the other hand, the term "here" is generally recognized to be deictic. This explains the following contrast:

(82) John **shops locally**. Susan does too.

(83) John **shops around here**. Susan does too.

The hidden variable associated with "locally" permits a sloppy reading, in which Susan shops in a different area than John does. The sloppy reading is not available for "around here"; Susan must shop in the same area that John shops in. This contrast follows directly from the deictic nature of "here", and the non-deictic nature of the hidden variable.

Just as in the case of pronouns, the hidden variable can get a sloppy reading even if it is not subject-bound:

(84) a. If John needed advice, I would offer to **meet him at a local bar**.

b. If John needed advice, I would **meet him at a bar around here**.

(85) If Tom needed advice, I would too.

Consider the reading for (84)a in which I am offering to meet John at a bar "local to John". With antecedent (84)a, the sloppy reading is possible for (85): I may be suggesting to meet Tom in a bar in Tom's neighborhood. This reading is not possible with antecedent (84)b. The sloppy reading is possible despite the fact that the invisible variable is not subject-bound.

2.10 The Bound-Sloppy Correlation

It has frequently been argued that the distinction between strict and sloppy readings can be correlated with the distinction between bound and free variables. The basic constraint in the Sag/Williams account, for example, could be expressed as follows:

(86) A free variable in the antecedent must have a strict reading at the ellipsis site.
A bound variable in the antecedent must have a sloppy reading.

The facts presented thus far shows that this correlation cannot be maintained. That is, sloppy readings are possible for pronouns that are not generally held to be bound variables, as in example (3), repeated here.

(87) If Tom was having trouble in school, I would **help him**. On the other hand, if Harry was having trouble, I doubt that I would.

The pronoun "him" in the antecedent receives a sloppy reading, although it is not a bound variable. It corefers with Tom, which is not treated as an operator. Furthermore, it is not c-commanded by "Tom".

However, perhaps the problem is not the bound-sloppy correlation, but rather, the notion of a bound variable. I treat any variable that refers to an entity in the discourse model as a bound variable. This is made explicit in [Heim, 1982], where an operation of "existential closure" binds all variables that refer to "discourse entities". The only variables that do not refer to discourse entities are deictics, as discussed above. Deictics do not refer to linguistically introduced discourse entities, and thus are not subject to existential closure, and of course, they are not bound by other operators, either. Thus the bound-sloppy correlation is maintained, but it is only deictics that are treated as free variables. Instead of (86), then, we have:

(88) A deictic variable in the antecedent must have a strict reading at the ellipsis site. Other variables may be strict or sloppy.

This accords with the facts I have presented, while retaining the intuition that the bound-sloppy correlation corresponds to a fundamental distinction concerning pronouns. In addition, it leads to another prediction:

(89) Variables that do not permit sloppy readings are those that cannot vary in modal contexts.

Consider again this example:

(90) Speaker A: Do you think they will **like me**?

(91) Speaker B: Yes, I'm sure they will.

The variable "me" does not permit a sloppy reading. Consider the following example:

(92) Possibly, they will like me.

In this example, the variable "me" must refer to the same individual in all contexts the operator "possibly" ranges over. The phenomenon of "sloppy identity" becomes, from this perspective, a reflection of a very simple fact: non-deictic variables are not required to designate the same individual in different contexts.

Thus the strict/sloppy ambiguity is a reflection of the same ambiguity seen in modal contexts, the so-called referential/attributive ambiguity[Donnellan, 1966], illustrated by the following example:

(93) The President could have been a Republican.

On the referential (strict) reading, (93) is true only if it was possible that Bill Clinton was a Republican. The attributive (sloppy) reading is true if it was possible that some Democrat was president. All these ambiguities are assimilated to a single, processing ambiguity: the variable can be interpreted either before or after the different contexts are examined. This provides a uniform explanation of the fact that, in all these cases, deictic variables do not permit the ambiguity.

2.11 More Complicated Antecedents

In the cases considered so far, there has always been a single antecedent VP. Most studies of VP ellipsis have claimed that VP ellipsis *requires* a single VP as its antecedent [Hankamer and Sag, 1976; Partee and Bach, 1984]. In fact, there are a variety of other antecedence possibilities. In this section I look at cases involving "split antecedents", "long distance antecedents" and "missing antecedents". These possibilities are strikingly similar to the antecedence possibilities for pronouns.

2.11.1 Split Antecedents

In some cases, the elliptical VP refers to two distinct VP's in surrounding discourse. Consider the following example:

(94) Wendy is eager to **sail around the world** and Bruce is eager to **climb Kilimanjaro**, but neither of them can because money is too tight. [Webber, 1978]

The reading of interest is that Wendy cannot sail around the world, and Bruce cannot climb Kilimanjaro. Clearly, this reading will not be permitted by syntactically copying either of the preceding VP's to the ellipsis site; a combination of the two VP's is required. It is well known that plural pronouns can refer to split antecedents. Thus I will simply make the same combining operation available for VP ellipsis[13]. This combining operation is formulated as a general operation on semantic objects in a discourse model, as follows:

(95) $\{a \ldots b \ldots\} \Rightarrow \{a \ldots b \ldots [a,b]\}$

That is, given a discourse model $\{a \ldots b \ldots\}$, under appropriate conditions the discourse model can become $\{a \ldots b \ldots [a,b]\}$, where [a,b] denotes a combination of a and b. This is a natural and independently motivated operation, needed to account for plural pronouns such as "them" in example (94).

Another example is the following:

(96) I can **walk**, and I can **chew gum**. Gerry can too, but not at the same time. [Webber, 1978]

Here the elliptical VP is a predicate combined from "walk" and "chew gum". Here is a naturally-occurring example:

(97) So I say to the conspiracy fans: **leave him alone**. **Leave us alone**. But they won't. (*The Welcomat*, 2/5/92 p.25)

Consider the following variant of (96):

(98) I can **walk**, and I can **chew gum**. Harry and Phil can too.

This cannot mean "Harry can walk, and Phil can chew gum", although the following reading is possible: "Harry can walk and chew gum, and Phil can walk and chew gum."

This range of possibilities can be defined according to the following schema:

Properties:

[13] I will not define the structure imposed by this combining operation – For my purposes, it is sufficient to treat the combining operation as simple list-formation. While the summation operation discussed by [Link, 1983] may be preferable, I will not address this issue here.

Rule 1: a [P,Q] ⇒ Pa AND Qa

Rule 2: [a,b] [P,Q] ⇒ Pa AND Qb ("respectively" reading)

Rule 3: [a,b] [P,Q] ⇒ a[P,Q] AND b[P,Q]

Rule 1 states that, if a combined property is applied to a single entity a, each property is applied to a, as in example (96). If a combined property is applied to a subject that is also combined, an ambiguity arises. The first possibility is what I call a "respectively" reading, given by Rule 2 above. This reading arises in example (94). Finally, example (98) illustrates the reading given by Rule 3.

The "respectively" reading is possible only if the cardinality of the two combined objects is the same. In addition, there must be a salient mapping that would give rise to the "respectively" reading. That is why the "respectively" reading is not available for example (98). However, consider the following discourse:

(99) In general, I think the husband should cook dinner and his wife should wash the dishes. So tonight, Mr Smith will cook dinner and Mrs Smith will wash dishes. Tomorrow, Mr Jones will **cook dinner** and Mrs Jones will **wash dishes**. Or if they can't, then Mr and Mrs Wilson will.

Here the "respectively" reading is available. That is, "Mr and Mrs Wilson will" can be interpreted as "Mr Wilson will cook and Mrs Wilson will wash dishes".

A similar schema can be applied to cases involving combinations of entities.[14] For example, in cases where the subject and object are both combined objects, we have a similar pattern:

Individuals

Rule 1: x V [u,v] ⇒ V(x,u) AND V(x,v)

Rule 2: [x,y] V [u,v] ⇒ V(x,u) AND V(y,v) ("respectively" reading)

Rule 3: [x,y] V [u,v] ⇒ V([x,y],u) AND V([x,y],v)

These rules can be illustrated by the following examples:

[14] In Chapter Three, I will suggest that these two rule schemas both reflect the optional application of a "respectively" operator.

(100) (Rule 1) John$_1$ likes Susan$_2$ and Mary$_3$. He$_1$ likes them$_{2,3}$ because . . .

J like [Susan,Harry] $\Rightarrow$ like(J,S) AND like(J,M)

(101) (Rule 2) John$_1$ likes Susan$_2$ and Bill$_3$ likes Mary$_4$. They$_{1,3}$ like them$_{2,4}$ because . . .

[J,B] like [S,M] $\Rightarrow$ like(J,S) AND like(B,M)

(102) (Rule 3) I like Susan and I like Harry. My parents like them too.

[mother,father] like [S,H] $\Rightarrow$
like ([mother,father],S) AND like([mother,father],H)

For the above example, the distributed reading is not available; it can't mean "my father likes Susan and my mother likes Harry". This is because there is no salient basis for such a mapping. Consider the following example:

(103) Susan's mother$_1$ likes Susan$_3$, and Mary's mother$_2$ likes Mary$_4$. Their fathers$_{5,6}$ like them$_{3,4}$ too.

Here the "respectively" reading is available, in which Susan's father likes Susan, and Mary's father likes Mary.

In sum, there is a single operation for combining objects in the discourse model, applying to both individuals and properties.[15] In sentences with more than one such combined object, there is the possibility of additional readings, involving mappings between members of the combined objects.

2.11.2 Long Distance Antecedents

It is generally believed that memory for syntactic structure is relatively short-lived, while semantic information is retained over longer stretches of discourse. This has been described as "one of the

[15] Lappin points out (p.c.) that Rule 2 assumes a conjunction reading, while examples like (94) are in fact interpreted *disjunctively*, i.e., "neither of them can sail around the world *or* climb Kilimanjaro, respectively." Another way of describing this is as a conjunction where the negation is distributed to both conjuncts. The analogous cases involving pronouns exhibit the same reading.

(104) Susan's mother$_1$ likes Susan$_3$, and Mary's mother$_2$ likes Mary$_4$. Neither of their fathers$_{5,6}$ like them$_{3,4}$, though.

This means Susan's father doesn't like Susan, and Mary's father doesn't like Mary.

best-established results in the psycholinguistic literature." [Garnham and Oakhill, 1987] There seems to be no bound on the distance separating the antecedent from the elliptical VP. In a study of VP ellipsis occurrences in the Brown Corpus (discussed in Chapter Six), I found that about 5% of VP ellipsis occurrences were long-distance, i.e., where one or more sentences intervened between antecedent and VP ellipsis [Hardt, 1990]. One example is the following:

(105) I disagree with the writer who says funeral services **should be government-controlled**. The funeral for my husband was just what I wanted and I paid a fair price, far less than I had expected to pay. But the hospitals and doctors should be. (Brown Corpus)

Most of the literature on VP ellipsis focuses on examples in which the antecedent and VP ellipsis are in the same sentence, or in two adjacent, conjoined sentences. However, it is clear that VP ellipsis is not restricted to such configurations. In examples with one or more intervening sentences, it is clear that the VP ellipsis and its antecedent would not be contained within a single syntactic tree. Thus it would appear that the relation between antecedent and VP ellipsis is beyond the scope of syntactic rules, as traditionally conceived. Furthermore, it appears that the distance between antecedent and VP ellipsis can exceed the normal memory for syntactic structure.

Such long-distances examples have sometimes been noted in the literature. [Klein and Stainton-Ellis, 1989] report several naturally-occurring examples of long-distance VP ellipsis. They also point to examples which involve multiple occurrences of VP ellipsis, such as the following:

(106) A: You never **go swimming**.

(107) B: That's because I don't **look good in a swimming costume**. I might if I did.

2.11.3 Missing Antecedents

In this section I consider cases where there is no VP antecedent. This has widely been held to be impossible, and such cases present a fundamental problem for a syntactic copy approach. I will suggest that missing antecedent cases are only possible under specific pragmatic conditions, similar to the case of pronouns.

Deictics

In some cases, the antecedent is available from the non-linguistic context. For example,

(108) You shouldn't have. (on receipt of a present)

(109) I will if you will. (standing on the edge of a cold swimming pool, about to jump) [Chao, 1987]

(110) Not in my wastebasket you don't! (to someone about to toss something in the wastebasket) [Schacter, 1977]

In all these cases, the missing property is contextually available, and not linguistically present. This clearly contradicts the standard syntactic approaches to VP ellipsis, in which a syntactic VP antecedent is required. On the other hand, the lack of a linguistic antecedent is not surprising for a semantic approach, since semantics provides the interface between linguistic and non-linguistic phenomena.

Nominal Antecedents

In this section I consider cases where the antecedent is a noun or NP rather than a VP.

(111) David Begelman is a great **laugher**, and when he does, his eyes crinkle at you the way Lady Brett's did in The Sun Also Rises. (p. 90, *You'll Never Eat Lunch in This Town Again*, Julia Philips)

In this example, the noun "laugher" is the antecedent for the elliptical VP. Since a noun semantically denotes a property, this could be naturally accommodated by a semantic approach. However, not all nouns seem to provide antecedents for VP ellipsis. In this example, the property expresses the "main predication" of the sentence. A sentence of this form:

(112) NP is a N

is semantically equivalent to a sentence of this form:

(113) NP VP where [[VP]] = [[N]].

Thus we may treat "David is a laugher" as roughly equivalent to "David laughs".

Here are some other (constructed) examples of the same form:

(114) Harry used to be a great **speaker**, but he can't anymore, because he has lost his voice.

(115) People say that Harry is an excessive **drinker** at social gatherings. Which is strange, because he never does at my parties.

Nominalized verbs can also serve as antecedent for VP ellipsis:

(116) We should suggest to her that she officially appoint us as a committee and invite **faculty participation**. They won't, of course, . . .
(UPENN email message)[16]

(117) Today there is little or no OFFICIAL **harassment of lesbians and gays** by the national government, although autonomous governments might. (b-board msg)[17]

The following example also has a NP antecedent.

(118) [Many Chicago-area cabdrivers] say their business is foundering because the riders they depend on – business people, downtown workers and the elderly – are opting for the bus and the elevated train, or are on the unemployment line. Meanwhile, they sense a drop in **visitors** to the city. Those who do, they say, are not taking cabs.
(Chicago Tribune, 2/6/92) *(Gregory Ward, p.c.)*

The NP "visitors" seems to be the antecedent for the elliptical VP. One possible explanation for this is that NP's such as "visitors" are collections, not of individuals, but visiting events.

Gerundive NP's may provide the antecedent for VP ellipsis:

(119) **Seeing them** did not greatly surprise Enid either, though she would wish later she hadn't. (*You Must Remember This*, Joyce Carol Oates, p. 287)

The antecedent for the VP ellipsis "hadn't" is "seeing them", which is in subject position, and thus is presumably a syntactic NP.

(120) The candidate was dogged by charges of infidelity and **avoiding the draft**, or at least trying to.

Here, the phrase "avoiding the draft" appears to be an NP, since it is coordinated with the NP "infidelity". However, it also is the antecedent for the elliptical VP "trying to".

[16]This example was provided to me by Bonnie Webber.
[17]This example was provided to me by Ellen Prince.

The following example[18] suggests that the theme-rheme structure of the antecedent is relevant.

(121) A: Could you tell us about the party held in your home last night, ma'am?

(122) B: It was a very ordinary bash.

(123) A: What happened to one of your guests wasn't. ("Streets of San Francisco" rerun)

Here, the antecedent property for the elliptical VP is simply "ordinary". This corresponds with the fact that the noun "bash" is quite redundant. Thus the actual predication structure of (122) is:

λ x.ordinary, it

I suggested with respect to example (111) that the predicate nominal can make available the property that is not denoted directly by the VP, but is the property that is predicated of the subject, namely, the property denoted by the N-bar. The current example suggests it would suggest that a subpart of the N-bar can be made available as an antecedent for VP ellipsis – not an arbitrary subpart, but the non-redundant subpart. This would suggest, for example, a contrast between the following two discourses: in the first discourse, the elliptical VP can have the antecedent "Russian", because it is the non-redundant portion of the antecedent predicate nominal – "spy" is redundant. In the second discourse, "spy" is not redundant, and the only reading of the elliptical VP accesses the entire predicate nominal.

(124) A: What sort of spy is Jones?

(125) B: I heard he's a Russian spy!

(126) A: That's funny – none of his friends are. [Russian/?Russian spy]

(127) A: What does Jones do?

(128) B: I heard he's a Russian spy!

(129) A: That's funny – none of his friends are. [*Russian/Russian spy]

Predicate nominals can also provide the antecedent for VP ellipsis[19], as pointed out by Baltin [Baltin, 1991]:

(130) I consider Sally an excellent teacher, but I don't think Sam is.

[18] In this example, it is not clear whether the ellipsis is VP ellipsis or AP ellipsis.

[19] Again, the status of this example as VP ellipsis or AP ellipsis is uncertain.

Active/Passive Mismatches

Another possibility for VP ellipsis with no VP antecedent arises with active/passive mismatches. This is illustrated by the following naturally-occurring cases:

(131) This information could have been released by Gorbachov, but he chose not to. (Daniel Schorr, NPR 10/17/92)

(132) Business needs to be developed differently than we have in the past. (5/24/91 NPR "Morning Edition" interview)

(133) The ice cream should be taken out of the freezer, if you can. (heard in conversation)

(134) A lot of this material can be presented in a fairly informal and accessible fashion, and often I do. (Chomsky 1982, cited in [Dalrymple *et al.*, 1991])

In these examples, the antecedent for the elliptical VP is not a VP. In (131), the antecedent for the elliptical VP is something to the effect of "release this information" – this does not appear as a VP in the antecedent. However, it would be the VP in the active counterpart of the antecedent. I will term this the "active property". These examples show that, in a passive sentence, the active property can be the antecedent for VP ellipsis. [Hankamer and Sag, 1976] argue that this is not possible for VP ellipsis, based on examples such as the following:

(135) * The oats had to be taken down to the bin, so Bill did.

(136) * The children asked to be squirted with the hose, so we did.

However, while Hankamer and Sag judge these to be ungrammatical, these judgements are not universally accepted; for example, [Chao, 1987] reports that many speakers find (136) acceptable. In my judgement, these examples are somewhat degraded, while (131) - (134) are quite a bit better. It appears that one cannot categorically rule out such cases, as Hankamer and Sag do. Instead, I will allow the grammar to make these active properties generally available in the discourse model. To explain the differences in acceptability among the above examples, I will appeal to independently required preference factors governing the selection of antecedents. These preference factors are discussed in detail in Chapter Five. There, I will suggest an explanation for the above contrast based on these preference factors.

The converse is also sometimes possible, where the antecedent is the VP of the passive counterpart of the antecedent sentence:

(137) ?Max fired Harry, although it was Tom who should have been. ([Fiengo and May, 1992b])

(138) HARRY they fired, although it was TOM who should have been.

The contrast can be explained as follows: the main predication of a sentence is typically the VP being applied to the subject. However, certain forms, such as topicalization, can express some other predication. For example, in "MAX, they fired", the predicate λ *x.they fired x* is applied to Max. In this case, the property formed by topicalization is the main predication of the sentence, and this property is available as an antecedent for VP ellipsis.

It appears, then, that VP ellipsis is sensitive to the predication structure of sentences; when the main predicate is something other than a VP, that predicate can be an antecedent for VP ellipsis.

2.12 Conclusions

The following are the major observations of this chapter:

1. variables bound outside the antecedent VP can give rise to sloppy readings, violating the Sag/Williams account of VP ellipsis.

2. When the antecedent contains a variable under syntactic constraints, such as a wh-trace, reflexive, reciprocal, or NPI, those constraints can be ignored at the ellipsis site.

3. The only variables that *require* strict readings are deictics; these are variables that are subject to a general semantic constraint, namely, that they do not vary across contexts.

4. This allows a reformulation of the well-known correlation between bound variables and sloppy readings; any non-deictic variable is considered a bound variable, so that only deictics require strict readings. This is correlated with the fact that deictic variables do not interact with modal operators.

5. The antecedent is not necessarily a single syntactic VP: split antecedents are possible, as well as deictic cases of VP ellipsis. This parallels the antecedence possibilities for pronouns, and thus supports the view that VP ellipsis is a proform.

