VP Ellipsis and Contextual Interpretation

Daniel Hardt

Department of Computer and Information Sciences
University of Pennsylvania
Philadelphia, PA 19104 USA
(hardt@linc.cis.upenn.edu)

April 23, 1992

Abstract

A computational account of VP ellipsis is described, in which VP's are represented in the discourse model as contextually dependent semantic objects. It is argued that this approach can handle examples that are not allowed by alternative accounts. An implementation is defined in terms of extensions to the Incremental Interpretation System. The treatment of VP ellipsis is analogous to that of pronominal anaphora. It is suggested that the recency and salience constraints commonly thought to apply to pronominal anaphora might apply in a similar way to VP ellipsis.

1 Introduction

The problem of verb phrase ellipsis can be divided into two sub-problems:

Problem (1): how is an antecedent selected?

Problem (2): given a particular antecedent, how is it to be reconstructed at the ellipsis site?

Most work on VP ellipsis has dealt with Problem (2), concerning the copying or reconstruction of a particular antecedent. A wide variety of approaches to this problem have been proposed, including surface structure accounts ([14], [18]), "syntactic" LF ([5]), and semantic ([25], [28], [15], [21], [4], [23], [24]). However, I will argue that there is a natural level of representation that has not been pursued, which I will call the "properly semantic" level. I will show that this alternative has significant empirical advantages over other approaches to Problem (2). In addition, the approach suggests some possible ways of addressing Problem (1), which concerns selecting among alternative potential antecedents. This problem has been largely ignored.

There is a variety of evidence that indicates that VP ellipsis is resolved at a semantic rather than syntactic level of representation. This evidence includes the possibility of deictic VP ellipsis, inferrable antecedents, non-syntactically parallel antecedents, cases where the antecedent is formed by combining two or more salient predicates, and cases where the antecedent is separated from the target by one or more intervening sentences([27],[11]). However, existing semantic accounts have important empirical problems. For example, I argue in [11] that they do not permit pronouns to "switch reference" from antecedent to target in examples such as[1]:

(1) a. I told John$_i$ that I didn't expect **him**$_i$ to **fail his**$_i$ **exam.**

  b. I told Bill$_i$ that I did. [expect him$_j$ to fail his$_j$ exam]

Similarly, most existing accounts[2] do not permit a pronoun to be bound by different binders in antecedent and target, as in:

(2) Every boy$_i$ in Bill's class hoped Mary would ask **him**$_i$ out, but a boy$_j$ in John's class actually knew that she would. [ask him$_j$ out]

It is interesting to note that none of the existing semantic accounts qualify as "properly semantic" according to some fairly standard criteria. The modifications required to comply with these criteria, I will argue, are exactly the ones needed to solve these empirical problems. The criteria I have in mind are the following two general requirements for semantic representation, imposed in Montague's[20] "Universal Grammar":

Condition (1): The logical form language must be "dispensable".

---

[1] In examples of VP ellipsis, the antecedent is in bold, and the target, or reconstructed material, is bracketed.

[2] A possible exception is the account of Prüst et al([23], [24]). I discuss problems with this account in section 4.

Condition (2): Semantic representations must have contextual parameters.

None of the existing semantic accounts satisfies both of these requirements. As Partee and Bach[21] argue, the Sag/Williams account does not satisfy Condition (1), because it imposes an "alphabetic variance" condition, making essential reference to the syntax of logical form expressions. This condition is also imposed in Partee and Bach's account, and a similar condition arises in a very different setting in the account of Dalrymple, Shieber and Pereira[4].[3] Only Lappin's account [17] explicitly removes the alphabetic variance condition, bringing this account in accord with Condition (1). However, semantic representations do not have contextual parameters in Lappin's account, or in any of the other accounts.

Thus, although there is a persistent intuition that VP ellipsis requires a semantic treatment, no existing account is "properly semantic" in the sense required by conditions (1) and (2). In this paper I will describe such an account, in which the semantic representation of a VP is a three-tuple $< DM_{in}, P, DM_{out} >$, consisting of a property P and input and output discourse models. A key feature of this approach is that the antecedent is reconstructed at the ellipsis site as a semantic object which includes contextual dependencies. These contextual dependencies can be resolved independently in the antecedent and the target. This is required for examples such as (1) and (2).

