AI-TR 537

# TOWARDS A COMPUTATIONAL THEORY OF DEFINITE ANAPHORA COMPREHENSION IN ENGLISH DISCOURSE

CANDACE LEE SIDNER

# Towards A Computational Theory of

# Definite Anaphora Comprehension in English Discourse

by

**Candace Lee Sidner**

**Massachusetts Institute of Technology**

**June, 1979**

c Candace Lee Sidner, 1979

This report is a revised version of a dissertation submitted to the Department of Electrical Engineering and Computer Science on May 16, 1979 in partial fulfillment of the requirements for the Degree of Doctor of Philosophy.

This report describes research done at the Artificial Intelligence Laboratory of the Massachusetts Institute of Technology. Support for the laboratory's artificial intelligence research is provided in part by the Advanced Research Projects Agency of the Department of Defense under Office of Naval Research contract N00014-75-C-0643. Research in this report was also conducted in part at SRI International and was supported by the National Science Foundation under Grant No. MCS 76-22004, and by Advanced Research Projects Agency of the Department of Defense under contract N00039-78-C-0060 with the Naval Electronic Systems Command.

## Abstract

This report investigates the process of focussing as a description and explanation of the comprehension of certain anaphoric expressions in English discourse. The investigation centers on the interpretation of definite anaphora, that is, on the personal pronouns, and noun phrases used with a definite article *the, this,* or *that.*

Focussing is formalized as a process in which a speaker centers attention on a particular aspect of the discourse. An algorithmic description specifies what the speaker can focus on and how the speaker may change the focus of the discourse as the discourse unfolds. The algorithm allows for a simple focussing mechanism to be constructed: an element in focus, an ordered collection of alternate foci, and a stack of old foci. The data structure for the element in focus is a representation which encodes a limited set of associations between it and other elements from the discourse as well as from general knowledge.

This description of focussing allows the following hypothesis of anaphora comprehension to be stated and supported. Definite anaphora are signals which the speaker uses to tell the hearer what element in the discourse is the current discourse focus; at the same time, the element in focus constrains which anaphoric expressions can be used to signal the focus. This hypothesis is supported by five results which are presented in this report:

-a means for distinguishing definite noun phrase used anaphorically from those used non-anaphorically.

-a means for distinguishing pragmatic anaphora from bound variable and inter-sentential anaphora.

-rules which use the focussing mechanism for the interpretation of pragmatic anaphora.

-reduction of the search for inferences which support the interpretation chosen for an anaphor.

-a data structure which represents the element in focus and indicates which items can be associated with the focus and which phrases can be used to mention those items.

This report also establishes other constraints which are needed for the successful comprehension of anaphoric expressions. The focussing mechanism is designed to take advantage of syntactic and semantic information encoded as constraints on the choice of anaphora interpretation. These constraints are due to the work of language researchers; and the focussing mechanism provides a principled means for choosing when to apply the constraints in the comprehension process.

To My Father
who first taught me the joy of mathematics
and the wonder of language

**Acknowledgements**

I would like to express my gratitude to everyone who contributed to this work, and provided support and encouragement to its author:

-to Jonathan Allen, my advisor, for good advice, for patience taken to understand ideas in progress, especially for prodding me to be clear in my thinking and writing.

-to my readers: Barbara Grosz, for many hours of discussion on the contents of this work as well as its form, and for encouragement and understanding of the nature of this work, and Peter Szolovitz, for careful review and many insightful questions.

-to "The Natural Language Group" in the AI Lab, Bob Berwick, Beth Levin, Mitch Marcus, Dave McDonald, and Kurt VanLehn, for insightful discussions, as well as for helping me to get an education; I also thank Bob Berwick and Ellen Hildreth for extensive editorial critique on drafts of this report.

-to "The Natural Language Group" in the AI Center at SRI International, Barb Grosz, Gary Hendrix, Jerry Hobbs, Bob Moore, Ann Robinson, and Jane Robinson for providing me with tools to test out my ideas and with the opportunity to share the concepts of my research.

-to the MIT Artificial Intelligence Lab for numerous fine tools, people with interesting ideas and a place to work. In particular, thanks to Ira Goldstein for encouraging the PAL project out of which this report grew.

-to my friends, many mentioned above, and especially to Chuck, Bruce, Edwina, Eric, Ellen, Amy, Peter, Rita, Jawaid, Martha, Harold, Bob, Helen, for everything from participating in long hours of technical discussions, to providing consolation and support when I despaired, and to joining in celebration. And to my family for continued encouragement and support.

# TABLE OF CONTENTS

**1. Introduction to the Problem of Anaphora Comprehension in English ..... 12**

1.1 Statement of the Problem from a Computational View ..... 12
1.2 A Definition of Antecedence ..... 14
1.3 Approach of This Report ..... 17
1.4 Artificial Intelligence Research on Anaphora ..... 24
1.5 Approaches to anaphora in linguistics ..... 27
1.6 Focus, Topic and Comment ..... 34
1.7 Discourse Approaches to Anaphora ..... 38
1.8 Some Recent Psychological Work on Anaphora ..... 41
1.9 Focus in Discourse ..... 44
1.10 Guide to Remaining Chapters ..... 48

**2. The Definition of Focus ..... 50**

2.1 Focus and Discourse ..... 50
2.2 Focussing for Definite Anaphora Comprehension ..... 52
2.3 The Representation of Focus ..... 55
2.4 The Focussing Process: Finding The Discourse Focus ..... 59
2.5 Examples of the Use of the Expected Focus Algorithm ..... 69
2.6 Rejecting the Default Focus ..... 70
2.7 Inferencing in Focussing ..... 73
2.8 Focus Sets ..... 76
2.9 An Algorithm For Focussing ..... 77
2.10 The Focussing Algorithm for Focus Recognition: Example Dialogues ..... 81
2.11 Focus Movement ..... 83
2.12 Backwards Focus Movement ..... 87
2.13 Syntactic Constructions for Focus Movement ..... 90
2.14 The Focussing Algorithm for Focus Movement: Example Dialogues ..... 91
2.15 Conclusions ..... 95

**3. Comprehension of Anaphoric Definite Noun Phrases ..... 97**

3.1 Co-specification and Definite Noun Phrases ..... 97
3.2 The Nature of Non-Anaphoric Definite Noun Phrases ..... 100
3.3 How Anaphoric Defnps Co-specify ..... 106
3.4 Implicit Backwards Specification ..... 112
3.5 An Algorithm for Interpreting Defnps in Discourse ..... 122
3.6 Distinguishing Generics ..... 125
3.7 Focussing and Generic Specification ..... 136
3.8 Conclusions ..... 139

**4. Focus in Pronoun Anaphora Resolution ..... 142**

4.1 Focussing and the Characteristics of Pragmatic Anaphora ..... 142
4.2 The Basic Rules ..... 144
4.3 An Actor Focus ..... 152
4.4 Focus and Sentence Syntactic Restrictions ..... 156
4.5 Examples of the Focussing Algorithm for Actor and Discourse Focus 160
4.6 Rules for Possessives ..... 165
4.7 More About the Representation of Focus ..... 168
4.8 Restrictions on Co-specification Using Focus ..... 173
4.9 Pronouns Which Have No Co-specifications ..... 176
4.10 The Problem of Parallelism ..... 179
4.11 Comparison with Other Recent Computational Algorithms ..... 183
4.12 Conclusions ..... 187

**5. Co-present Foci in Anaphora Disambiguation ..... 189**

5.1 The Occurrence of Co-present Foci in Discourse ..... 189
5.2 Rules for Co-present *the one...the other* ..... 191
5.3 Movement with Co-present Foci *The one...the other* ..... 193
5.4 Exophoric Demonstrative Uses of *This* and *That* ..... 196
5.5 Use of *This* and *That* for Co-present Foci ..... 196
5.6 *This* and *That* in Focus Movement ..... 203
5.7 Using the Focus Movement Algorithm ..... 213
5.8 *This* Noun Phrases Used Non-Anaphorically ..... 215
5.9 Conclusions about *This* and *That* ..... 216

**6. Some Experimental Systems Using Focus ..... 218**

6.1 Use of Focus in Discourse Understanding ..... 218
6.2 Focussing in the PAL System ..... 219
6.3 Focussing in the Task Dialogue Understanding System ..... 225
6.4 Conclusions ..... 232

**7. Summary and Conclusions ..... 234**

7.1 The Relation of Focussing to Anaphora Comprehension ..... 234
7.2 Limitations of Focussing in Anaphora Comprehension ..... 236
7.3 Further research ..... 237

**8. List of Cited Works ..... 240**

**Appendix I. Algorithms ..... 248**

**Appendix II. Co-specification and Reference ..... 260**

## TABLE OF FIGURES

Fig. 1.1. Reference Links Between Noun Phrases ..... 15
Fig. 1.2. Co-specification for indefinite noun phrases and definite anaphors ..... 17
Fig. 1.3. Instances of a General Meeting Concept ..... 20
Fig. 1.4. NP1 Precedes and Kommands NP2 ..... 28
Fig. 1.5. Task Representation for Box Assembly ..... 46
Fig. 2.1. Instances of a General Meeting Element ..... 56
Fig. 2.2. Control Flow for Focussing Algorithm ..... 78
Fig. 2.3. Action of Focus Mechanism for the Start of D25-2 ..... 81
Fig. 2.4. Focussing At D26-2 ..... 82
Fig. 2.5. Stack of foci at D31-8 ..... 89
Fig. 2.6. Focus mechanism with full database representation of focus ..... 94
Fig. 3.1. Contents of Sample Knowledge Network ..... 107
Fig. 3.2. Representation of network for discourse D13 ..... 115
Fig. 3.3. Generic and Specific Representations of D21 ..... 137
Fig. 4.1. Rule for Third Person Pronoun in Agent Position ..... 147
Fig. 4.2. Discourse/Actor Conditions for Agent Third Person Pronoun ..... 148
Fig. 4.3. Rule for Third Person Pronoun in Non-Agent position ..... 149
Fig. 4.4. Anaphora Co-specifications for D14-1 and 2a ..... 161
Fig. 4.5. Determining Actor and Discourse Foci of D14-1 and 2a ..... 163
Fig. 4.6. Determining the Anaphora Co-specifications in D14-1 and 2b ..... 164
Fig. 4.7. Rule for Third Person Personal Possessive Pronouns ..... 166
Fig. 4.8. Intensional and Extensional Specifications of *it* in D24-2 ..... 171
Fig. 4.9. Structural tree of Lockman expansions and continuations ..... 186
Fig. 5.1. Rules for Interpretation of *the one...the other* ..... 192
Fig. 5.2. Co-present focus links for D2 ..... 194
Fig. 5.3. Co-specification for co-present *this* and *that* ..... 202
Fig. 5.4. Interpretation of *This* Noun Phrases ..... 211
Fig. 5.5. Interpretation of *that* Noun Phrases ..... 212
Fig. 6.1. Modules of PAL ..... 220

Fig. I.1. Control Flow for Focussing Algorithm ..... 251
Fig. I.2. Rule for Third Person Pronoun in Agent Position ..... 253
Fig. I.3. Discourse/Actor Conditions for Agent Third Person Pronoun ..... 254
Fig. I.4. Rule for Third Person Pronoun in Non-Agent position ..... 255
Fig. I.5. Rule for Third Person Personal Possessive Pronouns ..... 256
Fig. I.6. Rules for Interpretation of *the one...the other* ..... 257
Fig. I.7. Interpretation of *This* Noun Phrases ..... 258
Fig. I.8. Interpretation of *that* Noun Phrases ..... 259
Fig. II.1. Reference Links Between Noun Phrases ..... 260
Fig. II.2. Co-specification for indefinite noun phrases and definite anaphors ..... 263

# 1. Introduction to the Problem of Anaphora Comprehension in English

## 1.1 Statement of the Problem from a Computational View

When speakers utter or write sentences, they use certain words to refer to people, places, objects, times, events and ideas which exist in the real world or in some possible world. When sentences are formed into units of two or more sentences, certain words "point back" to other expressions in the previous sentences. Those words which are used to point back are called anaphoric expressions; the words pointed to traditionally are called the antecedents. Recent studies of natural language in artificial intelligence and linguistics has shown that the comprehension of anaphoric expressions involves more than just that expression and the words it points to. One goal of this chapter is presenting a definition of antecedence which provides an adequate foundation for a study of anaphora comprehension.