6. Cases of “missing antecedents” are also possible, in which the antecedent is not a syntactic VP. In these cases, I suggest that the antecedent can be the “main predicate”, where predication is effected by means other than a syntactic VP. Again this parallels the case of pronouns; while there is a preference for an NP antecedent, an antecedent can be made available by other means.

All of these facts lead to the conclusion that the elliptical VP is a proverb, and is thus semantically identified with its antecedent.

Chapter 3

Formal System

3.1 Overview

In this chapter I describe a system of semantic derivation which implements the proverb theory of VP ellipsis, providing a solution to the empirical problems presented in the preceding chapter. The solution given is optimal in the sense that no special mechanisms must be stipulated for the case of VP ellipsis; rather, the phenomena of interest are shown to follow from quite general mechanisms governing semantic derivation and the interpretation of proforms.

A proform is interpreted by identifying an antecedent – a semantic object which has been stored, temporarily, in a discourse model. For example, an indefinite NP stores a semantic object in the discourse model, which a pronoun can select as its antecedent. Similarly, a VP stores a VP meaning in the discourse model, which can be selected by an elliptical VP as its antecedent.[1] Proforms are uniformly governed by a semantic identity condition. Typically, this means that a pronoun will denote the same individual as its antecedent, and an elliptical VP will denote the same property as its antecedent. However, meanings themselves involve dependencies on context: the meaning of an NP is a relation on contexts and individuals, which can be termed a dynamic individual. The meaning of a VP, a relation on contexts and properties, is termed a dynamic property. The basic constraint for VP ellipsis, then, is an identity of dynamic properties.

The dependency of a proform on its antecedent is often denoted by coindexing. Typically

[1]The approach in [Klein, 1984] is similar in that it extends the discourse model to include properties. They are not dynamic properties, however. Thus sloppy identity cannot be handled in that system as it is here. Instead, [Klein, 1984] essentially duplicates the Sag/Williams treatment of sloppy identity.

in linguistic theories, indexes are assigned prior to the semantic derivation. From a processing perspective, it is more natural to define a step in the derivation in which the antecedent is selected, based on relevant constraints and heuristics. To represent this, I will notate an anaphoric expression by an unindexed *parameter*, associated with an *assumption* representing constraints on its eventual referent. At some point in the derivation, the assumption must be *discharged*, at which time the parameter is replaced with a *discourse marker*, selected from the discourse model. In general, all interaction with the discourse model is mediated by the introduction and discharge of assumptions. [2] Notationally, this approach requires two different types of variables: parameters and discourse markers. Initially, a pronoun (or other anaphoric expression) is represented by an assumption together with a parameter. When the assumption is discharged, the parameter is replaced with a discourse marker. Parameters are notated in ordinary type, and discourse markers in **bold**.

The two central tenets of the approach are the following:

1. Meanings are stored in the discourse model.

2. Anaphoric expressions do not have their antecedents predetermined–they are determined at some stage during the derivation.

The system described in this chapter results from necessary implications of these basic tenets. The mechanism of assumption storage reflects point (2), because it allows an object to be represented in a derivation both before and after its antecedent is determined.

Point (1) requires that we store dynamic objects in the discourse model. This follows from the claim that meanings are *dynamic*, i.e., they may involve dependencies on context. This is what permits sloppy identity in the system, as will be described below.

Consider first the case of pronouns and their antecedents. As in [Heim, 1982], the notion of "definiteness" is central: an indefinite NP adds a semantic object to the discourse model, and a definite NP selects an antecedent from the discourse model. In the approach taken here, this is implemented with definite and indefinite assumptions, so that a definite assumption, upon discharge, selects an object from the discourse model, while an indefinite assumption, upon discharge, adds an object to the discourse model. To apply this to the case of VP ellipsis, it is necessary to stipulate that a VP is indefinite and an elliptical VP is definite. This means that a VP adds a dynamic

[2] This approach to the indexing of anaphoric expressions is based on a computational system of semantic interpretation developed by [Pereira and Pollack, 1991].

property to the discourse model, and an elliptical VP must select a dynamic property from the discourse model.

Consider the following discourse:

(1) **A man** walked in. He was tall.

The indefinite NP "a man" is the antecedent for the pronoun "he": the indefinite "a man" adds an individual to the discourse model, which can be selected by the pronoun. Consider now the case of VP ellipsis:

(2) John **walked**. Harry did too.

The situation is parallel: the occurrence of the VP "walked" adds a property to the discourse model, and the VP ellipsis can select that property as its antecedent. In general, an indefinite expression is one whose meaning is saved in the discourse model.

I will describe a system of derivation based on a function application rule and a rule for assumption discharge. The discharge of an assumption either adds an object to the discourse model or accesses an object in the discourse model. In addition, the system includes a "discourse" rule which combine sentences together. Once I have described these rules, I show how several examples from Chapter Two are derived.

3.2 Two Assumption Types: Definite and Indefinite

There are two ways an expression can interact with the discourse model: an indefinite expression adds its meaning to the discourse model, and a definite expression accesses an object in the discourse model. Since interaction with the discourse model is implemented through the assumption store, we have two types of assumption: definite and indefinite. This is based on a lexical feature, *def/indef*.

The meaning of an expression is an assumption-sense pair, where every unindexed parameter in the sense has an assumption associated with it. For example, the semantic representation of a pronoun "she" is the following:

"she" $\{<x,def,female \ \& \ sing>\}:x$

The sense is the parameter x. The feature *def* requires that, at some stage in the derivation, an antecedent must be selected in the discourse model, where the antecedent is consistent with the other semantic features, i.e., female and singular. Consider a discourse model in which a single female individual, $\mathbf{m}_1$ is available. The parameter x is replaced with the discourse marker $\mathbf{m}_1$, and the definiteness requirement is discharged. (The parameter can also be replaced by a variable bound by an quantifier or wh-operator in the current context.)

In general, then, an object with a semantic feature *def* will add an assumption of the form <x,def,features> to the assumption set of the derivation, where *features* denotes features such as number and gender, and x is the parameter. This *def* assumption must be discharged at some stage in the derivation, causing the parameter x to be replaced with an individual in the input discourse model. In simple cases, the antecedent is a discourse marker, and the antecedent has an empty assumption set. In some cases, the antecedent may have contextual dependencies, which will be represented as undischarged assumptions. These undischarged assumptions are added to the assumption set. In such cases, there is a possibility for "sloppy identity".

Consider now an indefinite NP: the indefinite "a man" adds some individual m to the discourse model, with an empty assumption set. Again, this follows a lexical feature, in this case *indef*, associated with the indefinite article "a".

"a" $\{<$ indef, x$>\}$:λ P.x | P(x)

The property P is associated with the common noun. (The notation x | P(x) is to be read, "x of type P".) To see how this works, consider the lexical meaning of the noun "man":

"man" $\{\}$:λ x.man(x)

The two expressions are combined as follows: the two senses are combined by function application, and the two assumption sets are combined by set union, giving the following object:

"a man" $\{<$ indef, x $>\}$:x|man(x)

The discharge of the *indef* assumption results in an individual with an empty assumption set, and the discourse marker **e** is added to the discourse model, together with the type man.

"a man" $\{\}$:**e**|man(**e**)

The object added to the discourse model by an indefinite NP is a function from contexts to individuals. If the assumption set is empty, as in the above case, it is a constant function from contexts to individuals; that is, it will determine the same individual in any context. In this case, the sense is represented as a discourse marker. If the assumption set is non-empty, the object may not be a constant function. This reflects the fact that undischarged assumptions represent contextual dependencies. In this case, the sense would be a parameter rather than a discourse marker.

When the indef expression is an NP, the antecedent expression is typically a discourse marker, with an empty assumption set. VP's are also indefinite, and introduce semantic objects with non-empty assumption sets into the discourse model. The presence of undischarged assumptions associated with an object in the discourse model permits "sloppy identity". This will be illustrated below, when we extend the system to handle VP ellipsis.

3.3 Rules for VP Ellipsis

3.3.1 Syntactic Background

The syntactic structure of an elliptical VP is an INFL node, dominating an auxiliary verb, and a VP node, dominating an empty category. It has long been recognized that VP ellipsis does not require "affixal identity"; for example, [Sag, 1976] cites examples from [Quirk, 1972] such as the following:

(3) John **understands the situation** and surely Peter should.

The explanation for this is that the VP node always dominates the base form of verb, and all verbal inflection is under the INFL node. This permits example (3) to be treated as an identity of VPs. However, as Quirk also points out, there are cases where the elided VP appears to require an inflected form. For example,

(4) John may be **questioning out motives**, but Peter hasn't.

Here, the elided VP would be "questioned our motives". This could perhaps be explained by claiming that the auxiliary "hasn't" signifies the perfect tense, so that the two structures would be:

[progressive, neg]$_{INFL}$ [question our motives]$_{VP}$

[perfect, pos]$_{INFL}$ [question our motives]$_{VP}$

This explains many of the differences in tense between antecedent and elided VP. This explanation would be available to either a syntactic or semantic identity condition. However, consider the following example:

(5) Last year before the election, Smith **predicted that he was going to win**. And next year before the election, I'm sure he will again.

Here, syntactic identity is violated, because the elided VP cannot be "predicted that he *was* going to win", but rather, "predict that he *is* going to win". Semantically, the VP meaning in each case is the same, i.e.;

λ x.x predict at time t x win at time t$'$,t$' >$ t.

3.3.2 Semantic Rules for VP Ellipsis

Above I described a system in which a pronoun can have an indefinite NP as its antecedent. The indefinite, because of a lexical feature *indef*, adds the NP meaning to the discourse model. The pronoun, because of a lexical feature *indef*, selects an individual from the discourse model.

To extend this to cover VP ellipsis, it is sufficient to make the following stipulations:

1. The overt VP is *indef*.
2. The elliptical VP is *def*.

An elliptical VP is represented as a property variable:

{<P,def>}:P

The semantics for the auxiliary verb "do" is as follows:

"do" {}:λ P.PRESENT(P)

I will go through a simple example.

(6) John walked. Bill did too.

The semantic representation for the VP "walk" is:

"walk" {<indef>}:λ x.walk(x)

The indef assumption is discharged, adding this object to the discourse model:

{}:λ x.walk(x)

We continue the derivation of the sentence, arriving at:

PAST(walk(John))

Next, we derive "Bill did P too." The elliptical VP is represented:

P {<P,def>}:P

Next, the def assumption is discharged, causing P to be replaced with the stored property, resulting in:

P {}:λ x.walk(x)

This is combined with the subject, giving:

Bill P {}:walk(Bill)

This is then combined with "did"[3] giving:

Bill did P {}:PAST(walk(Bill))

Just as with indefinite NP's, a VP adds its assumption-sense pair to the discourse model. Any undischarged assumptions associated with the VP will be stored along with the sense. Whenever there is more than one assumption in the assumption set, an ambiguity arises, since there is no ordering on the discharge of assumptions. It is this ambiguity that accounts for the possibility for strict and sloppy readings under ellipsis. As an illustration of this, consider example (7):

(7) Tom thinks he is smart. Harry does too.

The VP "thinks he is smart" is represented as follows:

thinks he is smart {<x,def,male>,<indef>}:λ y.think(y,smart(x))

[3]I am treating the auxiliary uniformly as a propositional operator. I don't address the issue of how this semantic derivation is mapped from the surface syntax. One possibility is to treat INFL uniformly as a raising verb, as proposed by [Koopman and Sportiche, 1991].

There are two assumptions: a *def* assumption one associated with the variable "he", and an indef assumption associated with the matrix VP "thinks he is smart"[4]. The *def* assumption, upon discharge, determines the referent of the pronoun. The *indef* assumption, upon discharge, stores the current VP meaning in the discourse model, as an assumption:sense pair. If the *indef* assumption is discharged first, the following assumption:sense pair is stored in the discourse model:

$\{<$x,def,male$>\}:\lambda$ y.think(y,smart(x))

Since the *def* assumption is as yet undischarged, this allows a sloppy reading. That is, when the VP meaning is recovered by an elliptical VP, the pronoun referent will be determined in the elliptical VP context. Alternatively, the def assumption could be discharged first, determining the referent of the pronoun to be "Tom". This leaves the following object:

$\{<$indef$>\}:\lambda$ y.think(y,smart(Tom))

Now the *indef* assumption is discharged, storing the following object in the discourse model:

$\{\}:\lambda$ y.think(y,smart(Tom))

If this object is recovered by an elliptical VP, it will only allow the strict reading, i.e., "thinks Tom is smart". The strict/sloppy ambiguity arises from the different orders of assumption discharge.[5]

3.4 Sloppy Identity for NP Anaphora

Next, I consider the possibility for sloppy identity for NP anaphora. In principle, sloppy identity is possible whenever there is an object of this form:

$\{<$x,def,. . .$>,<$indef$>\}$:. . .x. . .

The indef assumption will cause the object to be saved in the discourse model. If this is done before the def assumption is discharged, the variable x can be reinterpreted when this object is accessed in the discourse model – if the context has changed, there will be a sloppy reading. If

[4] For simplicity, we ignore the assumption associated with the nested predicate "is smart".

[5] This is similar to syntactic theories such as [Kitagawa, 1991; Lappin, 1992] in which VP copying and NP indexing are freely ordered.

the def assumption is discharged before the indef assumption, there can only be a strict reading. This should apply equally to individuals as well as properties, whenever the antecedent contains an embedded variable. The NP “a salesman who phoned her” might be represented:

{<x,def,female,sing>,<indef>}:y|salesman(y) and phone(y,x)

The embedded variable “her” could give rise to a sloppy reading in the following example:

(8) Susan would be polite to **a salesman who phoned her**.

(9) Mary would hang up on him.

The sloppy reading is “Mary would hang up on a salesman who phoned Mary”. This reading can be derived similarly to the sloppy reading in the VP ellipsis case above. The antecedent for “him” is the stored object:

{<x,def,female,sing>,}:y|salesman(y) and phone(y,x)

This will allow a derivation of the sloppy reading, in which x is interpreted as coreferential with “Mary”.

3.5 Derivation System

The derivation system consists of two main rules: function application and assumption discharge. There is also a “discourse rule” for combining sentences into discourse. Each object is represented semantically as an assumption:sense pair, and syntactically as a bracketed string.

3.5.1 A Function Application Rule

The basic rule of the derivation system is function application, defined as follows:

A(B) = C

This rule is subject to the following conditions:

1. type(sense(A)) = $\alpha \rightarrow \beta$
2. type(sense(B)) = α

3. sense(A) = λx.M

4. sense(C) = M[sense(B)/x]

5. Assm(C) = Assm(A) $\cup$ Assm(B)

6. DM(C) = DM(A) $\cup$ DM(B)

The first four conditions are simple function application. In addition to this, the assumption stores and the discourse models are combined.

It is assumed that the semantic derivation is systematically related to the syntactic structure in a Montagovian fashion, so that the function application rule will be governed by syntactic bracketing in the usual way – thus, the above rule can apply only when there is the following syntactic bracketing: [α β]$_\gamma$, where α, β, and γ are the syntactic counterparts of A,B, and C, respectively.

3.5.2 Assumption Discharge Rules

In addition to function application, the derivation system includes a rule for discharging assumptions. There are two assumption types: indefinite and definite. The discharge of the indefinite assumption adds an object to the discourse model, and the discharge of the definite assumption accesses an object in the discourse model. The indefinite and definite assumptions apply indifferently to individuals and properties.

Indefinite: add object to Discourse Model

Given an assumption-sense pair A:s, and discourse model DM, where <indef> ϵ A, the discharge of <indef> results in A$'$:s, and discourse model DM$'$ such that

- DM$'$ = DM $\cup$ A$'$:s
- A$'$ = A - $\{$<indef>$\}$

The discharge of the *indef* assumption causes the current assumption:sense pair, A$'$:s, to be added to the output discourse model.

Definite: access object in Discourse Model

Given a assumption-sense pair A:s, and discourse model DM, where <x,def,features> ϵ A, the discharge of <x, def,features> results in A′:s′, such that

- the antecedent is the assumption-sense pair B:t, where B:t ϵ DM
- A′ = (A - {<x,def,features>}) $\cup$ B
- s′ = s[t/x]

The sense of the antecedent t is substituted for the parameter x. Also, any assumptions associated with the antecedent, denoted by B, are added to the assumption set.

This rule describes the case in which a definite variable finds its antecedent within the current discourse model. Another possibility is that the variable is captured by an operator that has scope over it. This is described as follows:

Definite: Access Operator as Antecedent

Given a assumption-sense pair A:s, and operator O_i, where O_i has scope over A:s, and <x,def,features> ϵ A, the discharge of <x, def,features> results in A′:s′, such that:

- A′ = (A - {<x,def,features>})
- s′ = s[i/x]

3.5.3 A Discourse Rule

Here I give a "discourse rule": a sentence is combined with preceding discourse only when its assumption set is empty, and the sense of S is simply conjoined with preceding discourse. Also, the objects in the discourse model contributed by S are combined with objects contributed to the discourse model by preceding discourse.

C = A B

This rule is subject to the following conditions:

1. sense(C) = sense(A) AND sense(B)
2. Assm(B) = {}
3. DM(C) = DM(A) ∪ DM(B)

This rule requires that the assumption set of a sentence B must be empty before it is combined with prior discourse (A). The discourse models of the sentence (B) and prior discourse (A) are combined together.

This rule should in principle be assimilated to the function application schema. This could perhaps be accomplished by defining a proposition as a function on subsequent discourse, along the lines of [Chierchia, 1992]. I will not investigate this here, as it involves issues not immediately germane to the current study.

3.5.4 Lexical Items

Lexical items have syntactic and semantic representations which fully define the way they can participate in derivations, based on the simple rule schema given above. Here are the semantic representations (assumption:sense pairs) of sample lexical items:

N

man = {}:λ x.man(x)

ADJ

smart = {}:λ x.smart(x)

NP

he = {<x,def,male & sing>}:x

it = {<x,def,inanimate & sing}:x

e = {<x,def>}:x

VP

walk = {}:λ x.walk(x)

e = {<P,def>}:P

AUX

do = {}:λ P.PRESENT(P)

did = {}:λ P.PAST(P)

would = {}:λ P.POSSIBLE(P)

V

help = {<indef>}:λ y,x. help(x,y)

visit = {<indef>}:λ y,x. visit(x,y)

think = {<indef>}:λ y,x. think(x,y)

3.6 Sample Derivations

I will show how several examples from Chapter Two are derived in this system. Each line of the derivation contains the following information: the syntactic string, syntactic category, the semantic representation (an assumption:sense pair). Each step of the derivation results from either function application or the discharge of an assumption. If an assumption is discharged, this is indicated at the end of the line. Otherwise, the derivation step is function application.

At certain stages of the derivation, the current contents of the discourse model (DM) is displayed, as a set of assumption:sense pairs.

I begin with example (7), which is repeated here:

(10) Tom thinks he is smart. Harry does too.

is smart, VP, {}:λ y.smart(y)

he is smart, S, {<x,def,male&sing>}:smart(x)

thinks he is smart,VP, {<indef>,<x,def,male&sing>}: λ y.think(y,smart(x)) (discharge-indef)

thinks he is smart,VP,{<x,def,male&sing>}:λ y.think(y,smart(x)) (discharge-def)

thinks he is smart,VP,{}:λ y.think(y,smart(**Tom**))

Tom thinks he is smart,S,{}:think(Tom,smart(**Tom**))

DM =

{{}:Tom,

{<x,def,male&sing>}:λ y.think(y,smart(x))

. . . }

P, VP, {<P,def>}:P (discharge-def)

does P, VP, {<x,def,male&sing>}:λ y.think(y,smart(x))

Harry does P, S, <x,def,male&sing>}:think(**Harry**,smart(x)) (discharge-def)

Harry does P, S, {}:think(**Harry**,smart(**Harry**))

Thus we derived the sloppy reading for this example, by adding the VP meaning to the discourse model before the pronoun assumption is discharged. When the VP meaning is recovered, the pronoun assumption is discharged with respect to the current state of the discourse model, determining the referent to be "Harry". Of course, this reading can also be derived by the Sag/Williams account, since the "sloppy" variable can be represented as a lambda-bound variable.

Next, I give a derivation of example (3), from Chapter One, repeated here:

(11) If Tom was having trouble in school, I would help him.

(12) If Harry was having trouble, I doubt that I would.