No reference to the syntax of logical form expressions is made in this approach, satisfying Condition (1). The representation of VP's as relations involving input and output discourse contexts satisfies Condition (2). So this account is more "properly semantic" than alternative semantic accounts, whose theoretical status is somewhat less clear. One consequence of this theoretical clarity is the ease with which the approach can be computationally implemented.

I will describe an implementation of this approach in terms of some simple extensions to the Incremental Interpretation System[22]. The fact that this system incorporates contextual dependencies, as required by Condition (2), makes it very simple to implement the approach. Indeed in an important sense there are no additional mechanisms required for VP ellipsis, over and above those independently required for pronominal and other forms of anaphora.

I begin with a brief overview of the Incremental Interpretation System. I then describe my extensions to

---

[3]The account of Klein [15], while couched in the DRT formalism, essentially duplicates the Sag/Williams approach, defining versions of the Derived VP rule and the Pronoun Rule in DRT terms. Sells [26] also suggests storing properties in a DRT-style discourse model, although he does not apply this to VP ellipsis.

this system which implement the type of approach to VP ellipsis I am advocating, and I describe the derivation of an example that cannot be accommodated by alternative accounts. Finally, I point out that the current approach suggests some promising avenues for progress on the neglected question concerning the selection of an antecedent VP.

2 Background: The Incremental Interpretation System

A semantic representation in the Incremental Interpretation (henceforth II) System is called a "Conditional Interpretation", which is defined as an assumption-sense pair, A:s, where A is a set of assumptions, and s is the sense. The sense can be thought of as the ordinary truth-conditional semantic representation. The assumption set consists of assumptions that have been introduced during the derivation, and must be discharged before the derivation is complete. The assumption set "represents constraints on how the sense may be further connected to its context." [22]

The process of interpretation is defined by a set of *structural rules* and a set of *discharge rules*. Structural rules build the conditional interpretation of a phrase compositionally, from the conditional interpretation of its parts. Discharge rules remove assumptions. In principle all rules have an input and output discourse model, but only the discharge rules actually interact with the discourse model.

The form of a structural rule is

$$P \sim A\text{:}s \text{ if } P_1 \sim A_1\text{:}s_1 \text{ and } ...\text{and } P_k \sim A_k\text{:}s_k$$

The $\sim$ denotes the interpretation relation between a node of a syntactic analysis tree (produced by the parser) and a node of a semantic derivation tree. P denotes a syntactic node, where its immediate constituents are denoted by variables P_1 through P_k. The rule schema is to be understood as stating a constraint that P receives the interpretation A:s if it has constituents P_1 through P_k, and these constituents have the interpretations indicated.

The form of a discharge rule is

$$P \sim A'\text{:}s' \text{ if } P \sim A\text{:}s$$

Here, $A' = A - \{R\}$, where R is the discharged assumption. The discharge of R, together with the current state of the discourse model, determines some modifications to s, resulting in s'.

The assumption storage mechanism is based on Cooper storage [3], which was applied to quantifier phenomena. In the II system, this mechanism is applied to several additional phenomena. Below, I will

describe the rules for pronominal anaphora and for quantifiers.

2.1 Rules for Pronominal Anaphora

The treatment of pronominal anaphora in the II system is similar to the approach in Discourse Representation Theory([13], [12]): indefinite NP's introduce new elements in the discourse model. Pronouns and definite descriptions find their referent among elements in the discourse model.

Four types of referential NP's are defined: pronouns, definite descriptions, indefinites, and names. They are represented as follows:

bind(x,pronoun,number/gender): x

bind(x,def,sort): x

bind(x,indef,sort): x

bind(x,name,N): x

In each case, the sense is represented by a parameter x, and a binding assumption expresses constraints on the way x will be replaced by an entity in the discourse model. This is achieved by discharging the bind assumption. The discharge rules are:

A, bind(x,pronoun,number/gender): S $\Rightarrow$ A: S[x/e]

A, bind(x,def,sort):S $\Rightarrow$ A: S[x/e]

A, bind(x,indef,sort):S $\Rightarrow$ A: S[x/e]

A, bind(x,name,N): S $\Rightarrow$ A: S[x/e]

In the case of pronouns and definite descriptions, the element e must be a salient element in the input discourse model, satisfying the constraints expressed in the binding assumption. An indefinite assumption causes a new element e to be added to the output discourse model. In each case, e is substituted for each occurrence of x in the sense S. At least for pronouns, there is a second possibility: instead of selecting e from the discourse model, some other, undischarged parameter can be selected. This allows a pronoun to be bound by a quantifier, as described below.