Two other goals will be accomplished in this chapter. First, a brief overview of the difficulties in anaphora comprehension will be given. The overview is followed by a discussion of the approach which this report will give, together with the claims which will be supported in the chapters that follow. Second, research relevant to comprehension of antecedence will be reviewed. Before either past approaches to anaphora or the approach of this report can be considered, our intuitive understanding of antecedence must be brought into a perspective which leads to a definition of antecedence. That perspective can be established best by understanding the difficulties in anaphora comprehension. For the moment, then, let us look at the puzzle of choosing the antecedent of an anaphor.

The antecedent of an anaphor is difficult to determine because there may be a multitude of possible antecedents in a given discourse. Despite the old grammar school adage that a pronoun refers to the last noun phrase mentioned of the same gender and number, the class of possible antecedents of an anaphor is large. Yet human speakers generally can recover the antecedent intended by the speaker. The task is first to propose the class, pare it down, and then choose the intended antecedent.

Other difficulties must be faced. Anaphora interpretation requires the use of the context surrounding the anaphor. The context can be simply the sentence in which

the anaphor appears, or it can be one of several preceeding sentences. For example, the antecedent intended for "the class" in the fourth sentence of the last paragraph is the class mentioned in the second sentence. Why does context introduce difficulties? A context has a structure which must be characterized if the antecedent is to be found. A brief inspection of any text will indicate that its context is very rich and the possible structures greatly varied. So the choice of which elements of a context to include in the structural description has challenged many researchers.

The richness of discourse is not the only difficulty in posing a structure for discourses. Whatever structure is proposed must be appropriate for use by a process which determines antecedents of anaphora. This work fundamentally assumes that anaphora interpretation is a process. Since processes have a complexity of their own, there may be two levels of complexity. One method for limiting complexity demands simplicity in the data structures which a process uses; at least sufficient simplicity to be stated in theoretical terms. Thus the context must be distilled into a structure that preserves the needed richness and yet does not add complexity to the process of anaphora interpretation.

Finally, anaphora interpretation involves making inferences. Some of the inferences appear trival: "the class" cannot be used to refer to any uses of "the antecedent" in the paragraph above, (as any human can reason) because antecedents are not the same things as classes. But such an inference requires knowledge, even if just a test for equality. It also must be chosen, from a large collection of inferences, as the proper one to make at that point in the discourse. Computational use of inference, as many artificial intelligence researchers have discovered, requires control: knowing which inference to make when. Since controlling inference is by no means well understood in general, solutions to more specific situations, such as anaphora interpretation, must be sought.

Proposals for the solution of these problems have been offered over the past ten years by linguists and artifical intelligence researchers studying natural language. This report claims that representations of knowedge, inference mechanisms and a structure for context can be joined together to form a theory of anaphora comprehension. Simply said, they are joined by means of focussing on what the conversation, or more generally, what the discourse, is about. By focussing on an aspect

of the conversation, the speaker makes clear what is being discussed and can either mention it again, refer to related or more general terms, or shift the conversation to some other item in the discourse. The item which is focussed on, when described computationally, serves as an index into the discourse structure as well as a means of controlling which entities inferences will be made about. How this provides a theory of comprehension, given the problems of a multitude of possible antecedents, context structure and control of inference, is the topic of this report.

## 1.2 A Definition of Antecedence

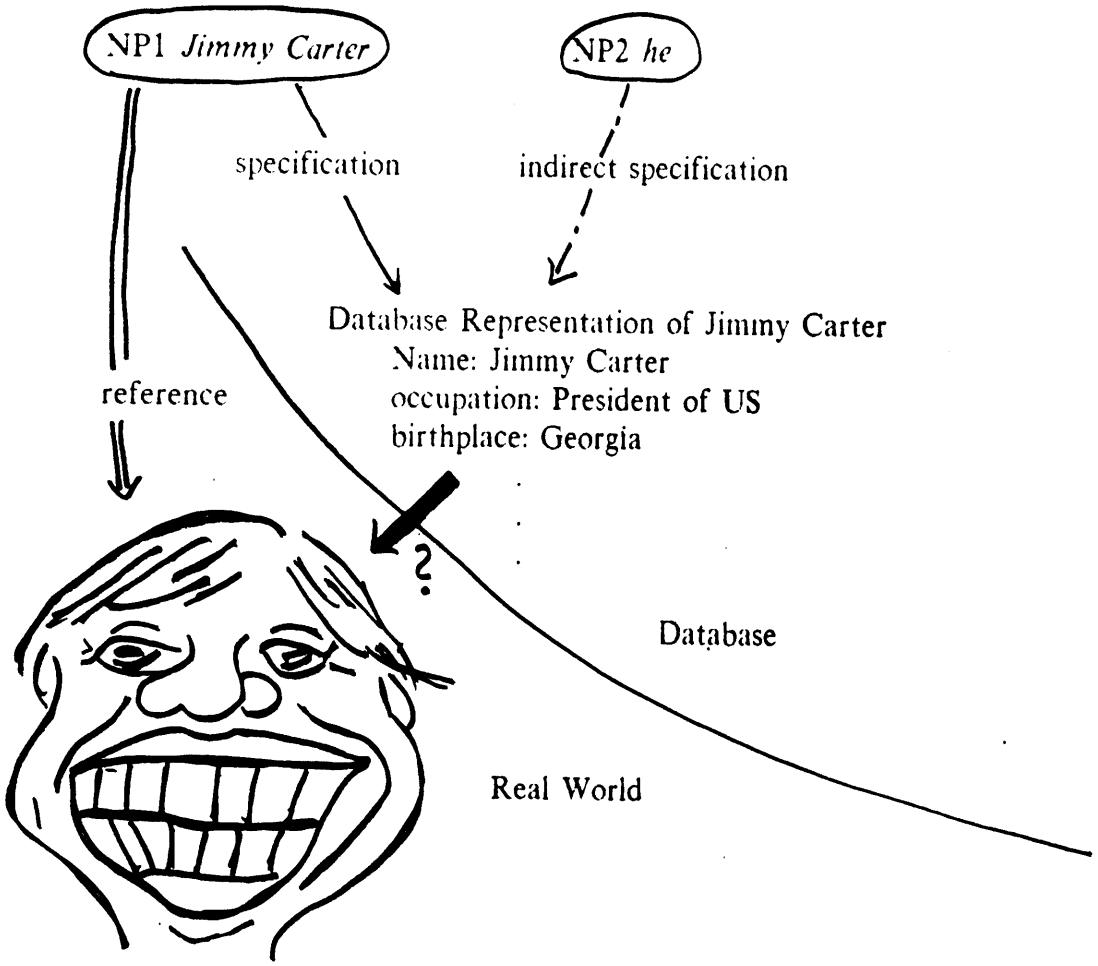
When we informally talk of noun phrases, we call them "referring expressions." In fact, noun phrases can be distinguished by two different uses. First, a noun phrase may be used refer to some entity in the real, or some imaginary, world. The noun phrase *Jimmy Carter* can be used to refer to Jimmy Carter, who is the current president of the United States. In figure 1 below, this relationship is expressed by the reference arrow. Names are a common means of referring, and definite anaphora[1] are another. If a speaker talks about Jimmy Carter, the speaker may continue to speak using the pronouns *he* or *him*, or a definite noun phrase such as *the President.*

Rather than consider the relation between a string of words and entities in the real world, I will define some relations between words, their interpretation in the sentence in which they occur, and elements of a database. Informally, I call the noun phrase and its interpretation based on syntax and semantics the bundle of a noun phrase. Specification is the relation between a noun phrase, including its syntactic and semantic interpretation in a sentence, and some database object. In other words, the bundle of a noun phrase specifies a database element which represents the real world object. The first use of *Jimmy Carter* specifies the database element shown in figure 1 as the database representation of Jimmy Carter. The database object is a representation of the real world person Jimmy Carter. By analogy to reference one can ask how the database object which is the specification refers to the real world. This question is, to my knowledge, an open question in the theory of cognitive science and will remain open

---

1. Definite anaphora include the personal pronouns, defnps used anaphorically and noun phrases containing *this* and *that.*

**Figure 1.1. Reference Links Between Noun Phrases**




---

in this report; the relationship in question is depicted by the filled arrow, which is labelled with a question mark, between the database representation and the person Jimmy Carter.

What can be said about the definite anaphor *he*? The bundle of the noun phrase *Jimmy Carter*, and its specification can be used to determine that *he* specifies the same database element as the *Jimmy Carter* bundle. In other words, *he* and its bundle co-specifies with the *Jimmy Carter* bundle to the specification of Jimmy Carter in the database. Because the specification of *he* depends *Jimmy Carter* and its specification, the specification of *he* is obtained indirectly. This indirect specification of *he* is shown in figure 1 by the dot-dash link.

A surprising effect of language bears on the specification relationship; noun phrases which we informally call referring expressions are not always used to refer. They are sometimes used a second way, namely, to construct something which can be talked about. An example will make plain this fuzzy talk.

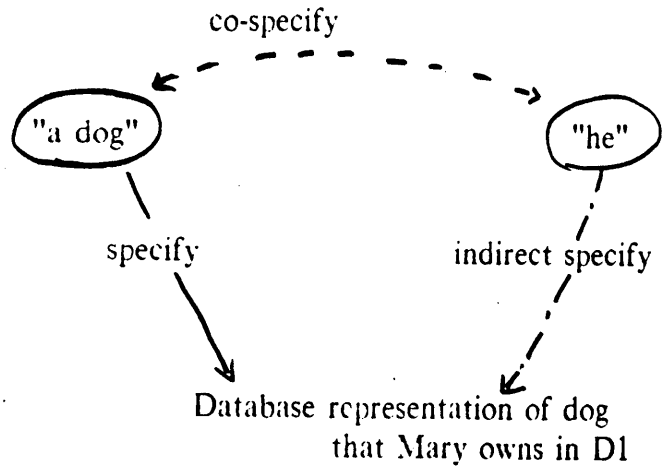
D1-1 Mary has a dog.
2 He's quite friendly
3 because he wags his tail a lot and wants to play.

In D1, the noun phrase *a dog* is not used to refer in the same way as a name, that is, *a dog* does not denote an entity in the world; instead the speaker in saying D1-1 constructs an item in the discourse about which more is said in D1-2 and 3. Informally people talk about the relation between *he* and *a dog* as being one of antecedence, but since *a dog* does not denote; some concept other than reference is needed to define the informal notion of antecedence.

The description of Mary's dog is what people intuitively understand as the antecedent of *he* in D1; the co-specification relation captures exactly this intuition. ***He*** specifies the same element as the bundle of *a dog*, i.e., the dog that Mary owns. In this case, the specification is just a representation of the dog which can be constructed from the discourse context. Thus the bundle of *he* co-specifies with the bundle of *a dog*, to the dog that Mary owns, as shown in figure 2.

To summarize, for noun phrases that are discourse constructs, the specification of the noun phrase bundle is a representation in the database derived from the discourse. A definite anaphor and its bundle will co-specify with the noun phrase bundle to the representation. For noun phrases which denote, co-specification will link the bundle of the referring noun phrase with the bundle of the pronoun. For the case of *Jimmy Carter* in figure 1, NP2 *he* co-specifies with NP1's bundle to the database representation of Jimmy Carter. As long as there is a person to whom the noun phrase *Jimmy Carter* refers and that person is represented in the database, the co-specification relation between an anaphor and an antecedent chooses the represented person. Said another way, the relation of antecedence between two discourse phrase bundles can be defined as the co-specification with relation. Hereafter, to eliminate imprecise terminology, I will drop the use of "noun phrase bundle" in co-specification and speak only of the noun phrase. When I do so, I am implicitly assuming that the syntactic and semantic interpretation of the noun phrase are part of noun phrases which act in a

**Figure 1.2. Co-specification for indefinite noun phrases and definite anaphora**




---

co-specification relation. Appendix II discusses the advantages of defining antecedence as co-specification.