As I showed in Chapter Two, the sloppy reading in this example is not derivable using the lambda abstraction mechanism of the Sag/Williams account, since the sloppy variable is not a lambda-bound variable. In the current approach, the sloppy reading can be derived straightforwardly, as shown below:

help, TV, {<indef>}:λ x,y.help(y,x).

help him, VP, {<indef>, <z,def,male>}:λ y.help(y,z). (discharge-indef)

help him, VP, {<z,def,male>}:λ y.help(y,z). (discharge-def)

help him, VP, {}:λ y.help(y,**Tom**).

I (would) help him, S, help(I,**Tom**)

.

.

If Tom was having trouble in school, I would help him, S, {}:if trouble(**Tom**) then help(I,**Tom**)

DM =
{ {}:Tom,
{<z,def,male>}:λ y.help(y,z)
. . . }

P, VP, {<P,def>}:P
would P, VP, {<P,def>}:P (discharge-def)
would P, VP, {<z,def,male>}:λ y.help(y,z) (discharge-def)
would P, VP, {}:λ y.help(y,**Harry**)
I would P, S, {}:help(I,**Harry**)
.
.
If Harry was having trouble, I doubt that I would, S, {}:if trouble(**Harry**) then help(I, **Harry**)

Next, example (13), repeated from Chapter Two:

(13) China is a country that Harry wants to **visit e**. And he will too, if he gets enough money.

In this derivation, there is a syntactic requirement on the wh-trace in the antecedent, namely, it must be bound by a wh-operator. This requirement is not imposed at the ellipsis site, since no syntactic material is copied.

The antecedent for the elliptical VP is:

{<z,def>}:λ x.visit(x,z)

The derivation proceeds as follows:
e, NP, {<z,def>}:z
visit, V, {<indef>}:λ x,y.visit(x,y)
visit e, VP, {<indef>,<z,def>}:λ x.visit(x,z) (discharge-indef)

wants to visit e, VP, {<z,def>}:λ x.want(x,visit(x,z))

Harry wants to visit e, S, {<z,def>}:.want(**Harry**,visit(**Harry**,z))

that OP_e Harry wants to visit e, REL, {<z,def>}:λ e.want(Harry,visit(**Harry**,z)) (discharge-def)

that OP Harry wants to visit e, REL, {}:λ e.want(Harry,visit(**Harry**,e))

country that OP_e Harry wants to visit e, CNP, {}:λ e.want(Harry,visit(**Harry**,e)) AND country(e)

China is a country that OP_e Harry wants to visit e,S,

{}:want(**Harry**,visit(**Harry**,**China**)) AND country(**China**)

DM =

{ {}:China,

{<z,def>}:λ x.visit(x,z),

. . . }

P, VP, {<P,def>}:P

will P, VP, {<P,def,future>}:P (discharge-def)

will P, VP, {<z,def>}:λ x.visit(x,z) (discharge-def)

will P, VP, {}:λ x.visit(x,**China**)

he will P, S, {<x,def,male AND sing>}:visit(x,**China**) (discharge-def)

he will P, S, {}, {}:visit(**Harry**,**China**)

Finally, we give a derivation for this example, also repeated from Chapter Two:

(14) Martha and Irv wanted to dance with each other, but Martha couldn't, because her husband was there.

Here, the antecedent VP is "dance with each other". This would not be possible at the ellipsis site. The semantic representation of the antecedent, however, is λ x.dance-with(x,y), where the free variable is bound by a quantifier outside the antecedent VP. As discussed in Chapter Two, I will follow recent work on reciprocals ([Heim *et al.*, 1991; Roberts, 1991]) in dividing "each other" into a *distributor* "each" and a variable "other". I will assume, then, the following structure for (14):

[Martha and Irv]$_{NP}$[each [λ x. wanted (x dance with other)]]$_{VP}$

Assuming that "other" represents a free variable (bound outside the VP), the following is the antecedent for the elliptical VP:

{<y,def>}:λ x.dance(x,y)

Given this, we can derive the desired reading by permitting the free variable in the antecedent to be reinterpreted at the ellipsis site. Here is the (abbreviated) derivation:

dance,V, {<indef>}:λ x,y.dance(x,y)

dance with other,VP, {<indef>,<y,def>}:λ x.dance(x,y) (discharge-indef)

.

.

.

Martha and Irv each wanted to dance with other,S,

{}:{Martha,Irv} DIST λx. want(x,dance(x,y))

DM =

{{}:Martha,

{}:Irv,

{<y,def>}:λ x,y.dance(x,y)

. . . }

P, VP, {<P,def>}:P

couldn't P, VP, {<P,def>}:P (discharge-def)

couldn't P, VP, {<y,def>}:λ x.dance(x,y) (discharge-def)

couldn't P, VP, {}:λ x.dance(x,Irv)

Martha couldn't P,S,{}:dance(Martha,Irv)

Thus, given an analysis under which a reciprocal is divided into a "distributor" and a variable, the desired reading is derived without any further stipulations, simply permitting the variable to be reinterpreted at the ellipsis site, as usual.[6]

[6]Lappin (p.c.) suggests that a syntactic reconstruction theory could derive this reading by applying "vehicle change" [Fiengo and May, 1992b], in effect allowing "each other" from the antecedent to change to "the other" at the ellipsis site.

3.7 Extending the System: Split Antecedents

I observed in Chapter Two that VP ellipsis occurrences sometimes have a "split antecedent", as in (94), repeated here in slightly modified form:[7]

(15) Wendy wanted to **sail around the world** and Bruce wanted to **climb Kilimanjaro**, but they couldn't because money was too tight.

The VP ellipsis is resolved by forming a list of properties L:

L = [**sail around the world**, **climb Kilimanjaro**]

The subject of the VP ellipsis, "they", also has a split antecedent, the list of individuals I:

I = [**Wendy**, **Bruce**]

To arrive at the desired reading, L is applied to I according to the following schema, which was described as "Rule 2" in Chapter Two:

[Rule 2:] (Properties) [a,b] [P,Q] $\Rightarrow$ Pa AND Qb

For the example above we have

[**Wendy**,**Bruce**] [**sail around the world**,**climb Kilimanjaro**]

which, by Rule 2 gives:

sail around the world(**Wendy**) AND **climb Kilimanjaro**(**Bruce**)

as desired.

A similar rule was invoked for the following example:

(16) John$_1$ likes Susan$_2$ and Bill$_3$ likes Mary$_4$. They$_{1,3}$ like them$_{2,4}$ because

Here the rule involves two pairs of individuals:

[Rule 2:] (Individuals) [x,y] V [u,v] $\Rightarrow$ V(x,u) AND V(y,v)

For the above example, we get:

[**John Bill**] like [**Susan Mary**] $\Rightarrow$ like (**John**,**Susan**) AND like(**Bill**,**Mary**)

Both of these rules can be eliminated in favor of a general "respectively" operator. The interpretation of "respectively" is defined as follows: Given a proposition P(X,Y) where X and Y are list-denoting expressions, RESP P(X,Y) is interpreted:

[7] I have changed "neither of them can" to "they couldn't" to avoid the issue of the interpretation of "neither". I continue to ignore the negation in the current form of the example.

RESP(P(X,Y)) = $\forall$ x ϵ X. P(x,f(x))

where the function f must be a one-to-one function from X onto Y. The list-denoting expressions X and Y can be either explicit lists or plural anaphoric expressions, either pronouns or proverbs. If X and Y are explicit lists, RESP can be either overt or "invisible". If X and Y are "plural" anaphoric expressions, RESP cannot be overt.

To see how the correct reading is derived with this operator, consider again (15). Resolving the VP ellipsis and "they", we have:

[**Wendy,Bruce**] couldn't [**sail around the world**,**climb Kilimanjaro**]

Now, we add the respectively operator:

RESP [**Wendy,Bruce**] couldn't [**sail around the world**,**climb Kilimanjaro**]

This gives the interpretation

$\forall$ x ϵ [**Wendy,Bruce**]. x couldn't f(x)

Here the function f is $\{<$**Wendy**,**sail around the world**$>$, $<$**Bruce**,**climb Kilimanjaro**$>\}$.

There is one other reading satisfying the definition of RESP, in which the function f is $\{<$**Wendy**,**climb Kilimanjaro**$>$, $<$**Bruce**,**sail around the world**$>\}$. This reading is not available, because this is not a "salient" mapping function.

If X and Y are explicit lists, the mapping function f must follow the surface order of the two lists. If they are anaphoric expressions, there is no surface order constrain f, so f is merely restricted to be a salient mapping.

One might think that, for anaphoric expressions X and Y, f must follow the surface order of the antecedents for X and Y. However, this is not always so, as shown by the following example, also from Chapter Two:

(17) Susan's mother$_1$ likes Susan$_3$, and Mary's mother$_2$ likes Mary$_4$. Their fathers$_{5,6}$ like them$_{3,4}$ too.

Here we have:

$\forall$ x ϵ $\{$**Susan's father**, **Mary's father**$\}$. x likes f(x)

The function f is $\{<$**Susan's father**, **Susan**$>$, $<$**Mary's father**,**Mary**$>\}$. In this case, "their fathers" does not have an explicit antecedent.

3.8 Conclusions

The system described in this chapter implements the proverb theory of VP ellipsis, in which there is a semantic identity condition relating an elliptical VP to its antecedent. The identity condition is stated as a general condition on proforms. The possibility of sloppy identity arises because the discharge of assumptions is freely ordered, so that, whenever an antecedent contains an anaphoric expression, a sloppy interpretation is possible. The system associates an indefinite feature with VP's as well as (some) NP's, and the indefinite feature causes an object to be stored in the discourse model, as a potential antecedent for anaphoric expressions.

Since no mechanisms specific to VP ellipsis are required, this approach is optimal from the standpoint of theoretical economy, and contrasts sharply with the Sag/Williams account, in which a special set of mechanisms are required to allow sloppy identity in VP ellipsis. In addition, this approach is empirically superior, in that it permits examples incorrectly ruled out in alternative approaches. I have presented derivations of several such examples.

Chapter 4

Previous Accounts

4.1 Overview

I divide previous accounts of VP ellipsis into three categories: syntactic reconstruction accounts, lambda abstraction accounts, and semantic accounts. In Chapter Two I gave many examples that contradicted the syntactic reconstruction view; here I will address two primary motivations for syntactic reconstruction accounts. The first is that the elided material is subject to syntactic island constraints, and the second is that binding constraints apply to the elided material. I show that neither of these arguments is convincing.

Next, I discuss lambda abstraction accounts of VP ellipsis: this includes the Sag/Williams account, which has already been discussed in Chapter Two, and the equational approach of [Dalrymple *et al.*, 1991]. The equational approach is a major departure from the Sag/Williams account in that it does not associate strict and sloppy readings with an ambiguity in the antecedent. However, it retains the use of the lambda abstraction mechanism to explain sloppy identity. I briefly describe two important critiques of the lambda abstraction approach: first, [Webber, 1978] gives numerous examples which cannot be handled by the lambda abstraction approach. Secondly, [Partee and Bach, 1984] show that the constraints imposed by the lambda abstraction approach are incompatible with basic notions in model-theoretic semantics, such as compositionality and the need to refer to types rather than tokens in the semantics. I show that, in a sense, these two criticisms of lambda abstraction make the same point – the "free variable" constraint must be removed. That is, sloppy readings must be permitted for variables that are free within the antecedent VP.

Finally, I examine several alternative formulations of semantic identity, and discuss their relation to the proverb theory that I have presented.

4.2 Syntactic Reconstruction

The syntactic reconstruction approach to VP ellipsis [Ristad, 1990; Lappin and McCord, 1990; Kitagawa, 1990; Kitagawa, 1991; Fiengo and May, 1990; Fiengo and May, 1992b; Lappin, 1992; Lappin, 1993] is based on the view that an elided VP is indeed present at some syntactic level of representation, and that the syntactically reconstructed VP is identical to its antecedent. These authors argue that syntactic constraints are imposed on elided material in the same way they would apply to overt material: in particular, it has been argued that binding constraints and island constraints hold of material within the elided VP. To enforce these constraints, these accounts argue that a syntactic copy of the antecedent is present at the ellipsis site. In Chapter Two I presented various examples which would not permit syntactic reconstruction. In this section, I will focus on the arguments for syntactic reconstruction associated with binding theory and island constraints.

With respect to binding constraints, it is clear that they do not apply to elided material in the same way that they apply to overt material. One example of this involves strict readings for reflexives, as discussed in Chapter Two. Similarly, it is well known that binding constraints on names are not enforced under ellipsis. An example of this is the following:

(1) Harry got to Sue's apartment before she did. [Dalrymple, 1991]

The overt counterpart of this sentence would constitute a Principle C violation, while this violation is not observed in the elided version. Since the syntactic binding constraints are not enforced equivalently under ellipsis, a syntactic reconstruction approach must relax the identity condition in some way. To deal with examples such as (1), [Fiengo and May, 1992b] define an operation called "vehicle change", which allows certain differences between antecedent and elided material. One such change is the conversion of names to pronouns. Applied to (1), this gives:

(2) Harry got to Sue's apartment before she did [got to *her* apartment].

With a pronoun substituted for the name, there is no longer a binding theory violation. Similarly, strict readings for reflexives, which violate Principle A, are allowed by Fiengo and May as follows. Consider this example:

(3) John defended himself better than his lawyer could have.

Fiengo and May argue that a reflexive like "himself" must be divided into a pronominal and an anaphoric part, and that, under ellipsis, the pronominal part can be copied without the anaphoric part. In effect, this allows the conversion of reflexives to pronouns under ellipsis, just as vehicle change allowed the conversion of names to pronouns. This has the effect of making Principles A and C unobservable under ellipsis, while Principle B should still apply. However, Principle B is also violated under ellipsis, as shown by the following naturally-occurring example:

(4) Why do you **want him$_i$ to play chess**?

(5) I don't. He$_i$ does! (From film preview)

Here, a syntactic copy of the antecedent at the ellipsis site results in a condition B violation:

(6) He$_i$ does want him$_i$ to play chess.

In sum, all the binding theory principles can be violated in appropriate contexts, as predicted by the proverb theory. The syntactic theory of Fiengo and May requires stipulations whose sole motivation appears to be that they render Principles A and C unobservable under ellipsis; the Principle B violations remain unexplained.

The second motivation for syntactic reconstruction has to do with island constraints that are enforced under ellipsis. The following examples were mentioned in Chapter Two:

(7) John read everything which Bill did.

(8) * John read everything which Bill believes the claim that he did.

It has been claimed that a wh-trace appears within the elided VP, and furthermore, that the trace is subject to island constraints. [Chao, 1987] has suggested that VP ellipsis must be syntactically reconstructed in cases where wh-binding takes place within the ellipsis site. However, as I suggested in Chapter Two, it is equally possible to treat these cases as "pseudo-gapping", so that the trace is present and only the main verb is elided. In fact, I argued that the pseudo-gapping analysis of these sentences is superior to the VP ellipsis analysis. Thus these examples do not require that VP ellipsis involves syntactic reconstruction, since they are not properly analyzed as VP ellipsis.

The two major arguments for syntactic reconstruction are based on binding effects and island constraints. We have seen that they are not compelling. The syntactic reconstruction approaches are similar to the proverb theory in that, in general, they permit sloppy readings for any embedded variable in the antecedent, subject only to independent parallelism constraints. For example, they do not restrict sloppy readings to subject covariance, as is the case in lambda abstraction approaches. [Fiengo and May, 1992b] define a syntactic parallelism constraint to constrain sloppy readings. In Chapter Seven, I suggest that this parallelism constraint cannot be defined syntactically. The proverb theory treats sloppy identity for NP anaphora and VP ellipsis in a uniform fashion. This is not possible for a syntactic reconstruction theory, unless pronoun resolution is treated as syntactic reconstruction as well.

4.3 Lambda Abstraction Approaches

4.3.1 The Sag/Williams Account

The Sag/Williams account is developed originally in [Sag, 1976] and [Williams, 1977]. A similar approach is taken in [Keenan, 1971]. I have described this approach in Chapter Two, and I will briefly recapitulate its main points here.

The Sag/Williams account is based on two rules: the Derived Verb Phrase rule [Partee, 1975], which defines a lambda representation of a VP, and the Pronoun Rule, allowing a pronoun to be represented as a lambda-bound variable. This approach allows an appealing solution to the problem of sloppy identity, in examples such as:

(9) John$_i$ thinks he$_i$ is smart.

(10) Harry does too.

As we saw in Chapter Two, there are two readings: Harry could think either John or Harry is smart. According to the Sag/Williams account, the sentence can be represented as:

(11) John$_i$, λ x.x thinks he$_i$ is smart

The "pronoun rule" permits the pronoun to be replaced by a lambda-bound variable, resulting in:

(12) John$_i$, λ x.x thinks x is smart

These two possible antecedent representations determine the possible readings: (11) gives a strict reading, and (12) gives a sloppy reading.

A key point about this approach is that sloppy identity in VP ellipsis is explained in terms of an ambiguity in the representation of the antecedent VP. A subject-coreferent pronoun can be represented as either a bound or free variable. If it is represented as a bound variable, a sloppy reading results from the reapplication of the lambda expression, otherwise, we get a strict reading. While this account seems to provide an elegant solution to the problem of sloppy identity, [Dalrymple *et al.*, 1991] demonstrate that it must be fundamentally revised.

4.3.2 The Equational Approach

[Dalrymple *et al.*, 1991] point to examples such as the following (due originally to [Dahl, 1973]:[1]

(13) John thinks he is a fool. Harry does too, although his wife doesn't.

Consider the reading in which Harry thinks Harry is a fool, but Harry's wife doesn't think Harry is a fool. This requires that the first elliptical VP gets a sloppy reading, and the second elliptical VP gets a strict reading. But if the possibility of a sloppy reading is associated with an ambiguity in the representation of the pronoun "he" in the antecedent, this reading would be impossible. To generate the sloppy reading for the first ellipsis, the VP must be represented:

λ x.x thinks x is a fool

That is, the pronoun "he" must be represented as a lambda bound variable. However, this will also give rise to a sloppy reading for the second ellipsis – there is no way to get a strict reading for the second ellipsis, and still maintain the Sag/Williams account's identity condition on lambda representations of VP's. A similar criticism can be made of any account in which the strict/sloppy ambiguity reflects an ambiguity in the representation of the antecedent clause, such as [Reinhart, 1983; Fiengo and May, 1992b; Higginbotham, 1992].[2] In the equational approach of [Dalrymple

[1] A similar point can be made based on an example due to [Gawron and Peters, 1990]: "John revised his paper before the teacher did, and Bill did too."

[2] [Fiengo and May, 1992b] suggest a solution to this problem within their approach, which I will not address here.

et al., 1991] , the problem of cascaded ellipsis is solved, while retaining the lambda abstraction mechanism of the Sag/Williams account.

The equational approach resolves ellipsis by solving an equation for a variable which represents the elided material.[3] Consider the following example:

(14) John **walked**. Tom did too.

To resolve the ellipsis, the following equation must be solved:

(15) P(John) = walked(John)

That is: what property P *could have been* applied to "John" to result in the complete sentence, **walked(John)**? Solutions to the equation represent possible antecedents for the elliptical VP. In this example, with appropriate restrictions on the form of solutions, there is only one solution:

λ x.walk(x)

More generally, Dalrymple, Shieber and Pereira state the ellipsis problem as one of finding solutions to the equation

$P(s_1) = s$

where s_1 is the subject of the antecedent clause and s is the complete antecedent clause.[4]

Just as in the Sag/Williams account, the possibility of sloppy identity is explained by the lambda abstraction mechanism; a pronoun can get a sloppy reading only if it is represented as a lambda-bound variable.

The fundamental difference between the equational approach and the Sag/Williams account is that the equational approach is not an "identity" theory, in that it does not require that the elided VP be identical to any semantic or syntactic object appearing in surrounding discourse. Above, I stated the ellipsis resolution question for the equational approach as follows:

[3]This approach is intended to apply to a variety of ellipsis phenomena, including gapping, stripping, and sluicing, in addition to VP ellipsis. However, I will only discuss its application to VP ellipsis. In Chapter Seven, I suggest that these other ellipsis phenomena are essentially different from VP ellipsis.

[4]In the fact, the general ellipsis resolution equation is stated more generally: $P(s_1,s_2,\ldots,s_n) = s$, where P is an n-place relation. Typically, for cases of VP ellipsis, P will be a one place relation, i.e., a property, but there is no requirement in the equational approach that this be so.