2.2 Rules for Quantifiers

The treatment of quantifiers in the II system essentially duplicates that of Cooper[3]. A quantified NP is represented by storing a quantifier assumption, together with a parameter representing the sense. At some later stage in the derivation, the quantifier assumption is discharged, determining the scope of the quantifier. There are two general rules for quantifiers, governing the introduction and discharge of quantifier assumptions. A quantified NP is represented as:

bind(x,q,n): x

where x is a parameter, q is the quantifier, and n is the common noun. For example, "every jet" is represented

bind(x,every,jet): x

Simplifying slightly, the discharge of quantifier assumptions can be represented as follows:

bind(x,q,s): $p_t \Rightarrow$ (q s x) p

As an example,

bind(x,every,jet): fly(x) $\Rightarrow$ (every jet x) fly(x)

As mentioned above, when a pronoun assumption is discharged, its parameter is replaced either by an entity in the discourse model, or by some, yet undischarged parameter. A pronoun becomes "bound" by a quantifier if the quantifier parameter replaces the pronoun parameter in this way.

3 The Account of VP Ellipsis

I now describe a semantic account of VP ellipsis in terms of some simple extensions to the II system. The approach parallels the above approach to pronominal anaphora. I will define a rule to add VP-meanings in the discourse model, and a rule for recovering those VP-meanings to resolve an elliptical VP. Thus full VP's are analogous to indefinite NP's, in that they both typically introduce semantic objects into the discourse model, and elliptical VP's are analogous to pronouns, in that their interpretation requires the selection of an appropriate object from the discourse model. The discourse model will have two sets: SE, the set of salient entities, and SP, the set of salient predicates.

To add VP-meanings to the discourse model, I allow all lexical verbs to introduce an assumption which adds the VP-meaning to the discourse model. I call this binding assumption type "pred". It is discharged as follows:

A, bind(pred):S $\Rightarrow$ A: S

where

$$DM_{out} (SP) = DM_{in} (SP) \cup \{A:S\}$$

That is, the discharge results in the semantic representation of the VP (i.e.,the assumption-sense pair A:S) being added to the SP set of the output discourse model.

I add the requirement that all arguments except the subject must be filled before the assumption is discharged. That is, the discharge of this assumption is permitted only if the sense is of the form

$$P(\text{SUBJ}, a_1, ..., a_n)$$

where SUBJ represents an unfilled subject argument position, with the remaining arguments a_1 through a_n filled.

The assumption for recovering a VP-meaning is introduced by a lexical auxiliary verb; this assumption is termed "epred", for elliptical predicate.

The discharge rule is:

$$\text{bind(epred): AUX} \Rightarrow A{:}S$$

where $A : S$ is some element of the SP set in DM_{in}. That is, upon discharge of the epred assumption, an auxiliary verb is replaced by some VP-meaning in the input discourse model.

The crucial point in these rules is that the antecedent VP is represented as an assumption-sense pair, since it is the assumptions that represent dependencies on context. For example, the representation of the VP "help him" might be

$$\text{bind(x,pronoun,male): help(SUBJ,x)}.$$

This expresses the constraint that the object position must be filled by some entity in the discourse model according to constraints of pronominal reference. Two copies of this VP, as antecedent and target in VP ellipsis, could allow the pronoun to refer to different entities, depending on the state of the current discourse model.

4 An Example

I describe the derivation of example (2), which is repeated here.

(2)   Every boy$_i$ in Bill's class hoped Mary would ask him$_i$ out, but a boy$_j$ in John's class actually knew that she would. [ask him$_j$ out]

The derivation is displayed in Figure 1[4], in the form of a simplified derivation tree. The derivation tree is defined as follows: each node contains a conditional interpretation, a current discourse model, and a derivation rule R, such that the node is consistent with the application of R to the node's daughters. For brevity, the discourse models are not displayed, and only certain rules are labeled.

The antecedent VP "ask him out" is represented as

bind(pred),
    bind(y,pronoun,male): ask-out(SUBJ,x).