Given these definitions, we are in a position to state the central question to which this report seeks an answer: what noun phrase can be used in the co-specification relation with a definite anaphor, i.e., what noun phrase and accompanying sentence interpretation play the role of *Jimmy Carter* or *a dog*?

## 1.3 Approach of This Report

Suppose that enough structure can be found in a discourse to indicate what it is that the speaker is talking about and to determine when the speaker switches to something else. What the speaker is talking about can be used to decide the co-specification relation of anaphora as long as each anaphor is related in one of several simple ways to what the speaker is talking about. This approach, if possible, offers a solution to the problems of search and inference. Search for a co-specification is minimized because the search centers on what the speaker is talking about, rather than

all the elements mentioned in the discourse. Inferencing is simplified because it is used to confirm the choice of the co-specification search process rather than to direct the search. Let us consider in more detail what such a proposal consists of.

The description of the communication process given here contains four simple assumptions which are generally true and will be taken as true in this work. First, the speaker is assumed to be communicating about something. This assumption implies that the speaker is not speaking gibberish, that there are referring expressions and either requests, questions, assertions or acknowledgements being made. The something which the communication is about will be called the focus of the discourse.[1] Second, the speaker assumes that the hearer can identify the focus of the discourse. The speaker wants to communicate about something, and for the communication to occur, the hearer must be able to distinguish what the speaker is communicating about. Third, the speaker is not trying to confuse or deceive the hearer. The speaker uses referring expressions with the intention of referring to someone or something, or with the intention of describing something or some event. In Gricean (Grice [1975]) terms, the byword is "Be perspicuous." Finally the speaker assumes the hearer has certain knowledge about the real world which can be used to reason about referring expressions during the communication process. Recent research (Cohen [1978], Clark and Marshall [1978]), and the well known work of Searle [1969] and Austin [1962], describe models of the speaker's knowledge of what the hearer believes. In this chapter, the weakest form of such a model is assumed: the speaker assumes the hearer has enough real-world knowledge in common with the speaker to know the entities in the real world which the speaker refers to and to know the discourse elements which the speaker mentions. The speaker draws on that knowledge in constructing a message for a hearer. These four assumptions will play an important part in the discussion of co-specification interpretation which follows.

---

1. Discourse will be defined in chapter 2. For the moment an intuitive use of the word is used. I do not want to suggest that only one thing can be communicated in a discourse, for speakers do direct their attention from one thing to another. For the moment, I will speak of the focus as the first center of attention in a speaker's discourse. As I will use the term technically later, I mean by focus what Grosz calls immediate focus.

This report makes the claim that the focus acts as an index function for referring expressions. For those referring expressions which are anaphoric, the focus indicates where to look for a co-specification. For those referring expressions which are names or descriptions of things outside the discourse, the focus acts as a center point for a process that generates representations of real world entities which fit the name or description and that chooses one such representation. However, the focus of a discourse alone is not sufficient to produce the indexing behavior. The focus must be used in conjunction with a hierarchical network of associations.[1] This network is a codification of some of the general knowledge speakers and hearers have about the real world. It indicates other concepts which are related to the focus. The network is a dynamic structure because the hearer adds to his/her general knowledge in the process of interpreting a piece of discourse. Focus must also act with a third piece of computational machinery, an inferencing mechanism, which is used to infer from general knowledge and other suppositions that a certain proposition is true.

An example will be helpful here. In the discourse below, the focus of discussion is the meeting of D2-1.

D2-1 I want to schedule a meeting with Ira.
2 It should be at 3 p.m.
3 We can get together in his office.
4 Invite John to come, too.

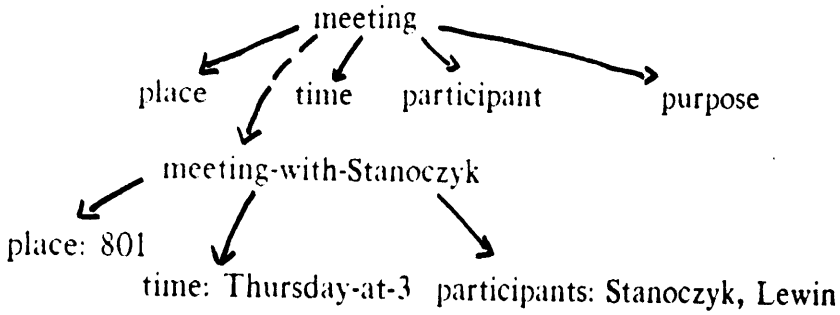
All four sentences give information about the focussed entity. Thus in D2 both sentences 3 and 4 make no direct reference to the meeting of D2-1. As human hearers, we know that these sentences are related to the rest of D2 because they give information about the focus *meeting*. In D2-3 there are three clues which connect this sentence and the rest of the discourse: the use of *get together*, the co-specification of *we* to the participants of the meeting, and *his office* establishing a place for a meeting. D2-4 introduces an additional participant which can be surmised from the use of *invite*, and the fact that the ellipsis of the event that John is invited to is the focus.

A piece of the hierarchical net needed for D2 is given below in Figure 3. A

---

1. On semantic networks see Hendrix [1975]; for frame systems, see Roberts and Goldstein [1977], and Bobrow and Winograd [1977].

**Figure 1.3. Instances of a General Meeting Concept**




---

prototype meeting has associated places, times, participants, and purposes. The relation between meeting and place is one of occurrence while the relation between meeting and purposes is one of causality. When D2-1 is encountered, the hierarchical net grows a new member: an instance of meeting from D2-1. It inherits the associated elements of meetings and some specific values for the participant element. D2-2 indicates that something (called *it*) will occur at a particular time. The focus of D2-1 is meeting, so the focus, meeting, is proposed as the antecedent of *it*. To confirm the proposal, the inference mechanism checks to see if meetings occur at times. Indeed they do, so the proposal of meeting as antecedent of *it* is accepted.

This explanation for the use of focus is not really so simple because the focus of a discourse changes. The interpretation of focus requires a means of recognizing that the focus has changed to some other element of the discourse. In D3 the focus begins on meeting, but the *it* in D3-3 has *my office* as its co-specification, not the meeting. Detecting this co-specification requires a means of noticing a movement of focus and using the inferencing mechanism to confirm the choice of co-specification. Focus movement detection will be described fully in chapter 2.

D3-1 I want to schedule a meeting with George, Jim, Steve and Mike.
2 We can meet in my office.
3 It's kind of small, but the meeting won't last very long anyway.

This report is concerned with the problem of determining which of the many discourse representations built up during sentence comprehension can be used to

determine the antecedent of a definite anaphor. Some of those discourse representations result from the interpretation of noun phrases used to refer to some object in the real or some imaginary world. Such a noun phrase and its sentence interpretation may specify some pre-existing database representation. A definite anaphor co-specifies with such a noun phrase and its sentence semantic interpretation to the database representation. Other discourse representations can be used to fix the specification of the definite anaphor. These representations result from noun phrases which act to construct a discourse element. For such cases, a definite anaphor co-specifies with the noun phrase and its sentence interpretation to a database representation, which is created from the sentence syntax and semantics. That is, the database representation is a discourse representation; it is the only specification there is for noun phrases acting as discourse constructs and the anaphora which co-specify with them.

In addition to the comprehension of pronouns like *it* and definite noun phrase anaphora like *the meeting*, focussing will be shown to provide an explanation for the definite anaphora of *this* and *that*, as used in the discourses below. Since these definite anaphora have received little treatment in any literature, an explanation of their behavior as part of focussing offers new insight into the nature of language.

D4-1 The axon may run for a long distance....
2 Man's longest axon runs for several feet, from the spinal column to muscles that control movements of the toes.
3 In spite of its great length, this axon, like all nerve fibers, is a part of a single cell.
4 It is living matter.

D5-1 I'm having a party tomorrow night;
2 it will be like the one I had last week.
3 That party was a big success
4 because everyone danced.
5 This one will have better food.
6 I've asked everyone to bring something special.
7 Want to come?

This report claims that co-specification comprehension makes use of a control mechanism based on what I have called the focus of the discourse. The mechanism relies on the focus being represented in the hierarchical net mentioned earlier and consults with an inference mechanism. This brief description of a focus and the focussing mechanism raises some questions which must be answered. What is the focus of the discourse, that is, how is it determined? What kinds of asssumptions about the

structure of the knowledge network must be made in order to use a focus for definite anaphor disambiguation? What inferences are used in the prediction of co-specification for pragmatic anaphora? How is the focus used?

This report makes several claims about the nature of definite anaphora:

1. Use of network relationships
   Focus and a knowledge network together determine a set of network relationships among elements of the discourse. These relationships indicate ways in which co-specification with the focus can be accomplished as well as how co-specification with elements associated with the focus can be accomplished.

2. The role of focus in co-specification
   Focus can be used to indicate which noun phrases are used to co-specify within a context and which are not.

3. Focussed inferencing
   Focussing controls the inference mechanism needed to determine an antecedent relationship between a focus and an anaphoric noun phrase because inferencing is used to confirm a hypothesized link between an anaphor and a focus.

4. Distinguishing pragmatic anaphora
   Focus, used with a representation such as Webber's [1978] and with information describing sentence syntactic constraints, such as c-command, can distinguish pragmatic anaphora from bound anaphora and sententially constrained anaphora.

5. Disambiguating pragmatic anaphora
   Focus, used with a representation such as Webber's and with information describing sentence syntactic constraints, such as c-command, can disambiguate the co-specification of pragmatic anaphora.

Each of these claims will be supported in the chapters that follow.

The discussion in this report will be limited to those phenomena affecting focus other than speech stress and prosodics. As will be shown later, stress and prosodics are a significant part of anaphora comprehension in discourse, so this limitation affects the feasibility of the forthcoming algorithm as a real psychological model of human behavior. However, not enough is known about stress and prosodics to discuss these

behaviors in computational terms. When they are better understood, the algorithm could be revised to incorporate their role.

The theory which results from answering these questions provides the details for the interaction of the focus, knowledge network and inference mechanism. Simply, the three processes used together can predict the co-specifications of pragmatic pronominal anaphoric expressions. For co-specification of non-pronominal definite noun phrases, the theory proposes that certain relations are needed in knowledge representation in order to determine the co-specification. Since the theory requires certain kinds of knowledge in the network, it makes predictions about the structure of human knowledge. Both predictions are testable and will indicate the strengths and weaknesses of the theory as a model of comprehension.

This theory can be validated in another way. Focus has been used to indicate how discourses are cohesive. In the work of Grosz [1977], the focus space is a mechanism for indicating those items that are in the focus of attention in a portion of a dialogue. The immediate focus of a sentence influences the listener's interpretation of subsequent sentences. Example D2 shows that some aspect of the focus is changed or referred to in every sentence of the discourse. If some texts are cohesive without a focus, or have focus connections but lack cohesion, the use of focus as an indicator of cohesion and the corresponding use of focus for anaphor interpretation would become more doubtful.

The theory to be presented here is partially predictive since it predicts the co-specification of anaphora, and partially descriptive since the semantics underlying focus choice cannot be stated in a completely explanatory model. Focus is not independent of the speech act a sentence encodes. In fact, focus appears closely related to the speech act since it is part of it (see Sidner, [1978] for further discussion). In D2 the speech act is intuitively regarded as a request to perform a scheduling act. The item to be scheduled is a meeting, and hence the focus is one aspect of the larger speech act request. Further research on the theory of speech acts may indicate constraints on how focus can move. In chapter 2, where the method of determining the focus is given, semantic concepts such as thematic relations [Gruber, 1976] are used. The entire field of semantics is in a state of flux, and its changing foundation suggests that new evidence

may permit some of the descriptive work here to be accounted for in a more theoretical framework.

Now we are in a position to examine the work on anaphora and focus which has preceded this report. Because the approach of this report ties together several phenomena, research will be reviewed in artifical intelligence and linguistics, covering both sentence syntax/semantics and focus/topic concepts, and in discourse understanding and psychology. In the discussion which follows, I will use the general term "antecedence" rather than "co-specification" in discussing research by other authors as most use antecedence in their work. In reviewing other work, we must continue to be aware that there are difficulties in determining which phrase and interpretation co-specifies with a definite anaphor.