What property P *could have been* applied to the parallel element(s) to result in the antecedent clause?

The "identity of relations" approach, exemplified by the Sag/Williams account, poses the question:

What property P *was* applied to the parallel element to result in the antecedent clause?

Now we consider again the case of cascaded ellipsis. The example is repeated here:

(16) John **thinks he's a fool**. Harry does too, although his wife doesn't.

We saw that there is a reading of this example which cannot be generated in the Sag/Williams account, namely, where John thinks John is a fool, Harry thinks Harry is a fool, and Harry's wife doesn't think Harry is a fool. This problem does not arise in the equational approach, because the choice of strict or sloppy reading has nothing to do with the actual representation of the VP in the antecedent clause. Rather, there is a separate choice made in solving the equation for each elliptical occurrence. The first ellipsis is resolved by solving the following equation:

P(John) = John thinks John is a fool

This has two solutions, corresponding to the strict or sloppy readings, just as in the Sag/Williams account, the VP antecedent has two logical form representations. Assume the sloppy reading is selected, so that the value for P is as follows:

(17) P = λ x.x thinks x is a fool

Then the second ellipsis must be resolved, by solving the equation for its antecedent clause, namely:

Q(Bill) = Bill thinks Bill is a fool

Again there are two possibilities, and the choice is completely independent of the choice made for the previous ellipsis. So there is nothing preventing the choice of a strict reading. Thus we select the following value for Q:

(18) Q = λ x.x thinks Bill is a fool

The case of cascaded ellipsis is a primary argument for the equational approach; in particular, it is motivates the rejection of the "identity of relations" view, that is, the view that VP ellipsis is governed by an identity condition. However, according to the proverb theory, where VP ellipsis is governed by an identity of *dynamic* properties, cascaded ellipsis presents no problem. As described in Chapter Three, the proverb theory permits pronouns within the antecedent VP to be resolved independently in the antecedent context and the VP ellipsis context.

In Chapter Two, I pointed out two major problems for the Sag/Williams account that are associated with the lambda abstraction account of sloppy identity; these are also problems for the equational approach. The first problem is that sloppy readings are possible for non-lambda bound pronouns, in examples such as the following, repeated from Chapter Two:

(19) If Tom was having trouble in school, I would **help him**.
On the other hand, if Harry was having trouble, I doubt that I would.

Here the pronoun "him" in the antecedent gets a sloppy reading under ellipsis. But it cannot be replaced by a lambda-bound variable, either in the Sag/Williams account or the equational account[5]. The second problem for the lambda abstraction approach is that sloppy readings are possible with pronominal anaphora as well, in cases such as:

(20) Smith makes **his children** go to bed at 8 every night.

(21) Jones lets them stay up as late as they want.

In resolving the pronoun "them", the pronoun "his" within the antecedent "his children" can be interpreted in a strict or sloppy fashion. The lambda abstraction approach provides no explanation for this.

The equational approach represents a major advance over the Sag/Williams account, because it rejects the notion that the strict/sloppy ambiguity in ellipsis is associated with an ambiguity in the antecedent. However, the equational account shares with the Sag/Williams account the empirical shortcomings of the lambda abstraction approach to sloppy identity. Next, I consider two critiques of lambda abstraction approaches.

[5]This assumes that the parallel elements are the subjects of the antecedent VP and the elliptical VP, although this is not required in the equational approach. If "Tom" and "Harry" could be treated as parallel elements, it would be possible for the equational approach to permit this reading, since "him" could then be represented as a lambda-bound variable. However, this would represent a radical departure for the equational approach, since the solution to the equation would no longer represent merely the elided material.

4.3.3 Two Critiques of Lambda Abstraction Approaches

Webber [Webber, 1978] gives several examples which are problematic for lambda abstraction approaches. Many of these examples have been discussed in Chapters Two and Three; I repeat one here:

(22) China is a country that Joe wants to **visit e**, and he will too, if he gets enough money. [Webber, 1978]

Webber argues that examples of this sort show that VP ellipsis cannot be described as syntactic identity, and suggests that the resolution mechanism sometimes involves inference. This example violates the "free variable" constraint of the above lambda abstraction accounts; this is the constraint that requires that a variable bound by an operator outside the antecedent VP must be bound by the same token operator at the ellipsis site. This constraint is the focus of the second critique of lambda abstraction: that of [Partee and Bach, 1984]. Here, it is argued that the "free variable" constraint is incompatible with basic notions in model-theoretic semantics, such as compositionality and the need to refer to types rather than tokens in the semantics. Partee and Bach leave this as an open problem, because they believe that this constraint is empirically necessary. Webber's examples show that it in fact must be removed. The proverb theory, by removing the free variable constraint, makes it possible to derive Webber's examples, and answers the theoretical criticisms raised by Partee and Bach.

4.3.4 The Lambda Abstraction Approach and Semantic Types

There is a more general problem that arises with the Equational Approach, or indeed, any approach in which sloppy identity is explained by lambda abstraction. This problem has to do with the mapping of syntactic categories to semantic types. The Equational Approach does not permit a regular correlation between syntactic categories and semantic types. The elided VP need not be a property in the Equational approach. Rather, the Equational Approach countenances an unboundedly large set of possible types for elided VP's.

To see this, consider the Equational analysis of the following example [Dalrymple *et al.*, 1991, pages 20-21]:

(23) The person who introduced Mary to John would not give her his phone number, nor would the person who introduced Sue to Bill. [Wescoat, 1989]

[Dalrymple *et al.*, 1991] present the following equation for this example:

P(pwi(m,j),m,j) = refuse(pwi(m,j),give(m,phone(j)))

Solving for P results in the following semantic object:

λ x. λ y.λ z.refuse(x,give(y,phone(z)))

Here, the missing object corresponding to the elided VP is a three-place relation, that is, an object of type $< e, < e, < e, t >>>$.[6] This is because, as in any lambda abstraction approach, each "sloppy" variable must be lambda-abstracted. Since there appears to be no bound on the number of "sloppy" variables in VP ellipsis, the Equational Approach must permit an unbounded number of semantic types for VP ellipsis. The same problem would arise with any generalization of the lambda abstraction approach.

4.4 Semantic Approaches

In the introduction to [Sag, 1976], Sag acknowledges that the observations of Webber and others indicate that a semantic identity account is required. Sag develops such an account in [Sag, 1981][7]. This treatment is broadly consistent with the spirit of the proverb theory. However, [Sag, 1981] does not reject the free variable constraint; rather, he introduces some additional complexity into the model-theoretic denotations of his system, for the purpose of capturing the constraint. I have argued that this constraint must be removed on empirical grounds, thus the semantic complications introduced by Sag are unnecessary. [Lappin, 1984] also suggests a semantic approach that modifies, without rejecting, the free variable constraint.

Another work with much in common with the proverb theory is [Rooth, 1981]. Rooth develops a Montague grammar approach, using the general framework of [Bach and Partee, 1980]. Rooth extends this framework by adding a discourse rule and treats VP ellipsis on an analogy with the quantifying-in rule for NP's. As in the present work, a VP occurrence stores a VP meaning, which

[6] I am ignoring intensionality in these semantic types.

[7] In addition, there are sketches of semantic approaches in [Fodor and Sag, 1982] and [Sag and Hankamer, 1984].

can be accessed to resolve an occurrence of VP ellipsis. Also, Rooth's treatment shares with the proverb theory the view that the elliptical VP, as a proform, has no internal syntactic structure. Rooth does not address the "free variable" constraint. Rooth finds the proform view of VP ellipsis problematic, because he believes there are cases of wh-binding within the elliptical VP. I have argued in Chapter Two that such cases are in fact pseudogapping rather than VP ellipsis. Thus they do not present a problem for the proverb theory.

Chapter 5

Computational Implementation

5.1 Introduction

This chapter describes two computer implementations, representing solutions to the two questions identified at the begin of this dissertation: the identity question and the location question. The implementation of a solution to the identity question is based on a system of semantic derivation developed by [Pereira and Pollack, 1991]. I extend this system to allow VP meanings to be stored in the discourse model, and to allow elliptical VP's to access those VP meanings. The resulting system is a partial computational realization of the mechanisms formally described in Chapter Three. Next I address the location question. I describe an algorithm for locating VP ellipsis antecedents, and I describe a computer program that partially implements this algorithm.

5.2 The Identity Question: Retrieving VP Meanings

In this section I describe a computer program which resolves VP ellipsis by retrieving VP meanings. I begin by describing a general computational system of semantic derivation, and then show how this system is extended to resolve VP ellipsis by retrieving the meaning of the antecedent.

As I stated in Chapter Three, the two central tenets of the proverb theory are the following:

1. Meanings are stored in the discourse model.

2. Anaphoric expressions do not have their antecedents predetermined–they are determined at some stage during the derivation.

The Candide system follows these tenets; indeed the approach in Chapter Three was in part inspired by the Candide system. As in Chapter Three, the Candide system represents meanings as assumption:sense pairs. However, Candide does not follow a necessary consequence of these tenets: it does not store assumption:sense pairs in the discourse model. In other words, context-dependent objects are not stored in the discourse model. I will modify the Candide system to store assumption:sense pairs in the discourse model. In addition, I will define rules for saving VP meanings in the discourse model (as assumption:sense pairs), and for recovering VP meanings from the discourse model.

5.2.1 Background: The Pereira and Pollack System

Pereira and Pollack's Candide system is implemented in Quintus Prolog. The system interprets syntactic expressions with respect to a discourse model. That is, the system defines a relation R(P, I, DM_{in}, DM_{out}), where *P* is an input syntactic parse tree, *I* is the output semantic interpretation of *P*, and DM_{in} and DM_{out} are the input and output discourse models.

The input syntactic parse tree is represented in a feature-structure notation. For example, the sentence "Arthur worked." is represented:

```
[issue = [predicate = [wordstem = work,
                       cat = verb],
          arg1 = [wordstem = arthur,
                  cat = propernoun,
                  role = subject]],
 type = declarative,
 polarity = positive,
 tense = past,
 tenseaspect = simple,
 progressive = no]
```

Candide interprets this syntactic representation by recursive calls to the predicate **interpret**. The clause for the interpretation of a tensed clause is:

```
interpret1(va(Pred,Args),_Fs,Int,Assms,Context,TContext0,TContext):-
      interpret_va_args(Args,Int1,Assms1,1,Context,TContext0,TContext1),
      interpret(Pred,Int1,Assms2,Context,TContext1,TContext2),
      append(Assms1,Assms2,Assms3),
      discharge(Int1,Assms3,TContext2,Context,Int,Assms,TContext).
```

I describe the argument list first: the first argument: **va(Pred,Args)** denotes an active verb **Pred** together with a list of arguments **Args**. (The subject is the first element of the list.) The variable **_Fs** can be ignored. **Assms** is the output assumption set. **Context** is the global discourse context, with **TContext0** the input discourse model, and **TContext** the output discourse model.

The body of the clause begins with a call of **interpret_va_args**, in which the arguments of the predicate (subject, object, etc.) are interpreted. Then the predicate itself (**Pred**) is interpreted. Any undischarged assumptions from these interpretations are appended together. Finally, these assumptions are discharged. Observe the "threading" of the **Tcontext** variables that represent current local discourse context. The output **Tcontext** of each predicate becomes the input **Tcontext** of the subsequent predicate. For example, in the call to **interpret_va_args**, the output discourse model is **TContext1**. This is the input discourse model for the call to **interpret** in the next line.

Consider now the interpretation of an indefinite NP; this will cause an individual to be added to the output discourse model.

```
interpret1(indefnp,Fs,Int,Assms,Context,TContext0,TContext) :-
access(arg1,Fs,Nom),
interpret(Nom,Int1,Assms1,Context,TContext0,TContext1),
nom_mod_complete(Assms1),
access(role,Fs,Pos),
gensym(r,Var),
discharge(Var,[bind(indef,Var,Int1,[pos(Pos)])|Assms1],TContext1,
    Context,Int,Assms,TContext).
```

This clause produces a new parameter by a call to **gensym**. Then an **indef** assumption for this parameter is discharged, adding an individual to the output discourse model, **TContext**.

Consider now the interpretation of a pronoun:

```
interpret1(pronoun,Fs,Var,[bind(pronoun,Var,Type,[pos(Pos)])],_Context,
  TContext,TContext) :-
     access(pntype,Fs,PnType),
     make_type(PnType,Type),
     access(role,Fs,Pos),
     gensym(r,Var).
```

The effect of this clause is to create a new parameter for the pronoun, together with an assumption constraining the eventual referent for the pronoun. We consider first the argument list: the first argument, **pronoun**, denotes simply that this is the clause for interpreting a pronoun. The second argument **Fs** is the syntactic (feature-structure) representation of the pronoun. This is the primary input argument to this clause. The next two arguments are the output arguments. First, **Var**, is a newly created variable representing the pronoun parameter. This is followed by the assumption list, which consists of pronoun assumption associated with this parameter. The assumption contains **Var**, the parameter, **Type**, which is number and gender, and **Pos**, the syntactic position of the pronoun.

In the body of the clause, the feature structure representation of the pronoun **Fs** is accessed to determine the semantic type, **Type**, and the syntactic position, **Pos**. Then **gensym** generates a new variable name for the pronoun.

As an example, a pronoun “he”, appearing in object position, is interpreted by this clause as follows: the **Type** is determined to be masculine, and a new variable $r1$ is created. This causes the following assumption to be added to the assumption set: **bind(pronoun,r1,male,pos(object))**. To determine the referent of the pronoun, the assumption must subsequently be discharged.

The **eliminate** predicate discharges assumptions, with a different clause for each assumption type. The following is the clause for the discharge of a pronoun assumption:

```
eliminate(pronoun,Int0,Assms0,TContext0,Context,Int,Assms,TContext,
```

```
        Current):-
    assm(Current,var,Var),
    assm(Current,type,Type),
    assm(Current,pos,Pos),
    resolve_pn(Type,Pos,Context,TContext0,Assms0,Ent,TContext,Var),
    unnest_poss_assm(Assms0,Current,Assms1),
    replace_sub(Int0,Var,Ent,Int),
    replace_sub(Assms1,Var,Ent,Assms).
```

The input interpretation is **Int0**: this is an arbitrary semantic representation containing a pronoun parameter whose assumption is to be discharged. This assumption is the variable **Current**. This assumption is examined to determine the parameter, **Var**, the pronoun type, **Type**, and its syntactic position, **Pos**. With this information, the predicate **resolve_pn** is called; this predicate selects an appropriate entity currently stored in the discourse model, based on various constraints and heuristics. The output interpretation **Int** is then formed from the input interpretation **Int0** by substituting **Ent** for **Var**; that is, the pronoun parameter is replaced by a discourse entity.

To summarize: the Candide system interprets syntactic representations in the context of an input discourse model, producing a semantic representation and an output discourse model. We have seen that a pronoun occurrence is interpreted by adding a pronoun assumption to the discourse model, which is later discharged, selecting an individual in the discourse model. Next, I will show how this system is extended to treat VP ellipsis in an analogous manner.

5.2.2 Extensions for VP Ellipsis

There are two extensions necessary for VP ellipsis: first, a rule is needed for adding VP meanings to the discourse model. Second, we need a rule for recovering a VP meaning to resolve a VP ellipsis. There is also a more general change to the model: we must allow objects with undischarged assumptions to be stored in the discourse model. This is necessary for sloppy identity.

Adding VP Meanings to the Discourse Model

As part of the interpretation of a VP, a semantic representation of the VP must be stored in the discourse model. This requires modifications to the **interpret** clause given above for a verb and

its arguments. The modified clause is as follows (modifications are within the dotted lines):

```
interpret1(va(Pred,Args),_Fs,Int,Assms,Context,TContext0,TContext):-
        interpret_va_args(Args,Int1,Assms1,1,Context,TContext0,TContext1),
        interpret(Pred,Int1,Assms2,Context,TContext1,TContext2),
        append(Assms1,Assms2,Assms3),
        % ---------------------------------------------- %
        %  add Int, assms for VP with same variable name
        gensym(p,Pvar),
        %  fix pred interpretation: get right argument vars
        fix_args(Int1,Inttosave,Assms3),
        %  get rid of syntactic assms, unmarked subj assm
        filter_assms(Assms3,SavedAssms),
        append([bind(pred,Pvar,Inttosave),bind(assms,Pvar,SavedAssms)],
               Assms3, Assms4),
        % ---------------------------------------------- %
        discharge(Int1,Assms4,TContext2,Context,Int,Assms,TContext).
```

The argument list of the clause is left unchanged, as are the first three lines of the clause, in which the arguments of the verb are interpreted, after which the verb is interpreted, and the assumption sets appended together. The first modification is a call to **gensym**, which creates a new variable for the VP meaning. For technical reasons which need not concern us here, there is a call to **fix_args**, performing some minor modifications to the form of the VP meaning, **Int1**, producing **Inttosave**. Similarly, a call to **filter_assms** performs technical modifications to the current assumption set, **Assms3**, resulting in **SavedAssms**. Assumptions containing **Inttosave** and **SavedAssms** are added to the assumption list. **Inttosave** represents the interpretation of the VP that is to be saved in the discourse model; **SavedAssms** represents the assumptions associated with that VP. When these assumptions are discharged, the assumptions and sense associated with the VP will be stored in the discourse model. Consider the following example:

(1) He killed him.

A call to the above **interpret** clause will produce:

Inttosave: **kill(a1,r1)**

SavedAssms: {**bind(pronoun, r1, male)**}.[1]

This assumption:sense pair will later be stored in the discourse model.

The following is a sample run of the program with this modification.

```
ARTHUR WORKED
*****Logical Form:
   declarative(work(e1))

*****Discourse segment context:
   assumptions(p1, [])
   property(p1, work(a1))
   entity(e1, t2 \ male \ name_of(t2,arthur))
```

In addition to the entity e1, representing "Arthur", there is a property p1 stored in the discourse model, representing "work(a1)"[2]. Also, the assumptions associated with p1 are stored: in this case the assumption list is empty.

To see a case where a stored assumption set is non-empty, consider the discourse:

(2) Arthur hit Lance. He killed him.

```
ARTHUR HIT LANCE
*****Logical Form:
   declarative(hit(e1,e4))

*****Discourse segment context:
```

[1] This is a slightly simplified representation.

[2] The subject position is represented by the parameter a1, which is implicitly lambda-abstracted. Thus this representation can be thought of as an abbreviation for λ a1.work(a1).

```
   assumptions(p5, []))
   property(p5, hit(a5,e4))
   entity(e4, t15 \ male \ name_of(t15,lance))
   entity(e1, t13 \ male \ name_of(t13,arthur))

HE KILLED HIM
*****Logical Form:
   declarative(kill(e1,e4))

*****Discourse segment context:
   assumptions(p6, .(bind(pronoun, r11, t17 \ male \ t, .(pos(object), [])))
   property(p6, kill(a6,r11))
   entity(e1, t13 \ male \ name_of(t13,arthur))
   entity(e4, t15 \ male \ name_of(t15,lance))
   assumptions(p5, .(bind(unmarked, b4, y \ animal \ t, []), []))
   property(p5, hit(a5,e4))
```

Property p6 is associated with the VP “killed him”. The assumption set for p6 contains the pronoun assumption associated with the parameter r11. If property p6 is the antecedent for a subsequent VP ellipsis, the pronoun assumption would be reinterpreted in the context of the VP ellipsis. I now turn to the recovery of VP meanings.

Recovering VP Meanings from Discourse Model

First, I define a new category of expression: the *proverb*, associated with auxiliary verbs such as “do”. Below is a clause defining the interpretation for proverb. This clause is like the ordinary clause for the interpretation of a verb and its arguments, except that there is only one argument – the subject, and the verb is a proform that is interpreted like a pronoun is.

```
interpret1(pva(Pred,Arg),_Fs,‘Int,Assms,Context,TContext0,TContext):-
   fs_type(Pred,proverb),
   resolve_pv((SavedProp),Pos,Context,TContext0,Assms,Ent,TContext,Var,
```

```
                SavedAssms),
    interpret_pva_subj(Arg,Int1,[Assms1],1,Context,TContext0,TContext1),
    replacesubj(SavedProp,Int1,Int2),
    discharge(Int2,[Assms1|SavedAssms],TContext1,Context,Int,Assms,
              TContext).
```

The predicate **resolve_pv** selects a property and assumption set from the discourse context.[3] The selected property is denoted by **SavedProp**, and the stored assumptions, **SavedAssms**. The subject is interpreted by **interpret_pva_subj**. Then the subject is added to the **SavedProp** representation by **replace_subj**. Finally **discharge** is called to discharge assumptions. The **SavedAssms** are simply combined with any currently undischarged assumptions, denoted by **Assms1**.