---

[4]Note: it is assumed that the auxiliary verb contributes tense and polarity. This contribution is ignored in the derivation for the sake of simplicity.
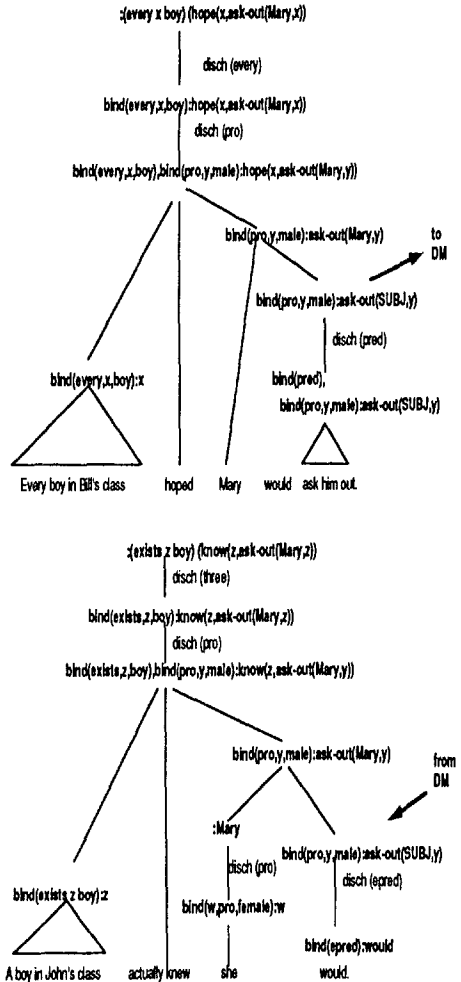

Figure 1: Derivation of Example (2)

The discharge of the pred assumption results in

bind(y,pronoun,male): ask-out(SUBJ,x)

being added to the discourse model. Later, the binding assumption for the pronoun is discharged, allowing it to be bound by the quantifier every boy. [5] In the interpretation of the elliptical VP, the auxiliary "would" is represented

bind(epred):would

The discharge of the epred assumption results in the selection of a VP-meaning from the current discourse model: in this case,

bind(y,pronoun,male):ask-out(SUBJ,x)

is selected. Later, the binding assumption for the pronoun is discharged, allowing the pronoun to be bound to "a boy".

This example receives a straightforward derivation in the system I am proposing. In other accounts, it appears that examples of this sort cannot be accommodated. It is clear, for example, that the example violates the alphabetic variance condition imposed in the Sag/Williams approach. This condition requires that the antecedent and target VP's be identical up to renaming of bound variables. In the example, the antecedent VP contains a free variable (him_i) which becomes him_j in the target, violating the alphabetic variance condition. Partee and Bach[21] adopt essentially the same alphabetic variance condition, as does Klein[15], so that their accounts also rule out the example. Lappin's[17] account explicitly rejects the alphabetic variance condition, replacing it with the following condition:

> For any two occurrences α and β of an open VP intension ϕx_i, α can serve as the antecedent of β iff both occurrences of x_i can be naturally interpreted as having the same intended range of possible values.

The example clearly violates this condition: since him_i ranges over members of Bill's class, and him_j picks out a member of John's class, they do not have "the same intended range of possible values".

Next, I consider the higher order matching approach of Dalrymple, Shieber and Pereira[4] (DSP). In this approach, a matching operation is performed to solve an equation, in which a second order variable represents the elliptical VP. To set up the equation, it is

necessary to determine the antecedent clause and the "parallel elements", and DSP provide no method for making this determination. Typically, with VP ellipsis, there are two adjacent clauses, in which the second clause contains an elliptical VP. Then the first clause is the antecedent clause and the two subjects are the parallel elements. Applying this to the current example, we have "Mary would ask him_i out" as the antecedent clause, and "Mary" and "she" as parallel elements. The equation to solve is

$$P(Mary) = \text{ask-out}(Mary, him_i)$$

In this case, the desired solution,

$$\lambda x.\text{ask-out}(x, him_j)$$

is not a possible solution to this equation, according to the matching operation used by DSP. This is the most straightforward method of determining parallelism to set up the equation, and it does not permit the derivation of the desired reading. However, it may be that an extended notion of parallelism might solve the problem. While this has not been investigated by DSP, such an approach has been advocated in another recent account, that proposed by Prüst et al([23],[24]). It appears that this account can accommodate the example, based on Prüst et al's requirement that if a pronoun p is bound to Q in the antecedent, the corresponding pronoun p' must be bound to a "structurally parallel" Q' in the target, where this is intended as matching syntactic and semantic structure. However, example (3) indicates that the two quantifiers need not be in structurally parallel positions. Indeed, example (5) shows that there is no requirement for a corresponding quantifier at all.