## 1.4 Artificial Intelligence Research on Anaphora

In artificial intelligence, Winograd [1971], Charniak [1972] and Rieger [1974] have each suggested an approach to the comprehension of anaphora in discourse. A closer review of their different approaches will provide some motivation for the approach of this report.

Winograd made two basic observations which formed the basis of a simple mechanism for interpreting pronominal anaphora such as *it* and *they*. First, he observed that pronouns had to agree with their antecedents in person, number and gender. Second, he observed the well-known heuristic that the antecedent of an anaphor is the last noun phrase that passes the person, number and gender test. To interpret pronominal anaphora in SHRDLU, he implemented this heuristic by means of separate lists (stacks) for each of the person, number and gender types of noun phrases in the discourse. When a pronoun was encountered, the most recently used noun phrase in the same person, number and gender stack was popped off the stack and used as the antecedent. This model of the comprehension of anaphora was very limited in its range of interpretation. Winograd himself suggested that inferences were needed to make a better model of comprehension since his model could not identify the correct antecedent of the anaphoric *they* in (1).

(1) The city councilmen refused the demonstrators a permit because they feared

violence.

Charniak observed from children's stories that many inferences are needed to interpret the antecedent of an anaphoric expression. The inferences are quite extensive as the following example indicates.

D6-1 Today was Jack's birthday. Penny and Janet went to the store. They were going to get presents. Janet decided to get a top. "Don't do that," said Penny. "Jack has a top. He will make you take it back."

In this case, there are two tops being considered, but *it* can refer only to the top that Janet wants to buy, since one does not return a top one has in favor of a new one. Charniak assumed that there were large collections of inference rules which he called demons. Demons knew what to do with a small piece of the total knowledge in memory and could invoke themselves whenever that knowledge was encountered. For example, a rule for D6 would be: look for someone not liking a present X and returning X. In the story above, Jack will not like the new top which Janet is going to buy as a present. When a sentence is encountered about taking something back, it can be concluded that the something is the present Janet wants to buy.

This approach gives rise to three difficulties. First, an immense number of demons is needed. Conceivably there may be demons for going to buy a gift and not having enough money, and for buying the wrong color, size, or model, for getting lost in the store, not finding the gift and so on. It is not clear whether there could be some generalized demon to incorporate all this information; that is, one cannot tell what knowledge should be in a demon and what should be left out. Secondly, as the approach stands, it may not be computationally feasible to have such a large cache of demons, but no other method seems to meet the needs of the demon strategy.

However, even if a cache of demons were available, a third difficulty blocks the way. If the story above were about model airplanes instead of tops, then duplication may or may not be a reason for not liking a gift since it is appropriate to have more than one model airplane. In fact, the D6 dialogue modified to discuss airplanes is somewhat confusing to people because duplication is not a reason for not wanting a present. On the demon approach, the demon for duplication would fire as well as the demon indicating that duplications are okay for presents which can be "used up," such as models. But it is unclear how control is to be handled when several

demons can fire. Yet without control, one cannot account for the manner in which people either understand or misunderstand the discussion of model airplanes. This last difficulty illustrates the main disadvantage of demons: they are difficult to control because there is an immense amount of knowledge, organized as demons, in which each demon can invoke itself independently of what the other demons do. There is no guarantee that any demon will invoke itself, nor is there a guarantee that several will not invoke themselves all at the same time. If several demons fire, there is no control over their functioning.

Rieger postulated that memory structures which were built from the discourse could be used together with other memory structures which were pre-stored in the memory to make inferences. During the inference process, a representation of a sentence as a memory structure was unified with a piece of knowledge in the database (also represented as a memory structure). The match of a variable to an entity in the knowledge structure established the antecedent of the phrase represented as a variable. To account for the reference of the definite noun phrase (hereafter, defnp) *the nurses* in D7, inferences are needed to deduce that John probably went to a hospital, was put to bed and was tended to by nurses and doctors.

D7-1 John was run over by a truck. When he woke up, the nurses were nice to him.

Rieger, like Charniak, was concerned with the inferencing needed to connect the sentences of a discourse into some coherent whole, and to disambiguate the references and antecedents of noun phrases. His algorithms suffer from the explosive number of inferences that can be made from each memory structure. Thus from the above example, one can deduce not only that John went to the hospital, but also that the truck was probably large, it probably had a driver, John probably was standing in the street and so on. A costly unification procedure must take place between each inference and the sentence in question in order to interpret the anaphoric and referential expressions and to connect the discourse. Furthermore, Rieger's system offers no explanation for how to decide when to terminate the inference process because it was not always clear when enough inferences had been made.

All three researchers argue convincingly that inferences must be made in order to choose antecedents for anaphoric expressions. Charniak and Rieger both suggest that the inferencing mechanism could also be used to understand how the discourses fit

together coherently. Their work indicates that one difficulty in interpreting antecedents is controlling the inferencing process. A second difficulty is reducing the search for likely antecedents. Winograd's stacks based on types contain every noun phrase mentioned explicitly in the discourse, but as the Rieger example above indicates, the referring expression may refer only implicitly rather than explicitly to something said previously. Furthermore, as indicated by Grosz [Deutsch, 1975], an antecedent may be several sentences and many noun phrases removed from its anaphor. Choosing the correct anaphor requires searching memory. These two problems interact: in Winograd's speculations and Rieger's approaches, the search for the correct antecedent requires that inferences must be made about each likely antecedent. The computations are exponential in the worst case. The first question for which this report seeks a new perspective is: what alternatives exist to these computational approaches?

The problem of determining the referential relationship for *the nurses* in D7, differs from finding the antecedents of pronoun anaphora as given in the Winograd and Charniak examples since in the latter case, the antecedent phrases are present in the text. Some anaphoric expressions do not have antecedents that are directly in the text. This observation suggests that the problems stated above are cannot be eliminated by finding a noun phrase that co-specifies with the noun phrase in question. Sometimes the problem is finding a noun phrase, which can be called the associator, that makes possible the inference that an anaphoric phrase co-specifies with some other concept related to the associator. As is evident from Charniak's and Rieger's work, a great many inferences may have to be made if one is not sure which noun phrase is playing the role of associator. The alternatives to the Rieger and Charniak approaches must include some explanation for associators. Now let us turn to some other research which has provided insights into the understanding of anaphora.

## 1.5 Approaches to anaphora in linguistics

Several other approaches to anaphora have preceded or appeared concurrently with the computational techniques of Winograd, Charniak and Rieger. In linguistics, anaphor comprehension has been restricted to cases of antecedents in the same sentence. Lasnik [1976] and Reinhart [1977] have suggested structures for sentences and rules

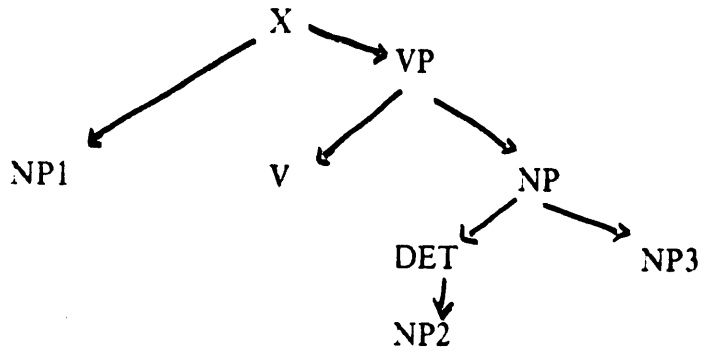
governing anaphora which use those structures.[1] Their rules do not generalize to antecedents in other sentences of the discourse because the structures they use are applicable only to single sentences. Lasnik, building on the work of Postal [1966], Langacker [1969], Wasow [1972] and others, has a rule of *precede and kommand* defining disjoint reference between noun phrases in a sen ence.

> If NP1 precedes and kommands NP2, and NP2 is not a pronoun, then NP1 and NP2 are disjoint in reference.
> Definition: A *kommands* B if the minimal cyclic node dominating A also dominates B.

Figure 4 shows NP1 and NP2 in a precede and kommand relation. NP1 precedes NP1 in the tree order, and X is a node dominating both NP1 and NP2. The minimal cyclicity of a node delineates sentence boundaries in constituent structure representations. This rule eliminates the co-reference in examples such as (2) and (3), since in both cases the first NP precedes and kommands the second, and the second is not a pronoun. However, as Reinhart [1978] points out, it does not eliminate (4).

---

**Figure 1.4. NP1 Precedes and Kommands NP2**




---

1. Lasnik and Reinhart speak of co-indexing as well as co-reference as the technical terms for antecedence relations. They are aware that co-reference is not an applicable term for the relation between some anaphora and their antecedent noun phrases. Co-indexing is the means of associating a pointer from an anaphor to its antecedent. It is similar in spirit to co-specify but does not make assumptions about semantic representations in databases.

2) * John$_i$ loves John$_i$'s mother.
3) * Mary$_i$ gave Mary$_i$'s friends a going away present.
4) * Near Dan$_i$, he$_i$ saw a snake.

Reinhart uses a rule of c-command. From the constituent structure of the sentence it defines classes of constituents in which disjoint reference must hold. These classes eliminate the need for a precedence relation which is the source of the difficulty with applying Lasnik's rule to (4). Her theory is an alternate to that of Wasow who regards (4) as acceptable.[1]

This report is not meant to provide an alternative account to the linguistic approach of Lasnik and Reinhart. Postal, Langacker, Wasow, Lasnik and Reinhart suggest that sentential anaphora are related to the structure of a sentence. The approach of this report comes together with the linguistic approach in cases such as (4) where the co-specification of the anaphor can be found outside the sentence in whatever discourse has preceded the sentence use. One may ask: why can't the approach to inter-sentential anaphora account for the sentential cases as well? As will be shown, inter-sentential co-specification relies on more than sentence structure; it must take advantage of relations among sentences of the discourse. Since these relations do not hold for sentential anaphora, the behaviors are different. Later in this chapter recent work on determining this structure will be reviewed.

Many computational systems (see Woods et al [1976], and Walker [1976]) have used semantic selectional restrictions on the noun phrases of a verb phrase in parsing and semantic interpretation. The approach is based on the semantic theory of Katz and Fodor [1963]. Katz and Fodor suggested that every word in the lexicon could be reduced to a set of primitive terms. Thus *bachelor* can be defined as an unmarried adult male (one sense of the word). Computational systems have been designed to incorporate this idea in various ways. Some actually reduce much of the language to primitives while others associate semantic categories (animate, inanimate for example)

---

1. Differences of opinion regarding the acceptablility of data are frequent. Sometimes the variation is due to dialectical differences, and other times it appears that more samples of native speakers are needed. This report will demand an account of examples like (4) since there are speakers who find (4) unacceptable.

with cases of verbs. These categories restrict which noun phrases can be selected with certain verb positions. Thus for *build*, the subject in an active voice sentence must be animate while the object must be inanimate. While selectional restrictions may exclude certain antecedents for a given anaphor, they are not sufficient for choosing an antecedent. In (5), selectional restrictions rule out *dog* as the antecedent for *it*, while in (6) either noun phrase in the first sentence can be the antecedent for *it* on the basis of semantic selectional restrictions.

(5) When his house burned down, Bill's dog died. Later Bill rebuilt it, but he never got another dog.
(6) Put the mud pack on your face. Notice how soft it feels.

An alternative account of anaphora is given by Kuno [1975] who attempts to address both sentential and inter-sentential anaphora with one set of rules. Kuno proposes several constraints on pronouns: command and precede (following Langacker and Postal), constraints on predictable theme (do not pronominalize the predictable theme), focus pronominalization (do not permit pronominalization in cases like "Among $John_i$, Mary, and Jane, $he_i$ is the tallest."), predictability requirement on backward pronominalization (do not pronominalize the lefthand noun phrase unless its referent is predictable from the preceding context), and the imitation tendency (if backward pronominalization is to apply, leave unpronominalized the noun phrase in the known part of the sentence). Of these, command and precede, focus pronominalization and backward pronominalization can be explained by the Lasnik-Reinhart analysis. For the others, Kuno requires a notion of theme and prediction, but these terms are not defined well enough to apply the rules with consistency.