Here is a sample execution, with the following discourse:

(3) Arthur worked. Gwen did.

```
  Next? 24.
ARTHUR WORKED
*****Logical Form:
   declarative(work(e1))

*****Discourse segment context:
   assumptions(p6, [])
   property(p6, work(a6))
   entity(e1, t18 \ male \ name_of(t18,arthur))

Next? 305.
GWEN DID
*****Logical Form:
   declarative(work(e3))
```

[3] Generally, **resolve_pv** simply selects the most recent VP meaning. In the following section, I describe an algorithm for selecting the antecedent.

```
*****Discourse segment context:
   entity(e3, t21 \ female \ name_of(t21,gwen))
   assumptions(p6, [])
   property(p6, work(a6))
   entity(e1, t18 \ male \ name_of(t18,arthur))
```

The proverb accesses the property p6, **work(a6)**.[4] The entity $e3$ is substituted in the subject position, giving the correct logical form, **work(e3)**.

Consider now the following discourse:

(4) Arthur worked and Gwen helped him.

(5) Lance worked and she didn't.

Consider the reading of the elliptical VP in which Gwen didn't help Lance; that is, where the antecedent VP is "helped him". This reading is not derivable in the Sag/Williams approach to VP ellipsis, since the "sloppy" pronoun "him" is not subject-coreferent.[5] This reading is derived in the following run of the program:

```
Next? 312.
ARTHUR WORKED AND GWEN HELPED HIM
*****Logical Form:
   declarative(and(work(e1), help(e3,e1)))

*****Discourse segment context:
   assumptions(p28, (bind(pronoun, r48, t69 \ male \ t, .(pos(other), [])))))
   property(p28, help(a28,r48))
   assumptions(p26, [])
   property(p26, work(a26))
   entity(e3, t68 \ female \ name_of(t68,gwen))
```

[4] Recall that there is an implicit lambda abstraction over a6, the subject.

[5] This example is similar to example (3) in Chapter Two; that example cannot be handled within the syntactic and lexical limitations of the Candide system.

```
   entity(e1, t66 \ male \ name_of(t66,arthur))

Next? 313.
LANCE WORKED AND SHE DIDN'T
*****Logical Form:
   declarative(and(work(e4), help(e3,e4)))

*****Discourse segment context:
   entity(e3, t74 \ female \ name_of(t74,gwen))
   assumptions(p29, [])
   property(p29, work(a29))
   entity(e4, t71 \ male \ name_of(t71,lance))
   assumptions(p28, .(bind(pronoun, r48, t69 \ male \ t, .(pos(other), [])),
    .(bind(unmarked, b18, y \ person \ t, []), [])))
   property(p28, help(a28,r48))
   entity(e3, t68 \ female \ name_of(t68,gwen))
   assumptions(p26, [])
   property(p26, work(a26))
   entity(e1, t66 \ male \ name_of(t66,arthur))
```

After the first sentence, the "Discourse Segment Context" includes two VP meanings, p28 and p26, with their associated assumptions. In addition, two entities, e1 and e3 are represented. In the interpretation of the second sentence, the elliptical VP accesses property p28 as its antecedent.[6] Associated with p28 is an assumption set, containing an assumption for the parameter r48, which is in the object position of "help" in property p28. Once the elliptical VP is resolved, the assumption for r48 is discharged, selecting "Lance", (entity e4) as the antecedent. This results in the logical form:

declarative(and(work(e4), help(e3,e4)))

[6]This is not the only possible reading; the property "work" could also be selected.

That is, “Lance worked and Gwen helped Lance”[7].

Thus the system can derive readings not permitted by lambda abstraction approaches such as the Sag/Williams account or the equational account [Dalrymple *et al.*, 1991].

5.3 The Location Question: Determining the Antecedent VP

In this section, I describe an algorithm for selecting the antecedent for VP ellipsis occurrences. The algorithm operates on VP meanings stored in the discourse model, in two stages. First, any impossible antecedents are eliminated. Next, preference levels are determined for remaining possible antecedents. The antecedent with the highest preference level is selected.[8]

I describe some possible extensions to the algorithm to deal with examples which do not have a VP antecedent: that is, cases of split antecedents, and “missing antecedents” described in Chapter Two. Next, I describe a computer program that implements a somewhat simplified version of the basic algorithm.

5.3.1 Structure of the Algorithm

The algorithm has access to VP’s stored in the discourse model, in addition to the syntactic representation of the current sentence, i.e., the sentence containing the elliptical VP. I will assume that the antecedents for NP’s have been determined, although I will not give an algorithm for this.

The VP antecedent selection function is:

```
A-Select(VPlist,VPE)
  VPlist := remove-impossible(VPlist,VPE)
  VPlist := assign-levels(VPlist,VPE)
  antecedent := select-highest(VPlist,VPE)
```

[7] Note that negation is currently ignored by the system.

[8] Similar approaches, combining structural constraints and attentional/pragmatic preferences, have been proposed for the resolution of pronoun occurrences, for example, [Guenthner and Lehmann, 1983; Brennan *et al.*, 1987; Lappin and Leass, 1991].

First, impossible antecedents are removed from the VPlist.[9] Then, the remaining items in VPlist are assigned preference levels, and the item with the highest preference level is selected as the antecedent. If there is more than one item with the same preference level, the item closest to the VPE, scanning left from the VPE, is selected.

The definition of the function **remove-impossible** is as follows:

```
remove-impossible(VPlist,VPE)
 For all v in VPlist
   if ACD(v,VPE) or
      BE-DO-conflict(v,VPE)
   then remove(v, VPlist)
```

There are two types of impossible antecedents: the first involves certain antecedent-containment structures, and the second involves cases in which the antecedent contains a BE-form and the VP ellipsis occurrence contains a DO-form. These are described in detail below.

Next, preference levels are assigned to remaining items in VPlist by the **assign-levels** function. (All items on VPlist are initialized with a level of 0.)

```
assign-levels(VPlist,VPE)
For all v in VPlist
  if related-clause(v,VPE) then
    v.level := v.level + 1
  if coref-subj(v,VPE) then
    v.level := v.level + 1
```

An antecedent is preferred if there is a clausal relationship between its clause and the VPE clause, or if the antecedent and the VPE have coreferential subjects. The determination of these preferences is described in detail below.

[9] VPlist denotes the properties stored in the discourse model. I have, somewhat arbitrarily, limited VPlist to properties appearing no more than two sentences before the elliptical VP.

Finally, the **select-highest** function merely selects the item on VPlist with the highest preference level. If there is more than one item with the highest preference level, the item nearest to the VPE (scanning left) is selected.

5.3.2 Impossible Antecedents

There are two types of impossible antecedents: the first involves certain types of containing antecedents, and the second involves cases in which the antecedent contains a BE-form and the target contains a DO-form.

Containing Antecedents

In general, an VPE cannot have an antecedent that contains it. Thus, in the following configuration, the VP headed by V is an impossible antecedent.

(6) [V [VPE] $]_{VP}$

For example:

(7) I [suppose [I [did]$_{VPE}]_S$ $]_{VP}$.

The VP headed by "suppose" is not a possible antecedent for the VPE. There is one apparent exception to this rule, when VPE is within an NP argument to V. This illustrated by the following example:

(8) Sandy [hit [everyone that Bill [did].$_{VPE}]_{NP}]_{VP}$.
[Sag, 1976]

Here, the antecedent does appear to be the VP headed by "hit". [10]

Be/Do Conflicts

If there is a "be-form" in the antecedent VP, with a "do-form" VPE, VP is not a possible antecedent. This is illustrated by the following example:

[10] In fact, such examples can be analyzed as pseudogapping rather than VP ellipsis, as suggested in [Lappin and McCord, 1990]. Thus, these examples do not in fact involve antecedent containment.

(9) John is smart. Harry does too.

Clearly, "is smart" is not a possible antecedent for the elliptical VP "does too". More generally, I define a rule for objects of this form:

(10) [BE Adj]$_{VP}$

These are not possible antecedents for VP ellipsis, unless the VPE itself is a "BE-form", as in:

(11) John is smart. Harry is too.

It is not clear what the semantic basis for this rule is, in view of the following contrast:

(12) John knows the answer. Harry *is/does too.

(13) John is aware of the answer. Harry is/*does too.

I will simply assume that all VP's (and elliptical VP's) are marked with a feature [BE+/-] and that the antecedent and VPE must match in this feature.

5.3.3 Preference Rules

The next stage of the algorithm is to order remaining possible antecedents based on preference rules. This is implemented as follows: each element of VPlist is initialized to a preference level of 0. Then for each element of VPlist, if a preference rule applies to that element, its preference level is incremented by 1. There are two types of preference rules: subject coreference, and clausal relations.[11]

Subject Coreference

If a VP has a subject that is coindexed with the subject of the VPE, that VP is preferred. For example:

(14) a. But, darn it all, why should we$_i$ [help a couple of spoiled snobs who$_j$ had [looked down their noses at us]]?

[11] These preference rules are merely heuristic, and clearly admit counterexamples. My claim is simply that they appear to improve the performance of the algorithm. Ideally, they could be incorporated into a general theory of discourse structure, and its relation to anaphora resolution.

b. But, in the end, we$_i$ did.

Here, the subject of the VPE "we$_i$" corefers with the subject of the VP "help a couple . . . ", causing this VP to be preferred as the antecedent. The configuration, then, is the following:

(15) [NP$_i$ VP . . . NP$_i$ VPE]

In this case, VP is preferred as an antecedent for VPE.

Comparative and Adverbial Clauses

If the elliptical VP participates in a particular clausal relationship to another clause, the VP within that related clause is a preferred antecedent. Two clausal types are dealt with: comparative clauses and adverbial clauses.

An example of a comparative clause is the following:

(16) Now, if Morton's newest product, a corn chip known as Chip-o's, turns out to sell as well as its stock did . . .

In general, in a configuration of the form:

(17) [NP VP Comparative [NP VPE]]

the VP is a preferred antecedent.

There is a similar preference rule for adverbial clauses, as in the following example:

(18) But if you keep a calendar of events, as we do, you noticed a conflict.

In the configuration:

(19) [NP VP ADV [NP VPE]]

VP is a preferred antecedent[12].

[12] In some of these configurations, it appears that there is an *obligatory* antecedent. In Chapter Seven, I suggest that some such configurations might be explained as cases of bound-variable VP ellipsis.

5.3.4 Selecting the Antecedent

The antecedent with the highest preference level is selected as the antecedent. If there is more than one element on VPlist with the highest preference level, the antecedent nearest to the VPE is selected. Thus in cases where no preference rules apply, the most recent antecedent is always selected, unless it is an impossible antecedent.

5.3.5 Extensions to the Algorithm

In the algorithm as I have described it, the antecedent is selected from a list of VP meanings in surrounding discourse. However, while it has often been claimed that VP ellipsis requires a syntactic VP antecedent, this is not always the case. Below, I discuss extensions to the algorithm to cover cases where the antecedent is split between two VP's, or where the antecedent is something other than a syntactic VP.

Split Antecedents

As discussed in Chapter Two, the antecedent is sometimes "split" between two previous VP's; thus it is sometimes necessary to form a combined property from two properties in the discourse model. Consider again example (94), repeated below:

(20) Wendy is eager to **sail around the world** and Bruce is eager to **climb Kilimanjaro**, but neither of them can because money is too tight. [Webber, 1978]

The subject of the elliptical VP "neither of them" itself has a "split" antecedent, and the two antecedents are the subjects of the VP's which must be combined. This suggests the following rule: if the subject of the elliptical VP refers to a combination of two subject NP's, the combination of the two VP's associated with those subjects is a preferred antecedent. This constitutes a simple extension of the subject coreference preference rule, which governed configurations of the form NP_i VP ... NP_i VPE. The extension now covers configurations of the form:

(21) NP_i VP1 ... NP_j VP2 ... $\text{NP}_{\{i,j\}}$ VPE

That is, the antecedent [VP1,VP2] is constructed, and is preferred.

Another combination rule is suggested by example (96), from Chapter Three, repeated here:

(22) I can walk and I can chew gum. Gerry can too.

The rule here is that combinations of VP's that have the same subject are to be constructed, and a preference rule is applied to them. That is, in the following configuration:

(23) NP_i VP1 . . . NP_i VP2 . . . NP_j VPE

[VP1,VP2] is preferred.

Thus combined antecedents for VP ellipsis are only considered in specific cases, which are defined by simple extensions to the subject coreference rule.

Active/Passive

In this section, I consider "active/passive conflicts". I argued in Chapter Two that such cases could not be ruled out entirely. In this section, I suggest ways in which the location algorithm could be extended to handle the acceptable cases while still ruling out those that are unacceptable.

First, I look at cases with a passive antecedent and an active VP ellipsis. I repeat the examples of interest here:

(24) This information could have been released by Gorbachov, but he chose not to. (Daniel Schorr, NPR 10/17/92)

(25) Business needs to be developed differently than we have in the past. (5/24/91 NPR "Morning Edition" interview)

(26) The ice cream should be taken out of the freezer, if you can. (heard in conversation)

(27) A lot of this material can be presented in a fairly informal and accessible fashion, and often I do. (Chomsky 1982, cited in [Dalrymple *et al.*, 1991])

To allow such examples, it is necessary to expand the set of possible antecedents, VPlist. Upon the occurrence of a passive sentence S, we now add to VPlist a VP′, where VP′ is the VP in the active counterpart of S. I term VP′ the "active property" of S.

The preference level of VP′ is initialized below that of other elements of VPlist. Thus, it will never be selected as the antecedent, unless some preference rule applies to it. One preference rule concerns clausal relationships between elliptical VP and the antecedent clause. The configurations for clausal relationships were given as follows:

- [NP VP Comparative [NP VPE]]
- [NP VP ADV [NP VPE]]

In this configuration, the preference level of VP is incremented. In addition, if NP VP is a passive sentence, we also increment the preference level of the "active property". For example, in (25), the active property *develop business* is added to VPlist, and its preference level is incremented by the clausal relationship rule. This permits the desired reading. If the clausal relationship is not present, the reading is sharply degraded:

(28) ? Business needs to be developed differently. And in the future, I'm sure we will.

Another preference rule is the subject coreference rule, which states that an elliptical VP with subject S will prefer an antecedent VP whose subject corefers with S. To apply to the above examples, this would mean that the "deep subject" of the active property corefers with the grammatical subject of the VP ellipsis. This is the case for (24). With a different subject, the example becomes degraded:

(29) ? This information could have been released by Gorbachov, but Yeltsin chose to instead.

Consider now example (26). This does not fall under the subject coreference rule, since the "deep subject" is unexpressed in the antecedent. However, one might argue that "you" is understood as the deep subject of the antecedent clause. This suggests a looser formulation of the subject coreference rule.

It is interesting to contrast the naturally-occurring (26) with the following example from [Hankamer and Sag, 1976]:

(30) * The oats had to be taken down to the bin, so Bill did.

This example clearly does not fall under the subject coreference rule, since "Bill" is in no sense the understood deep subject of the antecedent. The example improves if "Bill" can be so understood:

(31) Bill knew the oats had to be taken down to the bin, so he did.

Now I look at the reverse case, where there is an active antecedent and a passive VP ellipsis. Again, this is possible only in restricted cases. The following example seems marginal at best:

(32) ?* Max fired Harry, although it was Tom who should have been.
[Fiengo and May, 1992b]

However, I find the following variations quite acceptable:

(33) It was Harry who they fired, although it was Tom who should have been.

(34) Harry, they fired, although Tom's the one who should have been.

Intuitively, it appears that (33) and (34) are acceptable because the antecedent for the elliptical VP is the property being applied to Harry. This might be considered the "main predication" in these sentences, despite the fact that this property is not associated with a syntactic VP. To extend the algorithm to cover these cases, it would be necessary to identify alternative predication structures, such as topicalization, and introduce the predicate in such structures into the discourse model.

In Chapter Two,I argued that the active/passive conflict examples cannot be categorically ruled out. Here, I have suggested that their acceptability can be correlated with independently motivated preference rules, as well as a pragmatic notion of predication.

Noun Antecedents

In Chapter Two, I pointed out that a noun or NP could be the antecedent for an elliptical VP, as in example (111), repeated here:

(35) David Begelman is a great **laugher**, and when he does, his eyes crinkle at you the way Lady Brett's did in *The Sun Also Rises*. (p. 90, *You'll Never Eat Lunch in This Town Again*, Julia Philips)

The antecedent configuration here is

(36) NP be a N

where the common noun denotes a property that is the antecedent for a VP ellipsis. In such a sentence, a possible semantic representation is

(37) $N'(NP')$

where N′ is the property denoted by N. This property is the main predication in the sentence, so it is natural to expect that it is a possible antecedent for VP ellipsis, despite the fact that it is associated with a noun rather than a VP. The system of derivation must be expanded, so that the property N′ is added to the discourse model.

One way to do this relies on the observation that a predicative indefinite such as the one in (35) do not introduce individuals into the discourse model. Therefore, I will allow predicative indefinites to introduce properties into the discourse model, namely, the property associated with the N′, or common noun. In this case, this property is added to the discourse model.

(38) λ x.laugh(x)

Now the location algorithm can operate unchanged with respect to such sentences. No preference rules apply to such N′ properties, so they will only be the antecedent if they are the nearest possible antecedent to the VP ellipsis occurrence.

5.3.6 A Partial Implementation

In this section I describe a Common LISP computer program that partially implements the antecedent location algorithm described here. The algorithm consists of three major components:

1. Eliminating impossible antecedents
2. Ordering possible antecedents by preference rules
3. Making a selection based on recency

The program implements (1) and (3), but not (2), the preference orderings.

The main function for the antecedent-selection program, **find1**, is defined as follows:

```
(defun find1 (window vpe)
  (remove-imposs(getvps window)) vpe)
```

"Window" is a three-sentence context for the elliptical VP. The VP's occurring within that window are gathered by **getvps**, whereupon impossible antecedents are removed by the function **remove-imposs**, displayed below:

```
; remove containing antecedents and be-form antecedents
(defun remove-imposs (vplist vpe)
  (cond ((null vplist) nil)
    ((or (be-do-conf (car vplist) vpe)
      (vpcontains (car vplist) vpe))
        (remove-imposs (cdr vplist) vpe))
    (t (cons (car vplist) (remove-imposs (cdr vplist) vpe)))))
```

This function examines each element of vplist, removing that element if there is a "be-do conflict", or if the element contains **vpe** in a proscribed way. This proscribed configuration is an elliptical VP contained within a sentential complement of a containing VP. The remaining vps are in order of recency; thus the first VP on the list is the preferred antecedent.

In Chapter Six, I describe a test of this program, using examples from the Penn Treebank.

Chapter 6

Empirical Studies

In this chapter, I describe two sorts of empirical studies: first, I describe the collection of VP ellipsis occurrences in two corpora. Second, I describe the testing of the antecedent location algorithm, using examples collected from these corpora.

6.1 Locating VP Ellipsis Occurrences

I have performed two corpus searches for VP ellipsis occurrences. The first search was performed on a version of the million-word Brown Corpus where each word is tagged for part of speech. The second search was performed on the Penn Treebank: a completely parsed corpus of Wall Street Journal articles, containing about two million words.

6.1.1 Brown Corpus

I used the UNIX pattern-matching utility "grep" to locate cases of VP ellipsis in the Brown Corpus. The version of the Brown Corpus I used has each word tagged for part of speech. I defined search patterns for auxiliary verbs followed by a punctuation mark, i.e., period, question mark, exclamation mark, semicolon or colon, with at most one intervening non-verb. About 45%, or 315, of the the 752 sentences selected actually did contain VP ellipsis. The 315 cases of VP ellipsis found constitute .6% of the approximately 55000 sentences in the Brown Corpus. These examples cover the full range of types of material in the Brown Corpus, including both "Informative" (e.g., journalistic, scientific, and government texts) and "Imaginative" (e.g., novels,

short stories, and humor).