(3) Almost every boy_i in the class hope Mary will ask him_i out, but I know there are a few $boys_j$ who hope that she won't. [ask him_j out]

(4) Every boy_i in Mrs. Smith's class hoped she would pass him_i. In John's$_j$ case, I think she will. [pass him_j]

Examples (1) - (4) illustrate the flexibility required in interpreting pronouns within the antecedent and target VP's. I have shown how the proposed approach permits this flexibility. None of the alternative accounts discussed can accommodate these examples.

5 Constraints on Selecting an Antecedent

I have argued that the current approach has significant advantages over other approaches to problem (2), concerning the level of representation at which VP ellipsis

---

[5]The quantifier "every boy in Bill's class" is represented in the derivation as "every boy", for the sake of brevity. Similarly for "a boy in John's class".

is resolved. In addition, this approach suggests some possible ways of addressing problem (1), concerning the selection among alternative potential antecedents. Since the approach parallels the treatment of pronominal anaphora, storing semantic representations of both VP's and NP's in the discourse model, a natural hypothesis is that similar constraints govern the selection of an antecedent in both the pronominal and the VP ellipsis cases.

The problem of selecting among alternative VP antecedents has been virtually ignored in the literature.[6] The corresponding problem in pronominal anaphora has received a significant amount of attention. The Centering model ([6],[7],[1]) of pronominal anaphora is a leading example, applying a variety of constraints dealing with such factors as recency, salience, and attention. In addition, it is generally agreed that there are syntactic configurational constraints governing pronominal anaphora of the sort described in the "Binding Theory" of GB[2].

For each of these types of constraints, there are interesting parallels with the case of VP ellipsis.

Syntactic/configurational: It appears that VP ellipsis obeys the "precede and command" constraint, as pointed out by Jackendoff[8], ruling out examples such as

(5)   * Charlie will, if his mother-in-law doesn't leave town.

Recency: Just as in the pronominal case, the vast majority of cases involves an antecedent in the current or immediately preceding utterance. In a survey of VP ellipsis in the Brown Corpus[10], I found this to be true about 95% of the time.

Salience: VP's in less salient positions seem to be less available as antecedents for VP ellipsis. For example, Halliday and Hasan[9] give the following example:

(6)   A: The policeman paid no attention to the girl who was driving the car.

(7)   *B: Was she really?

Presumably the unavailability of the VP "driving the car" is related to the fact that it appears in a restrictive relative clause and is thus not particularly salient.

Attentional: There is evidence that a "center shift", i.e., shifting attention from one entity to another, might be correlated with the availability of VP antecedents. This is suggested by experimental work of Malt[19], who describes experiments that show, in her terms, that "changing the focus" in an intervening

sentence tends to make a previous VP antecedent inaccessible. Thus the first example (taken from Malt's experiment) was understood more readily than the second:

(8)   a. "I liked the Monet exhibit," Heather remarked.

       b. "It was really very nice".

       c. "I did too," Richard replied.

(9)   a. "I liked the Monet exhibit," Heather remarked.

       b. "Renoir is my favorite, though."

       c. * "I did too," Richard replied.

There is reason to believe, then, that constraints relating to factors such as recency, syntactic configuration, salience, and attention, might apply to VP ellipsis and pronominal anaphora in a similar way. A simplified version of these constraints is implemented in the pronoun case of the Incremental Interpretation System, and it would a simple matter to allow the same constraints to apply to VP antecedents.

6 Conclusions

While it has been argued by many that VP ellipsis is a semantic phenomenon, there is no existing account that satisfies some standard requirements on semantic representation, relating to the "dispensability" of the meaning-representation language, and the incorporation of contextual dependencies in semantic representations. In addition, existing semantic accounts have important empirical problems, not allowing pronouns to switch reference from antecedent to target with sufficient flexibility. The modifications necessary to comply with the standard requirements on semantic representation are exactly the ones needed to solve these empirical problems. I have described such a semantic account, showing that it handles the examples that are ruled out by alternative semantic accounts. The approach is easily implemented computationally, by some simple extensions to the Incremental Interpretation System. In addition, there is evidence indicating that the selection of a VP antecedent might be subject to the same sort of constraints that govern pronominal anaphora.