Partee [1972] offers a different account of anaphora and co-references. She claims that pronouns treated as cases of co-reference fail to account for the following cases of antecedence:

(7) John lost a pen yesterday and Bill found one today.
(*one* is not co-referential with the pen John lost.)
(8) John claimed to have found the solution to the problem, but Bill was sure he had found it.
(*It* may refer to the solution or to what Bill found which Bill believes to be the solution.)
(9) John wants to catch a fish and eat it for supper.

(*It* is claimed to be equivalent to "the fish John catches" rather than "the fish John wants to catch.")

(10) No one would put the blame on himself.
(*Himself* does not co-refer with *no one.*)

All these sentences are problematic and depend upon a semantic representation rich enough to account for transparent and opaque readings. Cases like (7), generally called *one* anaphora, are considered further in Webber [1978], which is discussed below. Sentences like (10), called bound variable anaphora, show clearly the need for scope identification as part of single sentence anaphora interpretation. Webber provides one formalism for this problem, and Partee [1978] suggests a similar one. Since bound variable anaphora appear to be distinct phenomena, I will not consider them in this report.

Partee [1978] has distinguished a class of antecedents which she calls the pragmatic antecedents: those with no linguistic antecedent, as in (11), and those where the antecedent is in an earlier sentence, as in (12). Among these, I will consider those antecedents which are pointed to by definite anaphora. With this choice, I leave aside the *one* anaphora for later study. The bound variable anaphora and pragmatic anaphora are related since one would like to explain why the preferred reading of *him* in (13) is, say: Archibald, if there has been previous discussion of that person, but, for some speakers, the preferred reading is each of the men if no one else has been discussed.

(11) Why is he [pointing] here?
(12) I couldn't reach Elliot. He is probably in Boston.
(13) Every man put a screen in front of him.

Karttunen [1968] defines a class of discourse referents to be: entities which have been asserted to exist in discourses; entities from the surrounding environment that have been observed to exist; entities that are determined by a common understanding which the participants believe they share with regard to their environment; and entities which can be inferred with some degree of certainty from the existence of another referent by an implication with which the listener is familiar. He suggests two rules for testing for semantic anomaly of defnps and pronouns. One rule states that either a noun phrase or a pronoun is anomalous if it does not refer to an existing discourse referent. The second rule states that for a noun phrase to refer to an existing discourse referent, it must be a defnp or a pronoun. With such a class of

discourse referents, one can ask: to which discourse referent does a defnp or pronoun refer? This report builds on Karttunen's work by partially answering the question of which discourse referent the defnp or pronoun co-specifies with.

Webber [1978] represents the possible antecedents of anaphora in discourse using a typed logic. She includes an extensive list of the range of discourse anaphora. A discourse has a collection of entities, categorized into individuals, sets, events, actions, states, facts, beliefs, hypotheses, properties, generic classes, typical set members, stuff, and specific quantities of stuff. When the speaker wants to refer to one of these entities in a discourse model, s/he can do so by using a definite pronoun whose referent is defined as an entity in the speaker's discourse model. One can also refer to an entity in a model by using a description of it. The speaker assumes that the hearer has a "counterpart" of the entity or can evoke one. Webber states that a definite anaphor A has E, a discourse entity, as its referent. A's antecedent is E's "ID": the unique description of E conveyed by the immediately preceding text.

Webber has provided two definitions for anaphor interpretation. The first, a definition of reference, picks out an entity already in a discourse model.[1] The second definition uses the notion of antecedence to create a formal representation for a text in which E's ID occurs as the antecedent for A. This representation can be used in disambiguation of discourse anaphora. The definition of antecedence in terms of a discourse ID is similar to the co-specification with relation because in the case of noun phrases which are discourse constructs the two terms define the same relation; the terms differ in the case of a noun phrase used to refer. Webber argues forcefully that surface structure phenomena and scope of quantification must be represented as part of anaphor resolution. Her representation provides a means of capturing the ambiguity in

---

1. This definition is misleading. When a speaker says "John went to the store. He is going to buy some milk," *he* does not refer to an entity in anybody's discourse representation; *he*, if it refers at all, refers to the person called John who presumably exists in the real world or in some other world such as a story. *He* also co-specifies with the noun phrase "John," according to the relation shown in figure 1. However, the confusion created by this definition is not too serious in Webber's work because she relies mainly on the notion of antecedence.

the relative clause of (14), which allows both *it* and *them*.[1]

(14) Three men who lifted a piano dropped {it, them}.

(Ex: lambda(v: set(man))[(Ey: piano) L v,y]) . D x,IT & /x/=3.[2]
Three men who (together) lifted a piano dropped it.

(Ex: set(lambda (v: Man)[(Ey: Piano) . L v,y])) . D x,THEM & /x/=3.[3]
Three men who (each) lifted a piano dropped them.

Webber's thesis provides a starting point for determining the co-specification of pragmatic anaphora: her formalism captures much of the surface syntactic and scope relations which must be used to decide among the possible antecedents of the anaphoric terms. Webber makes no claims about which anaphora will be chosen beyond the fact that the chosen antecedent must conform to the relations stated in the formalism. The question remains: what governs the choice of the antecedent of pragmatic anaphora?

In the literature previously discussed, there are several kinds of anaphora:

1. Bound anaphora (Webber, Partee)
2. One anaphora (Webber)
3. Sentential anaphora (Langacker, Postal, Lasnik, Reinhart)
4. Pragmatic anaphora and Discourse referents
   a. anaphora present in the discourse (Charniak, Rieger, Winograd, Webber)
   b. anaphora referred to by pointing or other deictic means (Partee)
   c. anaphora that are inferentially "present" in the discourse (Karttunen, Rieger)

Several approaches to anaphora interpretation have been presented:

---

1. Some readers will find the second reading of (14) unacceptable without the presence of *each* following *who*. The Webber notation does not offer an indication of what the scope behavior is which causes some readings to be preferred for some speakers.
2. The notation is read: There is an x such that it is a set of men who lifted a piano and that set of men dropped it and the number of men is 3.
3. This is read as: There is an x such that it is a set, whose members each are men and have lifted a piano and the set of men have dropped them and the set contains 3 members.

1. Gender, number and person agreeement (everyone)
2. Syntactic structure constraints (c-command (Reinhart), kommand and precede (Lasnik))
3. Sentential variable scoping (Partee, Webber)
4. Discourse ID's (Webber)
5. Semantic selectional restrictions (Katz and Fodor, Walker, Woods)
6. Demons (Charniak)
7. Inferencing via underlying conceptualizations (Rieger)

The problem which still remains is: How can pragmatic anaphora be distinguished from sententially bound or from quantifier bound ones? How can a noun phrase which acts as the co-specification of the pragmatic anaphor be chosen? Can this problem be solved in a manner which allows for needed inferences and yet controls the inference process? Can the chosen method distinguish specification, that is, first uses of referring expressions, from co-specification?

## 1.6 Focus, Topic and Comment

Other work, not on anaphora, but on the notions of focus, topic and comment, provides a background for this report. Several uses of the notion of focus can be found. Quirk and Greenbaum [1973] claim that the focus of information in a sentence is signalled by intonation. The focus is said to occur in neutral position (a default stress position) when the chief prominence of the intonation is on the last open class item[1] in the clause. A contrastive focus may be placed at any other point, either on a phrase, a single word or a syllable. Supposedly focus indicates where new information in a sentence lies. In cleft sentences, such as (15), the intonation focus is the object of the *be* predicate and in pseudoclefts, such as (16) a similar phenomenon occurs. For existential *there* sentences, they observe that by using *there* as a dummy subject, a sentence can be transformed to place new information later in the sentences as in (17) below.

(15) It was his best suit that John wore to the dance last night.
(16) What John wore to the dance last night was his best suit.

---

1. Open class items are categories of non-function words. Thus nouns and verbs are open class while articles, pronouns and conjunctions are closed.

(17) Plenty of people are getting promotions. ==>
There are plenty of people getting promotions.

Clefts, pseudoclefts and *there* insertion clauses are syntactic indications of focus which will be considered in this report; prosodic clues will not be. While prosodics play an important overall role in pragmatic anaphora, this aspect will not be considered in this report since the tools for interpretation of prosodics are still limited. These phenomena must play a significant part in a full explanation of definite anaphora because D8-2 below is acceptable only with contrastive stress on the *he*.

D8-1 Robby made breakfast for the king.
2 Robby thought it tasted good, but he didn't think so.

Closely related to the concept of focus are the notions of topic, comment, and theme. The concept of *topic* as used by Lyons [1968] and Sgall [1973] is defined as what is talked about in the sentence, while the *comment* is what is said about the topic. Halliday [1967], speaking from the same point of view, uses the terms *theme* and *rheme*. Both Lyons and Halliday claim that topics are the first constituent to appear in the sentence. Lyons uses the example of "John ran away," and says that John is what the sentence is about while what John did was run away. He indicates that subjecthood is a default position. Halliday views the matter differently: theme-rheme is but one of three dimensions of sentence organization. Theme and rheme reflect the message organization so that the first constituent is always the theme. Thus the theme is different in:

(18) John's aunt left him this duck press. (unmarked theme)
(19) This duck press John's aunt left him. (marked theme)

The notion of *focus*, for Halliday, is part of the second dimension, the information dimension, of sentence organization. An information focus is realized as phonological prominence, specifically the assignment of tonic value to a given syllable. Hence, focus is closely related to the concept of primary stress. Information focus reflects the speaker's decision about where the main burden of the message lies. It is one kind of emphasis whereby the speaker marks out a part of a message block that s/he wishes to be interpreted as informative. Information focus provides for a given and new labelling of elements where *given* and *new* mean derivable or not derivable from the preceding discourse. Thus information focus is a discourse feature and not a sentence feature like theme. Halliday claims that the difference between *given* as a

focus notion and *theme* "can perhaps be best summarized by the observation that, while *given* means *what you were talking about, theme* means *what I am talking about.*"[1]

Sgall, Hajicova and Benesova [1973] speak of topic, comment and focus. Their use of the terms is similar to Halliday's use of theme, rheme and focus respectively. The difference is that topic-comment articulation (TCA) is a phenomenon which results from what they see as the *communicative dynamism* of language communication. They point to Firbas [1971] who states:

> Communicative dynamism is based on the fact that linguistic communication is not a static but a dynamic phenomenon. By communicative dynamism, I understand a property of communication, displayed in the course of the development of communication to be conveyed, and consisting in advancing this development. (page 135-136)

They describe three layers of TCA, the first for sentences that do not link to the preceding communication or situation, the second for sentences where a constituent does refer to some previous communication and the third for sentences where a miscommunication is repaired. For Sgall *et. al*, the focus is that part of the sentence in which no information is given which is bound, i.e. is related, to the previous context. Thus focus is somewhat different from Halliday's use of the term because the focus for Sgall is related to the previous context.

Grimes and Kuno each speak of theme. Kuno says that theme is "what the rest of the sentence is about."[2] In this he is in accord with Halliday especially because he believes, like Halliday, that the theme is not necessarily old, predictable information. Grimes says that the theme is the semantic choice for the point of departure for a sentence.[3] Grimes' idea of point of departure is what the speaker wants to say from a particular viewpoint, whereas theme is a signalling device for the cohesive structure in a discourse. There are marked and unmarked versions of theme depending on the mode of a sentence: in a declarative the subject is the unmarked theme while in an imperative the verb is the unmarked theme. In this view, Grimes sees theme as a

---

1. Halliday [1967], *op. cit.*, page 212.
2. Kuno [1975], *op. cit.*, page 277.
3. Grimes [1975], *op. cit.*, page 324.

semantic notion which gets realized in the syntax by its initial placement.

Chomsky [1971] speaks briefly of *focus* and *presupposition*. Focus is the word, under normal intonation, with main stress which serves as the point of maximal inflection of the pitch contour. Thus in "Is it John who writes poetry?" John is the focus, while the presupposition is that someone writes poetry. Chomsky's use of focus corresponds to focus of contrast as used by Quirk and Greenbaum, discussed above.

Akmajian [1973], building on Chomsky's use of focus, indicates that focus can be used to explain the *it* in (20a) and (20b).

(20a) Pratt roasted a pig in the fireplace last year, but none of his friends realized it.
(20b) Pratt roasted a pig in the fireplace last year and Whitney did it too.