The following is a typical matching example:

(39) a. Not that Mr. Hedison does not **make the most of his role**.

b. He does, and more.

The fact that the Brown Corpus is tagged for part of speech greatly facilitated the search process; in particular, the specification of a non-verb between auxiliary and punctuation mark greatly reduced the number of false positives to manually sort through.

An example of a false positive for this pattern:

(40) By July 1, six weeks from now, motel-keepers all over the nation **will**, by 6 p.m., **be switching** on that bleak – to motorists – sign, " No Vacancy ".

Nor did the search pattern catch every example of VP ellipsis; an example of VP ellipsis that did not match the pattern is:

(41) "I couldn't **write** with them in the same room with me, but I could with Harold."

In addition, no examples of VP ellipsis with stranded "to" were found.

The cases were classified based on the distance of the antecedent from the ellipsis. There were three categories:

1. same sentence
2. immediately preceding sentence
3. a more distant sentence, ("Long-Distance")

The results of this categorization are displayed below:

VP Ellipsis *Category*	*Occurrences*	%
Same-sentence	200	64
Adjacent-sentence	99	31
Long-Distance	15	5

In 15 cases, or nearly 5% of the VP ellipsis examples, the antecedent was in neither the same sentence or the most recent. Two examples of this are the following:

> I disagree with the writer who says funeral services **should be government-controlled** . The funeral for my husband was just what I wanted and I paid a fair price, far less than I had expected to pay . But the hospitals and doctors should be.

> The thought came back, the one nagging at him these past four days. He tried **to stifle it**. But the words were forming. He knew he couldn't.

6.1.2 Treebank

Next, I describe a search in the Penn Treebank; a completely parsed corpus of Wall Street Journal articles, containing about two million words. To locate VP ellipsis occurrences in the parsed Treebank, I defined two lisp functions to identify VP ellipsis occurrences: **contains_vpaux** and **s_withaux**. The function **contains_vpaux** identifies auxiliary verbs that are parsed as VP's. Consider the following example:

```
(S
  (NP Congress)
  (VP (VP closed
        (NP this loophole)
        (ADVP (NP last year)))
      ,
      or
      (VP thought
        (SBAR 0
          (S (NP it)
             (VP did))))))).)
```

The structure **(VP did)** is identified by **contains_vpaux**. Of course the presence of the auxiliary verb is not enough to signal the occurrence of VP ellipsis; **contains_vpaux** looks only at the final element in the VP, checking to see if that element is an auxiliary or negation. This is sufficient to

determine that no main verb is present, since a main verb can never be followed by an auxiliary or negation.

The Treebank parses of VP ellipsis do not always take this form. In particular, modals are often immediate daughters to the S node, as in the following example:

(42) A lot of people think I will **give away the store**, but I can assure you I will not.

In the Treebank, the VP ellipsis is parsed as follows:

```
(S (NP I)
   can
   (VP assure
     (NP you)
       (SBAR 0
          (S
            (NP I)
              will
              not))))
```

Such examples are identified by **s_withaux**, which looks for an S with an auxiliary as immediate daughter, but no VP as immediate daughter. These two functions succeeding in identifying 318 examples of VP ellipsis from the Penn Treebank Wall St. Journal corpus. They also mistakenly identified about 700 cases as VP ellipsis. Many of these false hits can be eliminated by simple improvements to the VP identification functions. However, some problems appear more difficult. For example: the verb "do" is ambiguous between an auxiliary use and a main-verb use. In the main-verb use, "do" requires an object. However, when the object for the main-verb do is a trace, it is difficult to distinguish from VP ellipsis. In fact, this construction appears to be quite common, as in the following example:

(43) "...most of what I do t is very anonymous".

This is parsed in the Treebank as follows:

```
(S  ``
   (ADJP Most
      (SBAR (WHPP of
         (WHNP what))
       (S (NP I)
          (VP do))))
   (VP is
      (ADJP  very
      anonymous)
```

Since the Treebank parse does not include the trace, examples of this sort are indistinguishable from VP ellipsis in the Treebank, and must be removed manually.

6.2 Testing the Antecedent Location Algorithm

I have performed two empirical tests of the antecedent location algorithm. The first test was performed manually on examples collected from the Brown Corpus. The second test was performed automatically on the Penn Treebank, using the partial implementation of the algorithm described above.

6.2.1 Brown Corpus

The algorithm was manually tested on a set of 304 examples of VP ellipsis collected from the Brown Corpus. I have divided these examples into three categories, based on whether the antecedent is in the same sentence as the VPE, the adjacent (preceding) sentence, or earlier ("Long-Distance").

The algorithm selected the correct antecedent in 285, or 94% of the cases. For comparison purposes, I present results of an alternative strategy; namely, a simple linear scan of preceding text. In this strategy, the first verb that is encountered is taken to be the head of the antecedent VP.[1]

The results of the algorithm and the "Linear Scan" approach are displayed in the following table.

[1]These result are based on my intuitive judgement of the correct antecedent. There were a few cases in which the judgement was somewhat difficult; ideally, the correct antecedent should be determined by some independent fashion, perhaps using a group of informants.

Category		**Algorithm** No. Correct	**Linear Scan** No. Correct
Same-sent	196	193(96%)	172(88%)
Adj-sent	93	85(92%)	72(77%)
Long-Dist	15	7(47%)	2(13%)
Total	304	285(94%)	247(81%)

The algorithm performs considerably better than Linear Scan. Much of the improvement is due to "impossible antecedents" which are selected by the Linear Scan approach because they are closest to the VPE. A frequent case of this is containing antecedents that are ruled out by the algorithm. Another case distinguishing the algorithm from Linear Scan involves coreferential subjects. There were several cases in which the coreferential subject preference rule caused an antecedent to be selected that was not the nearest to the VPE. One example is:

(44) a. But, darn it all, why should we help a couple of spoiled snobs who had looked down their noses at us?

b. But, in the end, we did.

Here, the correct antecedent is the more distant "help a couple of...", rather than "looked down their noses...". There were no cases in which Linear Scan succeeded where the algorithm failed.

I will now look at sources of errors for the algorithm. The performance was poorest in the Long Distance category, in which at least one sentence intervenes between antecedent and VPE. In several problem cases in the Long Distance category, it appears that intervening text contains some mechanism that causes the antecedent to remain salient. For example:

(45) a. "...in Underwater Western Eye I'd have a chance to act. I could show what I can do".

b. As far as I was concerned, she had already and had dandily shown what she could do.

In this case, the elliptical VP "had already" means "had already had a chance to act". The algorithm incorrectly selects "show what I can do" as the antecedent. The intervening sentence causes the previous antecedent to remain salient, since it is understood as "(If I had a chance to act then) I could show what I can do." Furthermore, the choice made by the algorithm might perhaps

be eliminated on pragmatic grounds, given the oddness of "she had already shown what she could do and had dandily shown what she could do."

Another problem with the algorithm is illustrated by the follow example:

(46) a. "I didn't ask you to fight for the ball club", Phil said slowly.

b. "Nobody else did, either".

Here the algorithm incorrectly selects "fight for the ball club" as the antecedent, instead of "ask you to fight for the ball club". The subject coreference rule does not apply, since "Nobody else" is not coreferential with the subject of any of the possible antecedents. However, its interpretation is dependent on the subject "I" of "ask you to fight for the ball club". Thus, if one generalized the subject coreference rule to include such forms of dependence, the algorithm would succeed on such examples.

Many of the remaining errors involve an antecedent that takes a VP or S as complement, often leading to subtle ambiguities. One example of this is the following:

(47) a. Usually she marked the few who did thank you, you didn't get that kind much in a place like this: and she played a little game with herself, seeing how downright rude she could act to the others, before they'd take offense, threaten to call the manager.

b. Funny how seldom they did: used to it, probably.

Here the algorithm selects "call the manager" as antecedent, instead of "threaten to call the manager", which I determined to be the correct antecedent. It may be that many of these cases involve a genuine ambiguity.

The problem addressed here of locating the antecedent for an elliptical VP, has received little attention in the literature. Most treatments of VP ellipsis [Sag, 1976; Williams, 1977; Webber, 1978; Fiengo and May, 1990; Dalrymple *et al.*, 1991] have focused on the question of determining what readings are possible, given an elliptical VP and a particular antecedent. For a computational system, a method is required to determine the antecedent, after which the possible readings can be determined.

[Lappin and McCord, 1990] present an algorithm for VP ellipsis which contains a partial treatment of this problem. However, while they define three possible ellipsis-antecedent configurations, they have nothing to say about selecting among alternatives, if there is more than one

VP in an allowed configuration. The three configurations given by Lappin and McCord for a VPE-antecedent pair $<$ V,A$>$ are:

1. V is contained in the clausal complement of a subordinate conjunction SC, where the SC-phrase is either (i) an adjunct of A, or (ii) an adjunct of a noun N and N heads an NP argument of A, or N heads the NP argument of an adjunct of A.

2. V is contained in a relative clause that modifies a head noun N, with N contained in A, and, if a verb A$'$ is contained in A and N is contained in A$'$, then A$'$ is an infinitival complement of A or a verb contained in A.

3. V is contained in the right conjunct of a sentential conjunction S, and A is contained in the left conjunct of S.

An examination of the Brown Corpus examples reveals that these configurations are incomplete in important ways. First, there is no configuration that allows a sentence intervening between antecedent and VPE. Thus, none of the Long-Distance examples (about 5% of the sample) would be covered. Configuration (3) deals with antecedent-VPE pairs in adjacent S's. There are many such cases in which there is no sentential conjunction. For example:

(48) a. All the generals who held important commands in World War 2, did not write books.

b. It only seems as if they did.

Perhaps configuration (3) could be interpreted as covering any adjacent S's, whether or not an explicit conjunction is present.

Furthermore, there are cases in which the adjacent categories are something other than S; in the following example, the adjacent category is S$'$.

(49) I remember him pointing out of the window and saying that he wished he could live to see another spring but that he wouldn't.

Configurations (1) and (2) deal with antecedent-VPE pairs within the same sentence. In Configuration (1), the VPE is in a subordinate clause, and in (2), the VPE is in a relative clause. In each case, the VPE is c-commanded by the antecedent A. While the configurations cover two quite common cases, there are other same-sentence configurations in which the antecedent does not c-command the VPE.

(50) In the first place, a good many writers who are said to use folklore, do not, unless one counts an occasional superstition or tale.

(51) In reply to a question of whether they now tax boats, airplanes and other movable property excluding automobiles, nineteen said that they did and twenty that they did not.

In sum, the configurations defined by Lappin and McCord would miss a significant number of cases in the Brown Corpus, and, even where they do apply, there is no method for deciding among alternative possibilities.

6.2.2 Treebank

The antecedent-selection program was tested on 124 examples automatically collected from the Treebank, using the VP ellipsis identification functions described in Chapter Five. The examples were presented to the program together with a three-sentence "window" of preceding context. The program selected the correct antecedent in 84 cases, which is a success rate of approximately 73%. This is significantly lower than the rate of 94% in a manual test of the algorithm described above. There are two factors which might explain all or part of this difference. First, some of the difference can be attributed to the fact that the program is an incomplete implementation of that algorithm; the preference rules of the algorithm were not implemented in the program. Second, there are a number of problems resulting from syntactic assumptions in the Treebank parses that conflict with syntactic assumptions in the antecedent-selection program. Below, I give examples of both of these problems.

It appears that the success rate of the program would be substantially improved by the addition of the preference rules. Consider the following example:

(52) In an interview with the Washington Post in early October, the secretary said the Fed may be slightly more interested in curbing inflation than the administration is...

This example is parsed as follows:

```
(S (PP In
       (NP an interview
         (PP with
```

```
          (NP the Washington Post))
            (PP in
              (NP (ADJP early)
                 October))))
    ,
(S (NP the secretary)
   (VP said
     (SBAR 0
      (S (NP the Fed)
          may
         (VP be
          (ADJP  slightly
            more
            interested
             (PP in
               (S (NP *)
                  (VP curbing
                    (NP inflation))))
                (PP than
                  (X (NP the administration)
                    (VP is))))
```

The antecedent-selection program incorrectly selects “curbing inflation” as the antecedent. The correct antecedent is the VP “be slightly more interested...”. A preference rule for such comparative configurations was included in the algorithm described above. This rule would cause the correct antecedent to be selected in this case. However, it must be noted that the relevant structure of a comparative configuration is difficult to define based on the parses as given in the Treebank, and the implementation of this condition is therefore left to future work.

The other preference ruled described in the algorithm is the subject-coreference rule. This would allow the correct antecedent to be selected in examples such as the following:

(53) So why does Renaissance's computer like stocks with the dow at 2653.28, where it closed yesterday, when it didn't with the dow at 2200?

The antecedent-selection program incorrectly selects "closed yesterday" as the antecedent for the elliptical VP "didn't". Since "it" corefers with "renaissance's computer", the subject coreference rule would cause the VP "like stocks" to be selected as the antecedent.

To implement the subject coreference rule, we need an antecedent-location program for NP's. In future work, I plan to incorporate such a program.

There were several cases in which the Treebank identifies structures as VP's, although they are not possible antecedents for VP ellipsis. This caused several errors. For example:

(54) The framers hardly discussed the appropriations clause at the Constitutional Convention of 1787, according to Madison's notes. To the extent they did, their concern was to ensure fiscal accountability.

The program incorrectly selects "according to Madison's notes" as the antecedent, because the Treebank parse is as follows:

```
(NP The framers)
hardly
(VP discussed
    (NP the appropriations
       clause)
    (PP at
       (NP the Constitutional Convention
           (PP of
             (NP 1787))))
    ,
    (ADVP (VP according
             (PP to
                 (NP
                   (NP Madison)
                   's notes))))))).)
```

```
(
 (S (PP To
        (NP the extent
             (SBAR 0
                  (S
                    (NP they)
                      (VP did)))))))
```

Another example of this is the following:

(55) Can Sihanouk and Hun Sen knock off the Khmer Rouge still supported by China? He can't.

This is parsed as follows:

```
(SBARQ Can
   (SQ (NP (NP Sihanouk)
               and
               (NP Hun Sen))
       (VP knock off
         (NP (NP the Khmer Rouge)
                (VP still supported
                    (PP by (NP China)))))))
( (S
   (NP He)
   (VP ca n't)))
```

Here the VP "still supported by China" is incorrectly selected by the program. In both of these cases, the selected VP is not a possible antecedent. A possible remedy for this would be to exclude any VP's that are not immediate daughters of an S.

The availability of the parsed Treebank corpus made it possibility to automatically test a program based on the algorithm manually tested on the Brown Corpus. The program selected the correct antecedent in 73% of the examples – significantly lower success rate than the 94% rate of

the manual test of the algorithm. We saw that this difference can be traced in part to the fact that the program is only a partial implementation of the algorithm; the preference rules omitted from the program would have rectified false selections in many cases. In addition, many false choices resulted from parses in the Treebank that conflicted with syntactic assumptions of the program.

Overall, these empirical studies have resulted in a database of about 700 examples of VP ellipsis. In addition, tools have been developed for the automatic location of VP ellipsis occurrences and VP ellipsis antecedents. These tools can be used to collect further data for linguistic research. Based on the empirical results described above, it is expected that subsequent versions of the program will exhibit significant improvements in performance, making it possible to use in practical Natural Language Processing applications.

Chapter 7

Related Issues, Open Problems, and Conclusions

In this chapter I begin with a consideration of overt forms related to VP ellipsis: "do it" and "do so". Next, I look at VP relatives – constructions where the gap or variable is a VP. This extends the analogy between VP ellipsis and pronominal anaphora; just as pronouns can be either discourse-bound or bound within a sentence, VP variables can be bound in a sentence as well as discourse-bound. I then examine some issues involving VP ellipsis and focus: we see that there are cases involving focus and VP ellipsis where there is no non-elliptical counterpart with the same interpretation. This provides further evidence for the proverb theory. I describe an open problem concerning parallelism constraints on indexing. Finally, I discuss elliptical forms related to VP ellipsis. I speculate that the class of elliptical forms should be divided into two classes: proforms and conjunction forms. Both of these forms are subject to distinct, highly constrained interpretation mechanisms. I explore some consequences and potential problems with this hypothesis.

7.1 Related Overt Forms: "do it", "do so"

In this section I consider the forms "do it" and "do so". Like VP ellipsis, these forms anaphorically refer to a property, and there are many cases in which these forms can be used interchangeably with VP ellipsis. For example:

(1) Tom threw a ball.

(2) a. Harry did too.

b. Harry did so too.

c. Harry did it too.

Despite this similarity, there are important differences between these forms and VP ellipsis. The related forms involve the main verb “do”, while VP ellipsis involves an auxiliary. Thus, while VP ellipsis is possible with any auxiliary verb, including modals, this is not true for the related forms, as shown by the following examples:

(3) Tom arrived.

(4) a. Harry should too.

b. * Harry should so too.

c. * Harry should it too.

A similar pattern is found with infinitival “to”:

(5) Tom went to the meeting.

(6) a. Harry wanted to.

b. * Harry wanted to so.

c. * Harry wanted to it too.

Also, the “do” in the related forms does not allow do-support, or subject-aux inversion. This is allowed with VP ellipsis, as shown by the following examples.

(7) Tom went to the meeting, but

(8) Harry didn’t.

(9) * Harry didn’t it.

(10) * Harry didn’t so.

(11) Tom went to the meeting, but

(12) why did he?

(13) * why did he it?

(14) * why did he so?

Another important difference is that, unlike VP ellipsis, the related forms impose a semantic constraint on the antecedent, requiring that it be an action. For example:

(15) Smith knew the answer.

(16) a. Jones did too.

(17) a. *Jones did so too.

(18) a. *Jones did it too.

(19) The glass shattered.

(20) a. The plate did too.

(21) a. *The plate did so too.

(22) a. *The plate did it too.

Thus there are two major differences between VP ellipsis and the related forms: the related forms involve main verb "do", and they impose semantic constraints on the possible antecedents. It appears, therefore, that, the main verb "do" selects either an NP or ADV argument. Thus "do it" and "do so" are just special cases of this; other possibilities for "do NP" are: do this, do the work, do business, and other possibilities for "do ADV" are: do well, do badly, do ok.

7.2 VP Relatives: Bound VP Variables

Pronouns can be divided into two classes: sentence-bound pronouns and discourse- bound pronouns. The basis for this division is that bound pronouns must find an antecedent within a single sentence, while discourse pronouns can find an antecedent in another sentence. In this dissertation I have analyzed VP ellipsis as a property variable that is discourse-bound. In this section, I look at a construction in which an VP variable is bound by an operator, so that it must find its antecedent within the same sentence. Consider the following examples:

(23) "France opened the limit down, off at least 10% if you could calculate the index, **which you couldn't**," Mr. Clark, the Shearson trader, said early in the afternoon. (Penn Treebank)

(24) Don't like to bother no one unless we have to, **which I figger we do**, in your case. (Brown Corpus)

(25) But if you keep a calendar of events, **as we do**, you noticed a conflict. (Brown Corpus)

I will analyze the clause "as/which we do" as involving a property variable that is bound by the operator (as/which). One result of this is that the elliptical VP is not free to select its antecedent from the discourse model; the VP "keep a calendar of events" must be the antecedent.

It might appear that there is an alternative analysis, namely, that these examples are not VP ellipsis at all, but rather, they are the related forms "do-it" and "do-so". However, this analysis is not tenable, since it would require that this is main verb "do". The fact that modals are possible suggests that this is not the main verb "do".

(26) But if you keep a calendar of events, which/as you should, . . .

Furthermore, there is no requirement for an action antecedent.

(27) If Smith knew the answer, which/as Jones did . . .

(28) If the glass shattered, which/as the plate did . . .

This suggests that the variable is an empty VP, just as we see in VP ellipsis, except for the fact that, here, the VP variable is bound by the wh- or "as" operator. Thus the structure in general is: following:

$$_{VP}[\ldots]_i \text{ OP}_i \ldots {}_{AUX}[\ldots] {}_{VP}[\text{e}_i]$$

This is is analogous to other relative clauses, which take the general form:

$$_{XP}[\ldots]_i \text{ OP}_i \ldots {}_{XP}[\text{e}_i]$$

Note that the VP relatives have the additional restriction of a preceding AUX; I won't pursue the reason for this here.