7 Acknowledgements

References

[1] Susan E. Brennan, Marilyn Walker Friedman, and Carl J. Pollard. A Centering Approach to Pronouns, *Proceedings of the 25th Annual Meeting of the ACL*, 1987.

[2] Noam Chomsky. *Lectures on Government and Binding*. Foris. 1981.

[3] Robert Cooper. *Quantification and Syntactic Theory*. D. Riedel, Dordrecht, Netherlands. 1983.

[4] Mary Dalrymple, Stuart Shieber and Fernando Pereira. Ellipsis and Higher-Order Unification. *Linguistics and Philosophy*. Vol. 14, no. 4, August 1991.

[5] Robert Fiengo and Robert May. Ellipsis and Anaphora. Paper presented at GLOW 1990, Cambridge University, Cambridge, England.

[6] Barbara Grosz, Aravind Joshi, and Scott Weinstein. Providing a Unified Account of Definite Noun Phrases in Discourse. In *Proceedings, 21st Annual Meeting of the ACL*, pp. 44-50, Cambridge, MA, 1983.

[7] Barbara Grosz, Aravind Joshi, and Scott Weinstein. Towards a Computational Theory of Discourse Interpretation. ms. 1986.

[8] R.S. Jackendoff. *Semantic Interpretation in Generative Grammar*. Cambridge, Mass. MIT Press. 1972.

[9] M.A.K. Halliday and R.H. Hasan. *Cohesion in English*. London: Longman, 1976.

[10] Daniel Hardt. A Corpus-based Survey of VP Ellipsis, ms. University of Pennsylvania, 1990.

[11] Daniel Hardt. A Discourse Model Account of VP Ellipsis. Proc. ESCOL 1991. Baltimore, MD.

[12] Irene Heim. *The Semantics of Definite and Indefinite Noun Phrases*. Ph.D. thesis, University of Massachusetts-Amherst. 1981.

[13] Hans Kamp. A Theory of Truth and Semantic Representation. In Groenendijk, J, Janssen, T.M.V. and Stokhof, M. (eds.) *Formal Methods in the Study of Language*, Volume 136, pp. 277-322. 1980.

[14] Yoshihisa Kitagawa. Deriving and Copying Predication. ms. University of Rochester, 1990.

[15] Ewan Klein. VP Ellipsis in DR Theory. In J. Groenendijk, D. de Jongh and M. Stokhof, eds. *Studies in Discourse Representation Theory and the Theory of Generalized Quantifiers*, Foris Publications. Dordrecht, The Netherlands. 1987.

[16] Ewan Klein, Kate Stainton-Ellis. A Note on Multiple VP Ellipsis. *Linguistics* 27, 1119-1124. 1989.

[17] Shalom Lappin. VP Anaphora, Quantifier Scope, and Logical Form. *Linguistic Analysis*. Volume 13, No. 4, pp. 273-315. 1984.

[18] Shalom Lappin and Michael McCord. Anaphora Resolution in Slot Grammar, in *Computational Linguistics*, vol 16, no 4. 1990.

[19] Barbara Malt. The Role of Discourse Structure in Understanding Anaphora. *Journal of Memory and Language, 24*, 271-289. 1984.

[20] Richard Montague. *Formal Philosophy*. ed. Richmond H. Thomason, Yale University Press, 1974.

[21] Barbara Partee and Emmon Bach. Quantification, Pronouns, and VP Anaphora, in *Truth, Interpretation and Information*, Dordrecht 1984.

[22] Fernando Pereira and Martha Pollack. Incremental Interpretation. *Artificial Intelligence*. Vol. 50. no. 1, pp. 37-82. June 1991.

[23] Hub Prüst and Remko Scha. A Discourse Approach to Verb Phrase Anaphora. Proc. ECAI, 1990.

[24] Hub Prüst, Remko Scha, and Martin van den Berg. A Formal Discourse Grammar tackling Verb Phrase Anaphora. ms. 1991.

[25] Ivan A. Sag. *Deletion and Logical Form*. Ph.D. thesis, University of Massachusetts-Amherst. 1976.

[26] Peter Sells. Anaphora with Which. West Coast Conference on Formal Linguistics, vol 4. 1985.

[27] Bonnie Lynn Webber. *A Formal Approach to Discourse Anaphora*. Ph.D. thesis, Harvard University. 1978.

[28] Edwin Williams. *Discourse and Logical Form*. *Linguistic Inquiry*, 8(1):101-139. 1977.