In (20b) the focus on Pratt leaves the verb phrase as presupposed, so *did it* is anaphoric to the verb phrase. Here it is not the focus which is the source of the anaphor, but the presupposition. Similarly in (20a), the focus is empty, so there is no presupposition, and the whole clause is presupposed by *it*.

Recently Chafe [1976] has attempted to clarify the muddy waters of focus, topic and comment. Chafe speaks of nouns as having different statuses: given and new, focus of contrast, definite or indefinite, subject, topic and point of view. He tells us that given information is distinguished by the pronunciation of low pitch and stress, and by its being subject to pronominalization. Givenness is established by the speaker assuming something in the addressee's consciousness on the basis of either linguistic or extra-linguistic context. In contrast to given, the status of definite means that for a set of objects the speaker knows that the hearer already knows of them and can identify the particular member the speaker has in mind. Chafe uses the term *identifiable* as synonomous with definite.

Focus of contrast is recognized by the highest pitch and stress on a stressed syllable and distinguishes a noun from the shared background knowledge. The subject for Chafe is what we are talking about. He cites work by psychologists which shows that the subject is an effective prompt for understanding a sentence. He takes issue with the topic-comment idea: comment need not be new information since new

information can be in any position in the sentence. Chafe admits that in clefts and pseudoclefts, however, what the sentence is about is not in subject position. Chafe claims that the topic, as used by Lyons, Halliday and others, is simply a focus of contrast. Chafe points to Chinese, which he calls a "topic-prominent" language to show that topic is not so much what the sentence is about as "the frame within which the sentence holds."

While the observations in the work of Halliday, Grimes, Kuno, Lyons, Chafe, Sgall and Chomsky differ, each is seeking to explain what it is that speakers talk about. Most of the accounts are descriptive summaries; Chomsky and Akmajian give an indication of how the notion of focus might be used, although Akmajian's study is limited to those cases where a maximal inflection in the pitch contour occurs. All hint that the idea of given and new in discourse suggest that focus or topic might be helpful in resolving anaphora. Yet it remains unclear how the speaker and hearer use focus or topic to compute the antecedents of anaphora. This computation is the central concern of this report.

A concept related to focus is discussed by Chafe [1975]: the concept of foregrounding. Any of the noun phrases and the verb in the introductory sentence are foregrounded and can be discussed in a subsequent sentence. Foregrounding is indicated by low pitch and amplitude. Chafe indicates that foregrounded objects may be pronominalized (whereas non-foregrounded ones cannot). He also claims that two sentence boundaries act as a constraint on when a particular foregrounded object eliminates other objects as foregrounded, though he also suggests that constraint may last longer than two sentences. Chafe goes on to say that "we do not seem to be able to say precisely when a speaker must henceforth treat a lexical unit as no longer foregrounded." In this report a synthesis of focus treatments will be offered which accounts for the foregrounding phenomenon in anaphora comprehension.

## 1.7 Discourse Approaches to Anaphora

An alternative approach to linguistic accounts of anaphora disambiguation appears in research efforts to interpret stories using story "grammars" or story schema. This research aims at consideration of larger units of language than sentences. An initial specification of "grammars" for various texts can be found in Rumelhart [1975],

while story schema research is exemplified by Schank *et al* [1975] and Collins, Brown and Larkin [1977]. Rumelhart specifies a grammar, assumed to be implemented top-down, for stories; the grammar contains primitives such as "an action is an activity engaged in by an animate being or a natural force." The primitives are not completely specifiable since some seem to indicate deducing plans of actions in order for the grammar to be instantiated.

Collins *et. al.* suggest a model which is partially instantiated as the story progresses. The model is revised by the rebinding of variables via constraint satisfaction, resolution of conflicts (when inconsistencies arise in the model) and shifting from a question, which the model suggests but cannot answer, to a related question. Schank *et. al.*, in their story understanding program SAM, also assume a stereotypic model, called a script, of a given situation. The script is a set of pre-determined sequences of actions. Understanding a mundane situation is interpreted as matching the script against sentences of the story. For example a restaurant script contains information about going to a restaurant: if one enters a restaurant, one is likely to find a table, sit down, take a menu and order and so on. When a story talks about such matters, the script can be used to understand the story.

All of these researchers seem only peripherally concerned with anaphora. Rumelhart seems to take anaphora resolution for granted in his system and says little about it, while Schank seems to assume that anaphora can be resolved by variable binding within the script. The Schank model certainly can resolve some kinds of anaphora (it is not clear how it succeeds with *one* or *this-that* anaphora), but it does not explain the nature of the anaphoric process beyond the stereotypic situation. Collins *et. al.* present a model in which binding of variables is essential to the story interpretation. What they leave unexplained are the rules governing the choice of variable bindings and re-bindings for story entities. Though the model offers a new explanation of how people change their minds about a story as they interpret it, the details of anaphora interpretation have yet to be elaborated.

Williams [1977] approaches discourse grammars from the viewpoint of Chomskian linguistics. Williams assumes the principle that discourse grammar rules apply after sentence grammar rules. He argues that the discourse rules apply to the logical form of sentences following semantic interpretation of surface structure (see

Chomsky, [1976]). Williams discusses only the VP rules which assign a verb phrase in one sentence an anaphoric relation with an empty verb phrase in another sentence. He argues for the rule as one of interpretation rather than deletion. That is, in the base component when the verb is generated, a delta, a notation for an empty string, is generated in place of the verb; on the deletion account, all constituents are generated and then selectively deleted. Although brief, Williams' approach suggests that interpretation of pragmatic anaphora can be related to current linguisitic theory.

Halliday and Hasan [1976] approach text differently from the manner of Halliday's other work. They speak of the concept of *cohesion* in a text, where a text is any passage, spoken or written, of any length that forms a united whole. The *texture* of a text is what distinguishes text from non-text. Texture is provided by the cohesive relation that exists between elements of the text. The cohesive relations, called cohesive ties, are reference, substitution, ellipsis, conjunction and lexical cohesion. Cohesion is not a structural relation in the sense that structure links the parts of a sentence or a clause. Nor is it a discourse structure relation which links sentences as a paragraph. Cohesion, Halliday and Hasan stress, is a semantic relationship; "it refers to the relations of meaning that exist within the text and that define it as a text."[1]

In addition to cohesion, texture has two other components. One is the textural structure within a sentence. By this, Halliday and Hasan are referring to Halliday's work on theme systems, discussed earlier. The other is the discourse structure, that is, the structure which results in narratives, prayers, formal correspondence, novels and so on.

Halliday and Hasan say of reference that, other things being equal, the most probable target of a cohesive reference item seems to be the theme of the preceding sentence. However, they give only one example. This report will provide a range of data that suggests that something besides Halliday's notion of theme is the target. Halliday and Hasan do provide insights into the ways in which reference, ellipsis, substitution, conjunction and lexical cohesion form a cohesive relation, but they leave unexplained the rules that govern the use of any of these classes of cohesion.

---

1. Halliday and Hasan, *op. cit.*, page 4.

## 1.8 Some Recent Psychological Work on Anaphora

Both the discourse understanding literature and the focus literature suggest that knowing what the speaker is talking about may be one means for interpreting anaphoric expressions. Current research in psychology on antecedence and paragraph comprehension suggests a similiar direction.

In their work on paragraph comprehension, Carpenter and Just [1977, 1978] define the notion of a discourse pointer: a symbol in the comprehender's mind that indicates the current topic of the discourse or of the perceptual context. They use this notion to design a sequence of experiments to discover what role the discourse pointer plays in paragraph comprehension. In addition they suppose that the discourse pointer is integrative in the sense that it represents the old information in a sentence, so that when sentences mark the discourse pointer as new information instead of old, comprehension is slowed down.

Carpenter and Just devised several tests, a sample of which will be given to illustrate their methodology and theoretical conclusions. One test relates sentence comprehension to comprehension of a picture shown previously. The sentence uses clefts or pseudoclefts, which clearly mark new and old information. Their hypothesis is that sentences should be easy to comprehend if the linguistic structure marks as old the constitutent identified by the discourse pointer; the sentences should be difficult if the linguistic structure marks the discourse pointer as new. The results of the experiment can be interpreted as confirming this prediction when response time is used as a measure of difficulty.

Another test, done with sentence pairs, consists of a simple subject-verb-object (S-V-O) sentence followed by a cleft or pseudocleft sentence which provides new information in an acceptable constituent or an unacceptable one. Thus (21) might be followed by either (21a) or (21b) below:

(21) The ballerina captivated a musician in the orchestra.
(21a) The one who captivated the trombonist was the ballerina. (inappropriate)
(21b) The one who the ballerina captivated was the trombonist. (appropriate)

The results show that less time was needed to integrate pairs when the second sentence

and its information were appropriate to the context of the opening sentence. A similar test used intervening sentences. Integration by subjects required more time when the intervening sentences contained information unrelated to the discourse pointer.

Carpenter and Just's notion of a discourse pointer is similar to Chafe's notion of foregrounding and Halliday's concept of focus. However, they do not specify how a discourse pointer is chosen. Their tests are suggestive because the sentence pairs used contain cleft and pseudocleft patterns which seem to indicate the speakers use focussing in discourse. Their tests are only suggestive because the tests concentrate on the time taken to interpret the whole sentence rather than just the referring expressions.

Clark and Haviland in an article "Comprehension and the Given-New Contract" [1977] specify a maxim of antecedence: try to make an utterance so there is one and only one direct antecedent for any previously given information, and that antecedent is the intended antecedent. They do not define direct antecedent, but from their examples, it appears that a direct antecedent is a noun phrase in a previous sentence with the same head noun as the given noun phrase. In contrast, an indirect antecedent is a noun phrase which is related to the anaphoric noun phrase in some other way, as shown in D9.

D9-1 Horace got the picnic supplies out of the car. The beer was warm.

Clark and Haviland have experimented with comprehension of direct and indirect antecedents and found that comprehension is markedly slower for indirect antecedents. They also point out several breaches in the maxim of antecedence. Among these are the lack of any antecedent, covert breaches as in "H[illegible] you stopped beating your wife?", the lack of a bridge (a noun phrase which connects two sentences) and explicit violations[1] as in (22) below.

(22) Bill slipped me a bottle of gin, but the idiot told my wife about it.

Clark and Haviland also point out that the degree to which bridging is unacceptable seems to vary. (23) is easy to interpret while (24) is difficult:

---

1. Chapter 3 will offer an alternative explanation of antecedence which accounts for why (22) is acceptable to native speakers and is not a violation.

(23) The office was cool because the windows were closed.
(24) The haystack was important because the cloth ripped.

Clark and Haviland observe that both are easy if the right keyword is used as a prompt (air conditioning and parachuting).

The Clark and Haviland experiments show that the maxim of antecedence can be useful if some means of specifying the intended antecedent can be found. The problem is complex since this specification must account for why indirect antecedence takes longer, and why lack of a bridge causes unacceptable sentences.

Springston [1975] has studied the processing of sentence anaphora. He assumes three linguistic constraints are operative in finding the antecedent of a pronoun: the clause-mate constraint (reflexives are always used when an antecedent relationship is indicated between noun phrases in the same simple sentence), converse clause-mate constraint (simple pronouns cannot take as an antecedent a noun phrase that co-occurs in the clause) and the gender agreement constraint (pronouns must agree in gender with their antecedents). He makes an assumption about the nature of anaphor interpretation which does not hold in general, as Grosz [Deutsch, 1975] shows, but which provides some insights for single sentence interpretation: "examination of candidate anaphora[1] proceeds by clauses, with the candidates within the same clause as the pronoun being considered first."[2] If true, simple pronoun interpretation is an exhaustive search process. Since Grosz' examples show there is sufficient doubt that this is true inter-sententially, no conclusions regarding non-single sentence antecedents can be made. However, Springston's experiments permit some conclusions to be drawn. One experiment tests the backward and down constraint: a simple pronoun serving as subject of a clause cannot grammatically accept as an antecedent a noun phrase that is in a clause that is both to its right and subordinate to it. In (25), the backwards and down constraint (as well as gender constraint) eliminates *Joan* as the antecedent of *he*.