To interpret a relative clause, the gap is treated as a variable that is semantically identified with the head of the clause. This is quite similar to the proverb theory of VP ellipsis, where the empty VP is semantically identified with its antecedent; the only difference is that, in the case of a VP relative, the head of the clause is structurally identified as the antecedent; for VP ellipsis, there is generally no structurally designated antecedent.

The proverb theory predicts that VP relatives have the same interpretive possibilities as VP ellipsis. As evidence of this, observe that sloppy identity readings are possible for VP relatives:

(29) Tom **phoned his parents**, which Harry also did.

As the following example shows, variables bound outside the antecedent VP can also receive sloppy readings, which again patterns with VP ellipsis.

(30) Every student thinks Mrs. Smith will **fail him**, which, in Harry's case, I'm afraid she will.

Assuming that the VP variable is identified with its antecedent, that variable must be a dynamic property (i.e., relativized to contexts) to permit these readings. We see similar sloppy readings for NP relatives:

(31) Every student asked University Press to read **her paper**, which, in Mary's case, I think they will publish.

Finally, there are nominal relative clauses which have "split antecedents", as in the following example:

(32) A man entered the room and a woman went out who were quite similar. [Perlmutter and Ross, 1970]

Here, the antecedent for the NP relative "who were quite similar" is a plural object composed of the man and the woman. A similar possibility exists for VP ellipsis:

(33) John wanted to go to India and Harry wanted to go to China, which it turned out they couldn't.

The VP relative construction provides further support for the proverb theory, because it extends the analogy between VP ellipsis and pronouns: just as there are discourse pronouns and discourse VP ellipsis, there are sentence-bound pronouns and sentence-bound VP variables. Furthermore, the sloppy identity facts support the view that the basic notion of semantic identity must be dynamic.

7.3 VP Ellipsis and Focus

Next, I examine the interaction of VP ellipsis with focus. Consider the following example, due to [Kratzer, 1992]:

(34) A: You went to Block Island because I did. You went to Elk Lake Lodge because I did. And you went to Tanglewood because I did.

(35) B: I only went to TANGLEWOOD because you did.

The meaning of (35) is:

I went to Tanglewood because you went to Tanglewood, but
I did not go to Block Island because you went to Block Island, and
I did not go to Elk Lake Lodge because you went to Elk Lake Lodge.

Kratzer assumes that VP ellipsis is resolved by a syntactic LF copy rule. She shows that Rooth's approach is not able to derive the correct reading for this example. Kratzer shows that the reading can be derived by extending Rooth's approach with an additional indexing mechanism. However, I will show that the proverb theory permits the correct reading to be derived by Rooth's original approach. In addition, it makes it possible to avoid a questionable aspect of Kratzer's account, namely, allowing a focus marking within elided material. This contradicts a widespread assumption that ellipsis is closely related to destressing.

Interestingly, this example has no equivalent non-elliptical counterpart. There are two possibilities for the non-elliptical counterpart of (35): depending on whether "Tanglewood" is focused or unfocused when repeated. If it is focused, the non-elliptical counterpart is the following:

(36) I only went to TANGLEWOOD because you went to TANGLEWOOD.

Here, as Kratzer points out, the meaning is:

I went to Tanglewood because you went to Tanglewood.
I did not go to Tanglewood because you went to Elk Lake Lodge,
I did not go to Tanglewood because you went to Block Island,
I did not go to Elk Lake Lodge because you went to Block Island,
I did not go to Elk Lake Lodge because you went to Elk Lake Lodge,
I did not go to Elk Lake Lodge because you went to Tanglewood,
I did not go to Block Island because you went to Block Island,
I did not go to Block Island because you went to Elk Lake Lodge,
and I did not go to Island because you went to Tanglewood.

Perhaps the non-elliptical counterpart of (35) does not have focus on "Tanglewood" when repeated. Then the reading is the following:

(37) I only went to TANGLEWOOD because you went to Tanglewood.

Here we get the following reading:

I went to Tanglewood because you went to Tanglewood,
I did not go to Elk Lake Lodge because you went to Tanglewood, and
I did not go to Block Island because you went to Tanglewood.

The readings for (36) and (37) are derived according to Rooth's method for deriving p-sets, by generating a set of alternatives by replacing each focused constituent with each possible element in the domain of the appropriate type.

Kratzer assumes that VP ellipsis is resolved by syntactic copying at LF, and notes that, given this, the correct reading for (35) cannot be derived by Rooth's method. Thus, Kratzer suggests a modified version of Rooth's method, in which each focused expression is associated with a "designated variable". Next, Kratzer gives a "novelty constraint" that requires each focused expression at S-structure to be associated with a distinct designated variable. However, the novelty constraint does not hold at LF. This allows the correct reading to be derived for (35). On this approach, the S-structure representation is:

(38) I only $_{F2}$[went to TANGLEWOOD] because you did.

The VP ellipsis is resolved at LF by copying the antecedent VP, resulting in:

(39) I only $_{F2}$[went to TANGLEWOOD] because you $_{F2}$[went to TANGLEWOOD].

To determine the pset, we replace each focussed expressions with a designated variable, giving:

(40) I only went to v_2 because you went to v_2.

Thus the same designated variable occupies the two focus positions. We determine the p-set by replacing v_2 with all elements of the domain, giving the desired p-set:

go to Tanglewood because you went to Tanglewood,
go to Block Island because you went to Block Island,
go to Elk Lake Lodge because you went to Elk Lake Lodge.

If one adopts the proverb theory, Rooth's original method derives the correct reading for (35). On this approach, (35) is represented as follows:

(41) I only [went to TANGLEWOOD]$_i$ because you did P_i.

The elliptical VP is represented simply as a property variable P_i.[1] The pset is a set of properties:

$\{\lambda$ y.y [went to x]$_i$ because you did $P_i\}$

By substituting all the possible objects for x, we get the desired pset:

go to Tanglewood because you went to Tanglewood,
go to Block Island because you went to Block Island,
go to Elk Lake Lodge because you went to Elk Lake Lodge.

This example can be interpreted in the same way as the following example involving nominal anaphora:

(42) I only saw [PINK Edsels]$_i$ when you saw them$_i$

Here the pset is presumably of this form:

[1]This notation, in which P is coindexed with the antecedent VP, is a shorthand representation for the fact that the meaning of "went to TANGLEWOOD" was stored in the discourse model, and was selected as the antecedent for P.

I saw pink Edsels when you saw pink Edsels

I did not see green Edsels when you saw green Edsels

I did not see blue Edsels when you saw blue Edsels

etc. . .

This is computed with respect to the following representation:

$\{\lambda$ y.y saw [x edsels]$_i$ when you saw them$_i$.$\}$

Thus the proverb theory allows the correct reading for (35) in Rooth's original system; the reading is derived in the same way a similar example involving nominal anaphora is derived. In addition, the proverb theory avoids the need to represent elided material as focus-marked. In VP ellipsis cases such as this, where the antecedent contains a focused element, we have seen that there is no equivalent non-elliptical counterpart.

7.4 An Open Problem: Parallelism Constraints on Indexing

In this section I discuss some parallel indexing effects related to VP ellipsis. Many accounts have been motivated, in part, by a desire to capture these effects. I will discuss two attempts to capture these constraints: the first, due to Fiengo and May, is based on a syntactic reconstruction approach to VP ellipsis. The second, due to Helle Sim, is based on a DRT approach to VP ellipsis.[2]

Both of these approaches succeed in accounting for interesting aspects of parallel indexing effects. However, I will suggest that these effects are not categorical, as suggested in these accounts. Rather, they are tendencies that can be overridden by other pragmatic factors. Second, I will argue that these effects cannot be explained by reference to a syntactic reconstruction account of VP ellipsis, even if one believes that VP ellipsis is resolved by syntactic reconstruction, because similar effects obtain with other constructions where there is clearly no syntactic reconstruction.

7.4.1 Simple Case: A Single Pronoun

Many authors [Sag, 1976; Dalrymple *et al.*, 1991; Prüst *et al.*, 1991; Fiengo and May, 1992b] have pointed to parallel indexing effects that appear with VP ellipsis. A simple example is the following:

[2] The accounts in[Prüst and Scha, 1990; Prüst *et al.*, 1991] describe an approach in which parallelism constraints are imposed by matching both syntactic and semantic structure.

(43) John **saw her**. Harry did too.

The interpretation for "her" must be the same under ellipsis as it was in the antecedent: Harry and John must have seen the same person. However, it is not clear if this requirement has anything to do with ellipsis, since the non-elliptical counterpart of (43) seems similar:

(44) John **saw her**. Harry saw her too.

Without contrastive stress or deictic gesture, one would again expect that Harry and John saw the same person. In fact, it is possible to construct contexts in which such variables get "sloppy" readings under ellipsis:

(45) A: Does everyone have their library card?

(46) B: No! I **left it on the bus**!

(47) C: I did too!

The antecedent for the elliptical VP contains a pronoun "it", which refers to B's library card in the antecedent, but to C's library card in the ellipsis site.

From these simple examples, it seems clear that the ellipsis mechanism should not build in a constraint enforcing strict parallelism in examples like (43), since, first, it is merely a tendency, rather than a categorical constraint, and second, there is a similar tendency towards parallelism with no ellipsis.

7.4.2 Two Pronoun Cases

We now turn to examples in which the antecedent VP contains two pronouns, such as the following:

(48) John **said he saw his mother**. Bill did too.

Assuming that the two pronouns refer to "John" in the antecedent, one might expect the following four possible readings:

1. Bill said John saw John's mother.
2. Bill said Bill saw Bill's mother.
3. Bill said Bill saw John's mother.

4. *Bill said John saw Bill's mother.

It has been pointed out by many authors ([Sag, 1976; Dalrymple *et al.*, 1991; Fiengo and May, 1992b; Sim,] that the fourth reading appears to be systematically missing in such examples. I will discuss two attempts to capture these constraints: the first, due to Fiengo and May, and the second, due to Helle Sim.

Fiengo and May argue that the elliptical VP is a syntactic copy of its antecedent, subject to a constraint on changes in indices, to the effect that an index may be changed from i in the antecedent to j in the ellipsis site only if i and j are resolved in parallel fashion. That is, i and j must be part of "parallel dependencies". To express this, Fiengo and May define two types of indices: independent and dependent, or α and β occurrences. Intuitively, a β occurrence is dependent on an α occurrence for its interpretation, and it is only dependent indices that can change from antecedent to ellipsis site. Fiengo and May define an indexical dependency ID as a sequence of elements, an index, and a structural description, as follows:

$$< (c_1^{\alpha}, c_2^{\beta}, \ldots, c_n^{\beta}), I, SD >$$

The elements $c_1^{\alpha}, c_2^{\beta}, \ldots, c_n^{\beta}$ all bear the index I; intuitively, all the β occurrences are dependent upon c_1^{α}. The SD can be thought of as the phrase marker containing the sequence of elements $c_1^{\alpha}, c_2^{\beta}, \ldots, c_n^{\beta}$. The sequence of elements must begin with an α occurrence, followed by one or more β occurrences, all of which are dependent on the α occurrence. All the elements in this sequence must of course bear the index I, and there must be no other occurrences of the index I in the structural description SD. Fiengo and May define the notion of a "parallel dependency", which they term *i-copy*:

> Two indexical dependencies ID and ID' are *i-copies* iff ID and ID' vary from each other in at most the value of I.

Now consider again the example, to see how this allows readings (1) through (3), but disallows (4). One possibility is that the antecedent VP can be copied to the ellipsis site with the indices unchanged. This will result in reading (1) above. If an index is changed, it must, according to Fiengo and May, be part of a dependency that is parallel to a dependency in the antecedent: in Fiengo and May's terminology, they must be *i-copies* A dependency can be thought as two

(or more) coindexed categories, together with the syntactic material intervening between the two categories. Now consider reading (2):

(49) John$_1$ said he$_1^{\beta}$ saw his$_1^{\beta}$ mother. Bill$_2$ [said he$_2^{\beta}$ saw his$_2^{\beta}$ mother].

Since the indices on "he" and "his" have been changed, we must check to see if they are in parallel dependencies: they dependencies are

<(John,he,his),1,[NP,V,NP,V,NP]>
<(Bill,he,his),2,[NP,V,NP,V,NP]>

These two dependencies are i-copies, thus the reading is permitted.

For reading (3), we have:

(50) John$_1$ said he$_1^{\beta}$ saw his$_1^{\alpha}$ mother. Bill$_2$ [said he$_2^{\beta}$ saw his$_1^{\alpha}$ mother].

Here only one index has been changed: "he$_2$". The two dependencies are

<(John,he),1,[NP,V,NP]>
<(Bill,he),2,[NP,V,NP]>

Again the two dependencies are i-copies, so the reading is permitted. Consider now the fourth reading:

(51) John$_1$ said he$_1^{\alpha}$ saw his$_1^{\beta}$ mother. Bill$_2$ [said he$_1^{\alpha}$ saw his$_2^{\beta}$ mother].

Here the changed index is on the second pronoun "his$_2$". It must be dependent upon Bill$_2$. The dependency relating his$_2$ to Bill$_2$ is

<(Bill,his),2,[NP,V,NP,V,NP]>

since "Bill" and "his" are the only elements bearing the index 2. There is no such parallel dependency in the antecedent, since, there is an α occurrence "he" intervening between "John" and "his" in the antecedent. Thus the dependency

<(John,his),1,[NP,V,NP,V,NP]>

is not realized in the antecedent, since it omits “he”, although it shares the index 1, and occurs with the structural description.

While the Fiengo and May approach is able to rule out the missing reading, it does so in a way that is incompatible with the present framework for two reasons. First, it requires syntactic reconstruction of the elliptical VP, while I have argued that VP ellipsis is a proform that is not syntactically reconstructed. Secondly, Fiengo and May define their dependency condition on the syntactic environments of the ellipsis and the antecedent. Since these can be separated by arbitrary amounts of material, this conflicts with the normal restriction of syntactic constraints to a single sentence.

Another way of ruling out the missing reading is suggested by [Sim,]. This approach does not require syntactic reconstruction or intersentential constraints. However, it does rely on an implicit weakening of the identity condition, in that a variable within the antecedent is treated differently at the ellipsis site than within the antecedent. Intuitively the constraint is that a “sloppy” variable is subject to a “locality requirement” to the effect that, at the ellipsis site, it find its antecedent in the smallest possible domain. In the current framework, this requires that any undischarged assumptions within the antecedent be marked with this “locality” requirement. The locality domain can be defined on the function-argument structure of the *sense*.

Consider this constraint with respect to the four readings given above. First, reading (1) is trivially permitted, because both variables within the antecedent are discharged before the VP meaning is stored, and thus there are no variables subject to a locality requirement. In reading (2), both variables are undischarged when the VP meaning is stored, and thus both are subject to the locality requirement. The variable “he” corefers with the subject Bill, which is clearly the most local antecedent, and the variable “his” corefers with “he”, again the nearest possible antecedent. In reading (3), only the pronoun “he” is left undischarged when the VP meaning is stored; “his” has been discharged, giving rise to the a strict reading. Since “his” is not a possible antecedent for “he”, we must expand the domain, whereupon “Bill” is found to be a possible antecedent. Now consider reading (4). Here “his” is the only undischarged variable. The most local antecedent is the variable “he”; since this is a possible antecedent, and it is the only variable in that domain, it *must* be the antecedent for “his”. But on reading (4), “his” does not corefer with “he”. This violates the locality constraint.

This makes it possible to rule out the missing reading without syntactic reconstruction or intersentential syntactic constraints. In Chapter Two, I discussed several examples in which intersentential syntactic parallelism constraints are violated. Also, I gave many examples in which syntactic reconstruction is impossible.

We have seen two approaches that account for the missing reading phenomenon: Fiengo and May define a parallelism constraint over syntactically reconstructed material, while Sim defines a constraint over DRT representations. On the proverb theory of VP ellipsis, the Fiengo and May parallelism constraint would not apply, since there is no reconstructed syntactic material at the ellipsis site. Even if one holds a syntactic reconstruction view of VP ellipsis, there are other constructions that exhibit the missing reading phenomenon, but do not undergo syntactic reconstruction. Consider:

(52) When asked about their whereabouts, John claimed that he visited his mother last night, and Bill did likewise/made the same claim.

While there is a sense in which the VP "claimed that he visited his mother last night" is the antecedent for "did likewise" or "made the same claim", it is difficult to imagine a theory in which this antecedence relationship is interpreted via syntactic reconstruction. Rather, there is a semantic relationship asserted between Bill's claim and John's. Still, we see the same missing reading phenomenon that we had with VP ellipsis. While there are many possible readings for (52), it does not have the reading in which Bill claimed that John visited Bill's mother.

Finally, it may be that, with a properly constructed context, the missing reading is available, and thus should not be ruled out by the grammar at all. Consider the following situation: John is suspected of murdering Bill's mother. Bill has claimed that John was visiting Bill's mother on the night in question. But John has presented as his alibi that he was home with his own mother that night. The district attorney says, in reference to the case against John:

(53) So where WAS John last night? John says he was at his mother's house, but BILL does too.

Here, the reading is: "Bill says John was at Bill's mother's house that night". This is the reading generally believed to be systematically unavailable. If it is indeed available, the mechanism described in this section becomes unnecessary.

While there are undoubtedly important parallel indexing effects which must ultimately be captured by a theory of discourse, it is not clear whether these are generally tendencies or categorical effects. What is clear is that they apply quite generally to discourses, and have nothing specific to do with VP ellipsis.

7.5 Related Elliptical Forms

There are a variety of elliptical forms that appear to be similar to VP ellipsis. In this section, I will suggest that all elliptical forms can be divided into two categories: proforms and conjunction forms.[3] This suggestion is quite speculative; supporting it would require the examination of a wide range of different elliptical constructions. Here, I will examine three other elliptical forms: gapping, pseudogapping and stripping. I will contrast these related forms with VP ellipsis, and discuss some general issues about the simple division of ellipsis into two classes, as well as some problems that arise.

The following examples illustrate VP ellipsis, gapping, pseudogapping, and stripping:

VPE:

(54) Tom **likes Mary**, and Harry does too.

Gapping:

(55) Tom **likes Mary**, and Harry Susan.

Pseudo-gapping:

(56) Tom **likes Mary**, and Harry does Susan.

Stripping:

(57) TOM **likes Mary**, not HARRY.

In gapping, the verb (including auxiliary) is elided. In pseudogapping, the main verb is elided, but the auxiliary is non-elided, and at least one obligatory element in the verb phrase is non-elided. Stripping involves the elision of everything except a single NP.

[3]The characterization of conjunction forms draws on work in [Rooth, 1992; Reinhart, 1991].

While I have analyzed VP ellipsis as a proform, I will argue that the proform analysis is not appropriate for the related elliptical forms. One piece of evidence for this is the following: while proforms can precede their antecedents, this is not possible for the related forms. Consider the following contrast[4]:

(58) Although Tom doesn't, Harry **likes Mary**.

(59) * Although Tom Susan, Harry doesn't **likes Mary**.

(60) * Although Tom doesn't Susan, Harry **likes Mary**.

(61) * Although not TOM, HARRY **likes Mary**.

Above, I discussed the case of VP relative clauses. In general, any proform should have a corresponding relative clause. There is no relative clause corresponding to the related forms:

(62) Tom **likes Mary**, which Harry does too.

(63) * Tom **likes Mary**, which Harry Susan.

(64) * Tom **likes Mary**, which Harry does Susan.

(65) * TOM **likes Mary**, which not HARRY.

Proforms like VP ellipsis permit an arbitrary amount of material intervening between antecedent and proform. The related elliptical forms require the antecedent and elided material to appear in adjacent clauses:[5]

(66) The claim that Tom will write books is surprising, but I think he will.

(67) * The claim that Tom will write books is surprising, and Harry articles.

(68) * The claim that Tom will write books is surprising, but he will articles.

(69) * The claim that Tom will write books is surprising, but not Harry.[6]

Elliptical expressions, I suggest, can be divided into proforms and conjunction forms, based on the following criteria:

Proform

[4]Examples of this form are discussed in [Lobeck, 1992].

[5]This contrast is pointed out in [Rooth, 1992].

[6]The reading of interest here is: "Harry will not write books." This reading, which would require non-adjacency, is unavailable. There is an available reading which does not violate adjacency: "The claim that Harry will write books is not surprising."

- Can apply backwards.
- Need not be adjacent.
- Can be relativized.
- Limited class of semantic types: individual, property, proposition.