(25) Bill claims that when he was living in London, Joan was a popular movie star.

---

1. Springston misuses the term *anaphor* throughout his work--he means antecedent.
2. Springston, *op. cit*, page 68.

Springston hypothesized that *Joan* would be processed and rejected faster since both gender and the backwards and down constraint apply. However, he found no significant difference between sentences such as (25) and the same sentence with its subordinate clauses reversed. Similarly, in sentences such as (26), he found no evidence for a constraint that would indicate that the presence of *own* would cause a faster prediction of the antecedent of the pronoun.

(26) Sam reported that Bill struck {his, his own} father.

These results indicate that a theory of anaphora interpretation must do more than solve the problem of pragmatic anaphora. Haviland and Clark and Springston all indicate that pragmatic anaphora require different methods of interpretation depending on their place in the sentential or discourse structure. The theory to be presented here will take into account these differences.

## 1.9 Focus in Discourse

A major contribution to the study of noun phrase disambiguation and discourse interpretation is the research of Grosz [1977, 1978]. Grosz uses the concept of focussing (explained below) as her major tool in these two areas of study. Grosz develops the concept of a focus space which is a subset of the whole knowledge space which is most relevant. Focus spaces are implemented as a partitioned semantic network representation of objects currently highlighted by the discourse and are computed dynamically during the interpretation of a dialogue. Partitioning eases computation; more importantly, partitioning highlights a part of the discourse by sectioning off a piece of the knowledge network. Items that are highlighted are said to be "in focus," that is, the item is represented in a focus space. Grosz observes that several focus spaces can be "open" at one time due to the nature of a discourse, but that only one of the spaces is "active," or currently selected. Focus spaces are open because the discourse has not indicated that discussion of the objects in them has ended.

In addition to the focus space, there is a plan for the general task which one of the speakers is trying to perform and which the speaker is trying to talk about. The task plan, or task model, is a data structure, outlining the subtasks needed to perform each task. This plan is used as part of the focussing mechanism. Those focus spaces

not currently selected may be re-selected for two reasons: the current focus space may represent a sub-task which has ended so that a jump back to an open space is forced; or the current task in the discourse may contain a referent back to an object in an open space and that space will be re-selected.

Grosz is able to account for the references in non-pronominal definite noun phrases using focus spaces. Such references occur within open focus spaces, and refer to objects in the focus spaces. Grosz also distinguishes between a global and an immediate focus. She says that "global focus refers to the influence of memory for the more general meaning conveyed by the preceding utterances of a discourse on the current sentence, while immediate focus refers to the interpretation of a listener's memory for the linguisitic form of an utterance."[1] The immediate focus of a sentence is used to expand an elliptical phrase in a subsequent sentence.

Using focus spaces to explain references, Grosz points out that some objects are implicitly in focus. For example, parts of an object are implicitly in focus when an object is in focus. A major part of her work is devoted to developing a noun phrase resolution procedure which can match noun phrases represented as semantic network fragments containing variables against a semantic network database. The focus space representation is used to constrain the search for a matching piece of the network. Noun phrase resolution done in this way is used to resolve non-pronominal defnps and to answer questions which arise in discourse.

Grosz' work leaves a number of open problems. Among them are how a focus can be used for pronoun disambiguation, specification of rules for what makes an immediate focus, and specification of a full set of rules for disambiguation of defnps. Certain cases of *the* article defnps are not covered, including attribution, references definable based on computations specified by modifiers (as in "the next boy in line"), and a more complete treatment of the highly complex phenomenon of generic/non-generic distinctions.

Grosz' work also considered focus movement which is determined by the

---

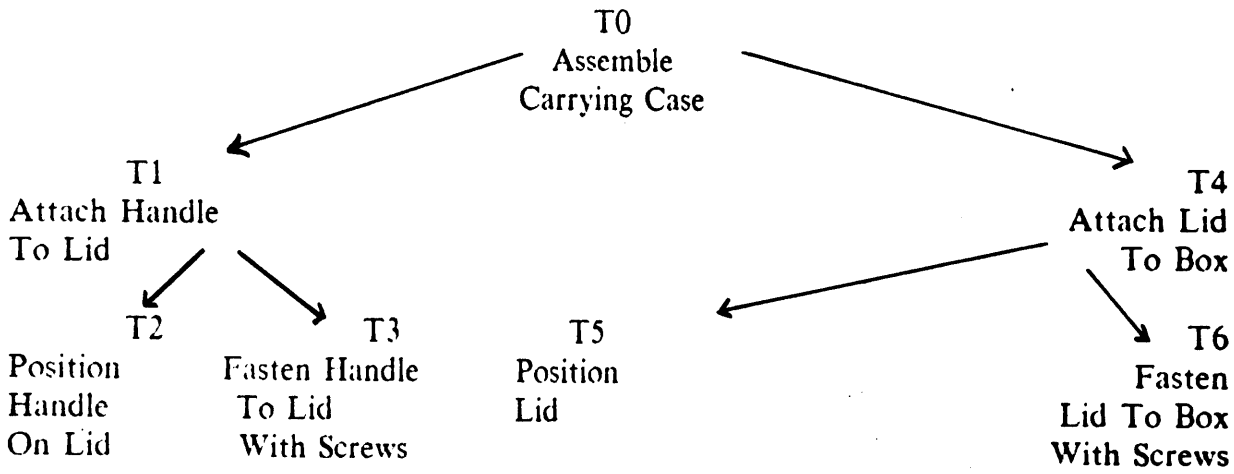
1. Grosz [1977], *op. cit*, page 5.

structure of the task. Unlike a script, the task model does not limit the exact form of a dialogue. It instead indicates the parts of the task which must be completed and the possible parts of the dialogue, ordering them hierarchically. In fact, the dialogue may contain some discussion which is not an aspect of the task, e.g., asking questions about the location of a tool needed for performing a task. This model can be used to decide when the discourse has moved from one focus to another. An example task will show how.

Figure 5 represents the task of putting a box together. If the discourse indicated that the current task point were T2, then reference made to *the screws* would indicate a move to T3. Grosz points out that not only the referential phrase is needed to indicate the move, but the remainder of the utterance is needed as well. Full utterances indicate the closing of a sub-task as in "The handle is fastened down with the screws," but Grosz does not discuss how to analyze them for their role in focus movement. Once T3 is closed and T4 and T5 begun, a reference to the screws indicates a move to T6. This example shows the use of a task plan to determine when focus movements have occurred. Grosz also points to the use of a phrase like *the screws* for moving to hypothetical worlds as in "Have you ever gotten this far and then realized that the screws don't fit?" Grosz leaves for later research how hypothetical world references could be recognized.

---

**Figure 1.5. Task Representation for Box Assembly**



Grosz' focus movement mechanism depends on the presence of the task model around which the dialogue is centered. Given such a model, many kinds of referential terms are completely resolvable and the movements in focus for the task are well defined. Task models and the use of focussing within dialogues about the task indicate those discourse situtations where focus moves and reference can be determined easily. They represent a clearly demonstrable part of the whole process of focus in reference disambiguation in general discourse. As might be expected, discourses with less pre-established frameworks have different properties.

All discourses seem to have frameworks which are known to the speakers of the discourse. Without such a framework, the speaker would have difficulty determining that s/he had said what s/he wanted to. Hearers may have difficulty interpreting references because the speaker's framework is not always available to the hearer. Rather than having a pre-established framework of a task model, the hearer in most situations has only his/her general knowledge of what is related to the discourse element which is focussed on. Sometimes in simple discourse, the hierarchical network of knowledge associated with the element in focus is sufficient to distinguish the specifications of the defnps. However, as I shall show in detail in chapter 3, the hearer uses other clues. The speaker often relies on a maxim of perspicuity to connect a defnp which does not seem on the surface to be related to the element in focus. The hearer also makes suppositions which are not necessarily true in general, but which explain the otherwise missing connection between a defnp and the element in focus. Hearers also rely on focussing to indicate that a defnp refers to some entity not mentioned in the discourse.

These techniques for reference disambiguation constitute a kind of guesswork by the hearer which relies on the structure of general knowledge, the element in focus and the nature of the communication process. As I will show, some of the focus related information, such as clefting and pseudoclefting, is syntactically marked. However, most of the phenomena of reference disambiguation can be detected only from consideration of general knowledge. In discourse with fewer well defined purposes than task dialogues, matters are less clear. In those cases, because the knowledge structure of a human speaker-hearer (or a computer speaker-hearer of some competance) is a bigger and richer structure than a plan for a task, the comprehension process for the hearer

can make fewer predictions concerning which parts of a discourse are actually completed. If the speaker does not indicate what is no longer under discussion, the hearer has more possibility for error or misinterpretation of a referential phrase. Furthermore, if the hearer does not have as much knowledge of the element in focus as the speaker, certain implicit focus connections (discussed in chapter 3) will not be obtainable, and hence the hearer will not be able to disambiguate certain references. These observations imply that the particular knowledge which actually appears in the knowledge network, as well as the form of the network, must be taken into account in reference disambiguation. As I will show in chapters 3 and 4, some texts are more easily interpreted when the speaker indicates explicitly the connection between the referential term and the element in focus, rather than leaving the connection to the hearer's general knowledge.

## 1.10 Guide to Remaining Chapters

This chapter is a brief introduction to the concept of focus and how it is used to interpret co-specifying expressions in discourse. The remaining portions of the report are needed to define focus and establish rules of the interpretation of anaphora. In the chapters which follow, focus will be examined for its role in several cases of co-specification. Chapter two will present a detailed account of how a focus is found, and how focus can move in a discourse. Chapter three will discuss focus and definite noun phrases, with rules for the interpretation of the anaphoric expressions. This chapter will build on the work of Grosz and on the observations of Karttunen on discourse referents. Chapter four will present rules for interpretation of personal pronouns. This chapter will build on the linguistic studies discussed here and pose an alternative to the computational approaches suggested so far by researchers in artificial intelligence. Chapter five will consider *this-that* anaphora and the other parallel foci such as *the one...the other*. An account of the interpretation of such anaphora uses the concept of focus to delineate co-specification. Chapter six will discuss two computer implementations of rules in chapters two, three and four and their relation to the larger problem of the interpretation of speech acts in a discourse.

The import of this report is the development of a definition of focus and the rules for the interpretation of definite anaphora. Focus has been considered and

focussing their attention on some element of the discourse. The element which is focussed on is elaborated by the clauses of the discourse. Often speakers' discourses can be quite different; their discourses are incoherent or at least hard to follow because:

> 1a) speakers talk about several elements without relating them or
> 1b) speakers talk about several elements without informing the hearer that several elements will be discussed at once or
> 2) there is no central element.

In a nutshell, discourses with properties (1) or (2) are not connected, that is, they lack a focus of attention, an element which is focussed on. The focus is then one of the connecting threads that makes a text or a set of utterances a discourse.

Focussing is a discourse phenomenon rather than one of single sentences. In general a single sentence is insufficient to capture all the information that a speaker wishes to tell a hearer. One might expect that when a speaker uses several sentences, the element of discussion would need to be re-introduced in each sentence. However, re-introduction is a highly redundant process and thus inefficient. Furthermore, in fact, when spreading the information over several sentences, the human speaker does not tightly relate all the sentences s/he speaks about a particular element. If re-introduction is not used, and still hearers claim to know what element is being talked about, there must be some means by which the focus connectedness occurs. In fact, there are two ways. The first uses special words which indicate to the hearer "that I am still talking about the thing I talked about in the previous sentence." Traditionally these signals are called anaphoric expressions. The second relies on assumed shared knowledge; the speaker assumes that some connections between the focus and some other elements are so common that s/he need not explicitly state what they are. Of course, there is a risk that the connections are no longer obvious, resulting in a set of sentences that are not connected.

Now a possible line of investigation becomes clear. If we can discover how the assumed shared knowledge is used, and if we can discover what the rules are for using words to signal the focus, we can predict whether a set of sentences will either be a discourse or incoherent or somewhere in between. Since researchers in artificial intelligence and psychology has been studying knowledge structures for some time, there are some results to build on and test. The approach taken here is to understand how

language use is constrained by the kinds of associations commonly available to both speaker and hearer.