Conj form

- Can't apply backwards.
- Must be adjacent.
- Can't be relativized.
- Arbitrary semantic type.

The interpretation of proforms is accomplished via the store and retrieve operations described in Chapter Three; these operations make essential reference to the discourse model. Conjunction forms, on the other hand, do not require access to the discourse model for recovery of missing material. Rather, they must appear in a specific syntactic configuration. The following interpretation mechanism, described in [Rooth, 1992], characterizes the interpretation of conjunction forms[7]:

$[\![\phi$ conj remnant $]\!]^\circ$ =
$[\![$ conj $]\!]^\circ([\![\phi]\!]^\circ, \triangle([\![\phi]\!]^f)([\![$ remnant $]\!]^\circ))$

The $\triangle$ operator performs lambda-abstraction over focused elements. $^\circ$ denotes the "ordinary" semantic value, while ^{f} denotes the "focus" semantic value.[8]

To illustrate, consider the derivation of the stripping example, (57).

$[\![$ TOM likes Mary, but not HARRY.$]\!]$ =
but (like(Tom,Mary), (λ x.NOT like(x, Mary)) Harry) =
but (like(Tom,Mary), NOT like(Harry,Mary).

The conjunction forms can also appear in other configurations, such as comparatives. I won't address the generalization of the conjunction interpretation operation to these forms. The

[7]The approach to conjunction forms sketched here is similar in spirit to the analysis of gapping in[Steedman, 1990].
[8]The reader is referred to [Rooth, 1992] for a discussion of focus semantic value.

interpretation mechanism enforces the constraints on conjunction forms listed above: it places no constraints on semantic type, but it does require adjacency and directionality. In addition, it requires focus on the corresponding elements in the left hand clause (ϕ). Focus determines the elements that the $\triangle$ operator will lambda-abstract over. The same operation applies to pseudogapping as well. The following is a derivation for example (56):

$[\![$ TOM likes MARY, and HARRY does SUSAN.$]\!]$ =
and (like(Tom,Mary), (λ <x,y>. like(x,y)) <Harry,Susan>) =
and (like(Tom,Mary), like(Harry,Susan).

If one considers the class of elliptical constructions uniformly, as in [Dalrymple *et al.*, 1991], it appears that the language interpreter is presented with a highly unconstrained problem. The division of ellipsis into proforms and conjunction forms suggests that each class is in fact highly constrained. Proforms are quite unconstrained in the relation of proform to antecedent; they need not be adjacent, and proforms may precede as well as follow their antecedent. Proforms are, however, highly constrained semantically, and the proform occurrence specifies the semantic type of the antecedent[9]. Conjunction forms, on the other hand, are unconstrained as to semantic type of the elided material. But they are very constrained in that the antecedent clause must immediately precede the ellipsis clause. Also, the corresponding elements must be unambiguously marked by focus. The $\triangle$ operator is similar to the matching mechanism of [Dalrymple *et al.*, 1991]. But it is applied here in a very constrained fashion: the parallel elements are completely specified by focus, and the antecedent clause must be conjoined to the ellipsis clause.

If this division of ellipsis into proforms and conjunction forms can be supported, the problem of ellipsis will, in a sense, be defined out of existence. In its place will be two separate, highly constrained mechanisms. Of course, the defense of this general picture requires examination of the whole range of ellipsis phenomena.

The case of pseudogapping appears to present some problems with this simple division. Sometimes, pseudogapping can apparently precede the antecedent, based on examples like the following [10]:

[9]This may be an overstatement, since one might claim that a pronoun like "it" can refer either to an individual or a proposition. In any case, if proforms exhibit any ambiguity, it is quite limited when compared to the variation of semantic types of conjunction forms.

[10]These examples are due to Shalom Lappin(p.c.)

(70) ? Before John does to Rosa, Max will propose to Lucy.

(71) ? Only because John did novels, would Max write plays.

These examples are difficult to evaluate; in (70), the non-elided object is a prepositional phrase that could be interpreted as an adjunct rather than an argument. This would then be VP ellipsis rather than pseudogapping, since pseudogapping requires a non-elided argument. A corresponding example with an unambiguous argument indeed appear unacceptable:

(72) * Before John will Rosa, Max will phone Lucy.

Another complication arises with (71), in that "did" has a main verb use in which it takes an NP argument, as in the following variant:

(73) ? Only because John did novels, would Max speak with him.

If one removes this ambiguity by replacing "did" with a modal, the example becomes clearly unacceptable:

(74) * Only if John will novels, will Max write plays.

Note that the corresponding VP ellipsis is acceptable:

(75) Only if John will, will Max write plays.

It is not clear, then, that either (70) or (71) is appropriately analyzed as pseudogapping. In (70), the non-elided prepositional phrase might be analyzed as an adjunct rather than an argument, in which case it is VP ellipsis rather than pseudogapping. In (71), there is a main verb use of "did" which is possible, so that there would be no ellipsis at all. When one looks at similar examples that are unambiguously pseudogapping, the backwards occurrences seem to be impossible.

Another interesting issue arises with so-called antecedent-contained deletion examples, like the following[11]:

(76) John invited everyone who Max did.

As discussed in Chapter Two, the proverb theory requires a pseudogapping analysis of these structures, since there would otherwise be a wh-trace within the ellipsis site[12]. It is not clear

[11] See [Sag, 1976; Williams, 1977; May, 1985] for the analysis of antecedent-contained VP deletion.

[12] See [Lappin and McCord, 1990] for arguments for a pseudogapping analysis of antecedent-contained deletion structures. [Jacobson, 1992] also proposes a treatment along these lines, in a categorial grammar framework.

how to apply the conjunction operation to this example. First, it must be generalized to apply to subordinate clauses as well as conjoined clauses[13]. If one assumes such a generalization, the correct reading can be derived if both "John" and "everyone" are focused, and there is a trace present as a remnant:

JOHN invited EVERYONE who Max did e

We can then apply the $\triangle$ operator, as follows:

$\triangle$ [[JOHN invited EVERYONE]]^f = λ <x,y> x invited y

Next, the extracted predicate is applied to the remnant [[Max did e.]]:

(λ <x,y>. x invited y) < Max, e> = Max invited e

This gives the correct resolution of the elided material. One problem with this relates to Rooth's focus requirement: it does not appear that focus is required on "everyone". This suggests that focus marking is simply one way of determining the lambda abstraction performed by the $\triangle$ operator. I won't explore ways of generalizing this here. This is an important problem for the proverb theory, since the proverb theory forces the pseudogapping analysis of examples like (76). Syntactic reconstruction theories are free to treat these examples as ordinary VP ellipsis. However, there are similar cases which must be treated as pseudogapping by syntactic reconstruction theories as well, as pointed out in [Fiengo and May, 1992a]. Consider:

(77) John criticized Mary, who Bill didn't.

Here again, it does not appear that focus is required for "Mary". To account for this, either pseudogapping should not be placed in the conjunction forms category, or the focus-marking requirement must be relaxed.

Finally, the pseudogapping analysis predicts that the following topicalized version of antecedent-contained ellipsis should be unacceptable:

(78) ? Everyone who Max did, John invited to his party too.

This suggests that the directionality constraint of the conjunction forms is loosened in the case of pseudogapping. Further work is required to see if the class of elliptical expressions can be divided into conjunction forms and proforms as suggested here.

[13]This is discussed by [Rooth, 1992] with respect to comparative clauses.

7.6 Conclusions

In this dissertation I have argued that VP ellipsis is a proverb, and is thus governed by a semantic identity condition. This suggests that it is in fact misleading to speak of ellipsis at all; rather, the missing VP is simply a variable, to be semantically interpreted just as other variables are. In this chapter, we have seen additional evidence for this: in the VP relative case and the focus examples, it is not possible to replace the missing VP with a syntactic copy of its antecedent. These cases provide further evidence for the proverb theory of VP ellipsis.

Bibliography

[Bach and Partee, 1980] E. Bach and B. Partee. Anaphora and semantic structure. In A. Ojeda, editor, *Papers from the parasession on pronouns and anaphora*, pages 1–28. CLS, Chicago, 1980.

[Baltin, 1991] Mark Baltin. Vp-ellipsis as non-lambda predicate ellipsis. New York University, December 1991.

[Barwise, 1987] Jon Barwise. Noun phrases, generalized quantifiers, and anaphora. In P. Gardenfors, editor, *Generalized Quantifiers*, pages 1–29. Dordrecht: Riedel, 1987.

[Brennan *et al.*, 1987] Susan E. Brennan, Marilyn Walker Friedman, and Carl J. Pollard. A centering approach to pronouns. In *Proceedings of the 25th Annual Meeting of the ACL*, 1987.

[Chao, 1987] Wynn Chao. *On Ellipsis*. PhD thesis, University of Massachusetts-Amherst, 1987.

[Chierchia, 1992] Gennaro Chierchia. Anaphora and dynamic interpretation. *Linguistics and Philosophy*, 18, 1992.

[Chomsky, 1981] Noam Chomsky. *Lectures on Government and Binding*. Foris, 1981.

[Dahl, 1973] Osten Dahl. On so-called 'sloppy identity'. In *Gothenburg Papers in Theoretical Linguistics 11*. University of Gothenburg, Sweden, 1973.

[Dalrymple *et al.*, 1991] Mary Dalrymple, Stuart Shieber, and Fernando Pereira. Ellipsis and higher-order unification. *Linguistics and Philosophy*, 14(4), August 1991.

[Dalrymple, 1991] Mary Dalrymple. Ellipsis and reconstruction. Xerox PARC, 1991.

[Dalrymple, 1992] Mary Dalrymple. The equational approach to ellipsis. Talk at Stuttgart Workshop on Ellipsis. University of Stuttgart, 1992.

[Davidson, 1980] Donald Davidson. The logical form of action sentences. In *Essays on actions and events*, pages 105–148. Clarendon Press, Oxford, 1980.

[Donnellan, 1966] Keith Donnellan. Reference and definite description. *Philosophical Review*, 75, 1966.

[Fiengo and May, 1990] Robert Fiengo and Robert May. Ellipsis and anaphora. Paper presented at GLOW, Cambridge University, Cambridge, England, 1990.

[Fiengo and May, 1992a] Robert Fiengo and Robert May. Ellipsis and apposition. CUNY Graduate Center and U.C. Irvine, 1992.

[Fiengo and May, 1992b] Robert Fiengo and Robert May. Indices and identity. CUNY Graduate Center and U.C. Irvine, 1992.

[Fodor and Sag, 1982] Janet Dean Fodor and Ivan A. Sag. Referential and quantificational indefinites. *Linguistics and Philosophy*, 5:355–398, 1982.

[Garnham and Oakhill, 1987] A. Garnham and J. Oakhill. Interpreting elliptical verb phrases. *The Quarterly Journal of Experimental Psychology*, 39A:611–627, 1987.

[Gawron and Peters, 1990] Jean Mark Gawron and Stanley Peters. *Anaphora and Quantification in Situation Semantics*. Center for the Study of Language and Information, Stanford, CA, 1990.

[Groenendijk and Stokhof, 1992] Jeroen Groenendijk and Martin Stokhof. Dynamic predicate logic. *Linguistics and Philosophy*, 1992.

[Guenthner and Lehmann, 1983] Franz Guenthner and Hubert Lehmann. Rules for pronominalization. In *Proceedings of the European Chapter of the ACL*, 1983.

[Haik, 1987] Isabella Haik. Bound vps that need to be. *Linguistic Inquiry*, 10, 1987.

[Hankamer and Sag, 1976] Jorge Hankamer and Ivan Sag. Deep and surface anaphora. *Linguistic Inquiry*, 7(3), 1976.

[Hardt, 1990] Daniel Hardt. A corpus-based survey of vp ellipsis. University of Pennsylvania, 1990.

[Heim *et al.*, 1991] Irene Heim, Howard Lasnik, and Robert May. Reciprocity and plurality. *Linguistic Inquiry*, 22(1), 1991.

[Heim, 1982] Irene Heim. *The Semantics of Definite and Indefinite Noun Phrases*. PhD thesis, University of Massachusetts-Amherst, 1982.

[Hestvik, 1992] Arild Hestvik. Subordination and strict identity interpretation of reflexives. 1992.

[Higginbotham, 1992] James Higginbotham. Anaphoric and common reference. MIT, 1992.

[Hirschberg and Ward, 1991] Julia Hirschberg and Gregory Ward. Accent and bound anaphora. *Cognitive Linguistics*, 2(2):101–121, 1991.

[Jacobson, 1992] Pauline Jacobson. Antecedent contained deletion in a variable-free semantics. In *Proceedings of the Second Conference on Semantics and Linguistic Theory*, Columbus, Ohio, 1992.

[Kamp, 1980] Hans Kamp. A theory of truth and semantic representation. In J. Groenendijk, T.M.V. Janssen, and M. Stokhof, editors, *Formal Methods in the Study of Language*, pages 277–322. Dordrecht, 1980. Volume 136.

[Kaplan, 1989] David Kaplan. Demonstratives. In Joseph Almog, John Perry, and Howard Wettstein, editors, *Themes From Kaplan*. Oxford University Press, 1989.

[Keenan, 1971] Edward Keenan. Names, quantifiers, and the sloppy identity problem. *Papers in Linguistics*, 4(2):211–232, 1971.

[Kitagawa, 1990] Yoshihisa Kitagawa. Deriving and copying predication. University of Rochester, 1990.

[Kitagawa, 1991] Yoshihisa Kitagawa. Copying identity. *Natural Language and Linguistic Theory*, 9(3):497–536, 1991.

[Klein and Stainton-Ellis, 1989] Ewan Klein and Kate Stainton-Ellis. A note on multiple vp ellipsis. *Linguistics*, 27:1119–1124, 1989.

[Klein, 1984] Ewan Klein. Vp ellipsis in dr theory. 1984.

[Koopman and Sportiche, 1991] Hilda Koopman and Dominique Sportiche. The position of subjects. *Lingua*, 85:211–258, 1991.

[Kratzer, 1992] Angelika Kratzer. The representation of focus. In von Stechow et al, editor, *The Handbook of Semantics*. de Gruyter, 1992.

[Ladusaw, 1979] William Ladusaw. *Polarity Sensitivity as Inherent Scope Relations*. PhD thesis, University of Texas, Austin, 1979.

[Lappin and Leass, 1991] Shalom Lappin and Herbert Leass. A syntactically based algorithm for pronominal anaphora resolution. IBM T.J. Watson Research Center and IBM Germany Scientific Center, March 1991.

[Lappin and McCord, 1990] Shalom Lappin and Michael McCord. Anaphora resolution in slot grammar. *Computational Linguistics*, 16(4), 1990.

[Lappin, 1984] Shalom Lappin. Vp anaphora, quantifier scope, and logical form. *Linguistic Analysis*, 13(4):273–315, 1984.

[Lappin, 1992] Shalom Lappin. The syntactic basis of ellipsis resolution. In *Proceedings of the Stuttgart Ellipsis Workshop*, Stuttgart, Germany, 1992.

[Lappin, 1993] Shalom Lappin. Ellipsis resolution at s-structure. In *Proceedings of NELS 23*, 1993.

[Levin, 1985] Nancy S. Levin. *Main-verb ellipsis in spoken English*. Garland Press, New York, 1985.

[Linebarger, 1987] Marcia Linebarger. Negative polarity and grammatical representation. *Linguistics and Philosophy*, 10:325–387, 1987.

[Link, 1983] G. Link. The logical analysis of plurals and mass terms: A lattice-theoretical approach. In R. Bauerle et al, editor, *Meaning, Use and Interpretation of Language*. de Gruyter, Berlin, 1983.

[Lobeck, 1992] Anne Lobeck. Licensing and identification of ellipted categories in english. In *Proceedings of the Stuttgart Ellipsis Workshop*, Stuttgart, Germany, 1992.

[May, 1985] Robert May. *Logical Form: Its Structure and Derivation*. MIT Press, Cambridge, MA, 1985.

[Nunberg, 1991] Geoff Nunberg. Varieties of deference. Talk at University of Pennsylvania, 1991.

[Nunberg, 1992] Geoff Nunberg. Two kinds of indexicality. In *Proceedings of the Second Conference on Semantics and Linguistic Theory*, pages 283–302, Columbus, Ohio, 1992.

[Partee and Bach, 1984] Barbara Partee and Emmon Bach. Quantification, pronouns, and vp anaphora. In *Truth, Interpretation and Information*. Dordrecht, 1984.

[Partee, 1975] Barbara Partee. Montague grammar and transformational grammar. *Linguistic Inquiry*, 6(2):203–300, 1975.

[Partee, 1991] Barbara Partee. Deictic and anaphoric pieces of meaning. Talk given at ASL/LSA Conference on Logic and Linguistics, 1991.

[Pereira and Pollack, 1991] Fernando Pereira and Martha Pollack. Incremental interpretation. *Artificial Intelligence*, 50(1):37–82, June 1991.

[Perlmutter and Ross, 1970] David M. Perlmutter and John R. Ross. Relative clauses with split antecedents. *Linguistic Inquiry*, 1:350, 1970.

[Prüst and Scha, 1990] Hub Prüst and Remko Scha. A discourse approach to verb phrase anaphora. In *Proc. ECAI*, 1990.

[Prüst *et al.*, 1991] Hub Prüst, Remko Scha, and Martin van den Berg. A formal discourse grammar tackling verb phrase anaphora. Department of Computational Linguistics, 1991.

[Quirk, 1972] Randolph Quirk. *A Grammar of Contemporary English*. Longman, 1972.

[Reinhart, 1983] Tanya Reinhart. *Anaphora and Semantic Interpretation*. Univ. of Chicago Press, 1983.

[Reinhart, 1991] Tanya Reinhart. Non-quantificational lf. In Asa Kasher, editor, *The Chomskyan Turn*, pages 360–384. Basil Blackwell, 1991.

[Ristad, 1990] E.S. Ristad. *Computational Structure of Human Language*. PhD thesis, Massachusetts Institute of Technology, 1990.

[Roberts, 1991] Craige Roberts. Distributivity and reciprocal distributivity. In *Proceedings from Semantics and Linguistic Theory I*, Cornell University, 1991.

[Rooth, 1981] Mats Rooth. A comparison of three theories of vp ellipsis. In W. Chao and D. Wheeler, editors, *Occasional Papers in Linguistics 7*. University of Massachusetts, Amherst, MA., 1981.

[Rooth, 1987] Mats Rooth. Noun phrase interpretation in montague grammar, file change semantics, and situation semantics. In P. Gardenfors, editor, *Generalized Quantifiers*, pages 237–269. Dordrecht: Riedel, 1987.

[Rooth, 1992] Mats Rooth. A theory of focus interpretation. *Natural Language Semantics*, 1:75–116, 1992.

[Ross, 1967] H. Ross. Constraints on variables in syntax. MIT Department of Linguistics and Philosophy, 1967.

[Sag and Hankamer, 1984] Ivan Sag and Jorge Hankamer. Towards a theory of anaphoric processing. *Linguistics and Philosophy*, 7:325–345, 1984.

[Sag, 1976] Ivan A. Sag. *Deletion and Logical Form*. PhD thesis, Massachusetts Institute of Technology, 1976. (Published 1980 by Garland Publishing, New York).

[Sag, 1981] Ivan A. Sag. Partial variable assignment functions, verb phrase ellipsis, and the dispensability of logical form. Stanford University, 1981.

[Schacter, 1977] Paul Schacter. Constraints on coordination. *Language*, 53:86–103, 1977.

[Sim,] Helle Sim. A drt analysis of vp ellipsis. University of Oslo, Norway.

[Steedman, 1990] Mark Steedman. Gapping as constituent coordination. *Linguistics and Philosophy*, 13(2):207–263, 1990.

[Tancredi, 1992] Christopher Tancredi. *Deletion, Deaccenting and Presupposition*. PhD thesis, Massachusetts Institute of Technology, 1992.

[Webber, 1978] Bonnie Lynn Webber. *A Formal Approach to Discourse Anaphora*. PhD thesis, Harvard University, 1978. (Published 1979 by Garland Publishing, New York).

[Wescoat, 1989] Michael Wescoat. Sloppy readings with embedded antecedents. Stanford University, 1989.

[Williams, 1977] Edwin Williams. Discourse and logical form. *Linguistic Inquiry*, 8(1):101–139, 1977.