The other important goal of this investigation is the formulation of rules governing the use of anaphoric expressions. A significant assumption is being made in pursuing the investigation: there are rules which govern how the speaker signals that s/he is talking about the same thing, and these rules which can be computed. It is possible that there are no computable rules, that some "oracle" determines what anaphoric expressions take as their antecedents. However, that doubt will be set aside for the remainder of the report because in the end, it will be clear that such a doubt is unfounded.

There are then two goals of investigation to be carried out: using focussing as a means for formulating rules for anaphoric expressions, and using the element in focus and the structure of assumed shared knowledge as a means for determining the constraints on language use. These goals are mutually dependent. Although one might guess at its truth, this claim cannot be argued a priori. In the investigation of how anaphoric expressions are actually used in discourse, I have observed this relationship. Because of that observation, these two goals will be studied together in the chapters which follow.

## 2.2 Focussing for Definite Anaphora Comprehension

Focussing can be described as a computational machine, part of which determines the antecedent of an anaphoric expression. This mechanism uses the element in focus to decide how certain phrases point back in the discourse and what they can point back to. The comprehending part of the machine makes decisions about antecedents based on a set of rules which depend on the focus. The rules which are embodied in this mechanism also permit the mechanism to detect some unacceptable uses of anaphora. Therefore, to be theoretically useful, the focus mechanism has two tasks to perform; it must simulate not only the hearer's behavior in understanding certain anaphors, but also the hearer's behavior in failing to understand others.

Focussing provides an important feature missing in earlier computational theories of anaphora: it controls the process of inference for anaphoric expressions.

Inferencing is controlled because the focussing mechanism restricts inferencing to providing evidence for what the focus mechanism picks out as the proper antecedent. Rather than allowing inferencing running wild every time some piece of information is encountered, inferencing is used to support what the focus mechanism predicts as the antecedence relation. Inferencing is still powerful because it can say "no, I can't support that idea--in fact I get a contradiction among other things I know given that antecedence relation."

Let us consider an example of how the focussing mechanism might work in the comprehension of definite anaphora.

D1-1 Last week there were some nice strawberries in the refrigerator.
2 They came from our food co-op and were unusually fresh.
3 I went to use them for dinner, but someone had eaten them all.
4 Later I discovered it was Mark who had eaten them.
5 Mark has a hollow leg, and it's impossible to keep food around when his stomach needs filling.

Suppose the focus for D1-1 through 4 is the strawberries mentioned in the first sentence. *They* is used to co-specify with the strawberries in each sentence. The focus mechanism would need to choose the strawberries as focus and then use a rule which says "A pronoun that is syntactically acceptable in place of the focus phrase co-specifies the focus unless the inference mechanism finds a contradiction for that co-specification." *They* is syntactically acceptable for strawberries. An inference mechanism used in conjunction with the focus would need to confirm that strawberries can come from food co-ops, can be fresh, used in cooking and eaten, that is, no contradiction in general knowledge results. If sentence 3 stated that the item mentioned with the pronoun were to be used for cleaning instead of dinner, the inference mechanism would come to a contradiction since the materials used for cleaning do not include food; the focus mechanism on such an alternative sentence would predict that the intended antecedent for *they* is not understandable. The focus mechanism would not need to make use of a structure of shared knowledge, but the inference mechanism would, to know that strawberries are food and fruits of a plant.

The story of focussing begins to show some complexity. Not only is a knowledge structure, in particular a network hierarchy, needed as a part of the focussing mechanism, but an inference mechanism, working in a particular way, is needed as well. That is not all. By D1-5, the discourse is no longer about strawberries; it is about Mark

and his hollow leg. Any reasonable discussion of focussing in discourse must take into account the fact that the focus changes in a discourse. In D1-3 someone is introduced. In D1-4, Mark is specified as the someone so that by the last sentence of D1, Mark has become the center of discussion.

Perhaps, one might claim, it would be better to ignore using a focus for interpretation of the antecedents of anaphora since focussing makes use of so many phenomena. However, this complaint fails to recognize that people use all the phenomena mentioned to determine antecedents of anaphora. Focussing offers a theoretical method for tying them together in a computational process that can be controlled and can make predictions about acceptable and unacceptable uses.

The focus mechanism will work only if it is coherent to talk about some element of the database as being in focus. How can the focus be chosen? This question has a non-trivial answer because there are a variety of phenomena which distinguish what it is that someone is talking about. One such phenomenon is syntax: there are constructions which mark focus, such as there-insertion sentences like (1).

(1) There once was a wise old king who lived on a mountain.

Another phenomenon which marks focus is speech stress and prosodics: it appears that these mark what the speaker is most interested in talking about. In (2) if contrastive stress is put on Jeremy, the hearer might expect that the next sentence will say more about him.

(2) I want one of JEREMY'S pictures.

Another phenomenon is case analysis: certain cases of the verb appear preferred as the place to indicate what is going to be talked about; in (3) the focus of the discussion is likely to be the speaker's turtle.

(3) I got a really pretty turtle this week.

Fourth, certain determiners like *this* and *that* are indicators of what is of interest and what is currently not in focus respectively, as in (4).

(4) I talked with this lady in the credit department, but she didn't give me much help with my order.

Finally, a fifth phenomenon is knowledge associations. If someone is talking about a clock and then [illegible].

(5) The dial isn't very well lighted.

the focus of discussion is still about that clock. It is the use of the focus mechanism and the focussed element which tie all these different phenomena into one mechanism.

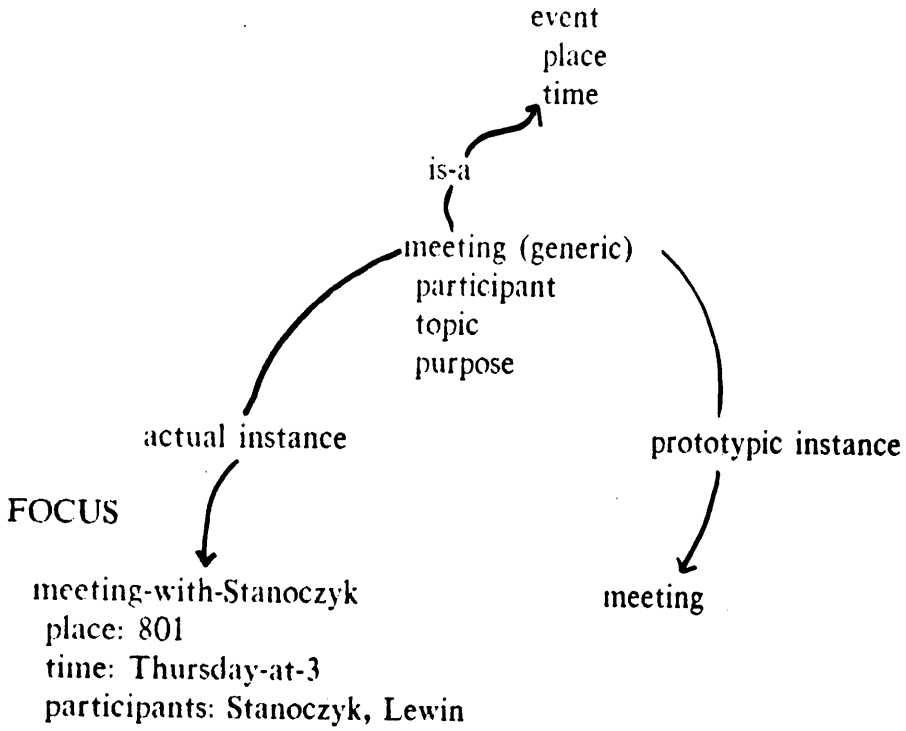
To use these phenomena properly, one must first determine how to find a focus. From a computational viewpoint, an algorithm for determining focus must be provided which takes these phenomena into account. This algorithm is a tool which will be used to state rules about anaphora interpretation. There is no alternative method but to include all the phenomena affecting focus, since without them, predictions of focus may be false. The algorithm which will be presented in this chapter has an important feature. This algorithm orders the phenomena discussed above; that is, each has relative importance depending on what other phenomena are present. This ordering means that not only are all the phenomena tied together, but that the presence of some phenomena make others less necessary in the process of establishing focus.

## 2.3 The Representation of Focus

First let us turn our attention to focus. What is meant in computational terms by the notion of focussing on an element of the discourse? One aspect of focussing is how the focus is represented. The element of the discourse which is focussed on, called the discourse focus, or simply the focus, is represented as a piece of a hierarchical associative network of elements. The network contains many elements, but the focus is the element which is selected as primary among them for a given part of a discourse.

What kind of relations exist among the hierarchical network? Every element in the net is associated with other elements. Some of the associations are "built-in" in the sense that they exist previous to the discourse. For example, **meeting** includes the built-in associations that a meeting has a **time**, a **place**, a set of **participants**, and a **topic** of discussion. The associations are special in the sense that an element has direct links to certain associated elements but not to all other elements; for example **meeting** has no associations to color, cost or age. Each phrase in a discourse is encoded as an instance of the generic element specified by that phrase. Thus *a meeting* is encoded as an instance of the generic network element of **meeting**. Using a hierarchical net, instances of generalized templates can be created, as in figure 1.

**figure 2.1. Instances of a General Meeting Element**




---

What kind of data structure is being suggested in figure 1? The data structure must support two kinds of links with the ability to inherit on both. One link expresses the is-a kind of relation; it allows properties from the conceptual description of one element to be inherited by another. Thus the generic **meeting** is a conceptual element which is-a kind of **event**. The second relation with inheritance captures the notion of an instance. This relation occurs between a conceptual element like **meeting** and a particular meeting like **meeting-with-Stanoczyk**. The instance is more than a kind of meeting, it is a particular copy or instance of its parent node. Instances fall into two categories: actual, representing some instance in a discourse, and prototypic, representing some instance which has traits identical to its parent. Anything which is said about a prototype can be "inherited" in reverse," that is, the properties of the prototype are properties which the parent node has as well. Prototypes will be discussed in chapter 3.

The data structure needed must have other properties. It must also allow for the embedding of structure within structures. If we are told that John is eating an ice

cream cone, the representation must show that the act of eating includes two sub-structures, one representing John and the other the ice cream cone. Since these sub-structures may be of interest separately from the eating, the representation must allow for each to be distinguished from the structure for eating. The data structure must also permit pieces of it to be partitioned off; the purpose of partitioning will be considered later. Finally the data structure must allow for a natural representation of scope of quantifiers; their role is considered in chapter 4.

The characteristics in the data structure are necessary for anaphora comprehension because loss of any characteristic has important effects on what anaphora can be comprehended. As different sorts of anaphora are considered, it will be clear which characteristics are necessary for the interpretation of each kind of anaphora.

Since a theory of focussing requires some representation, the associative network hierarchy has been chosen because it captures some of the basic relations which seem to exist in human memory. The hierarchical net description assumed here is not assumed to be psychologically real. Whether such a representation, or one similar, exists in some way in the human mind can be determined only by empirical study. The net model is assumed here because it makes possible a discussion of focus and associated elements, i.e. it is a "natural" data structure for focus. Because of its usefulness in discussing focus and encoding information, the kind of structure used is plausible and must either be explained away by a complete theory of human knowledge (if such a thing is possible) or be included as part of such a theory. Net structures like the one suggested here are generally part of AI representation languages (see OWL, [Martin, 1978], KRL [Bobrow and Winograd, 1977] among others). The point of this net model is to claim that whichever representation language is used, it must have the features mentioned to perform anaphora comprehension.

The illustration of focus in figure 1 is slightly misleading because it suggests that the focus is just the encoding of the specification of a particular noun phrase. In fact, the process which establishes the specification of a noun phrase in focus must have access to the syntactic and semantic forms of the phrase. There are two reasons for including syntactic and semantic forms. The first is that if syntactic and semantic forms are left out, some anaphors will appear ambiguous, when in fact they are not. Unnecessary ambiguity can be illustrated using D2 below.

D2-1 The first man on the moon became a national hero.
2 Due to his status, he rode in ticker tape parades, met public officials and was chased by autograph hunters.

The focus of this example is *the first man on the moon.* Suppose that the specification of that phrase is the focus as depicted below, without benefit of the referring definite noun phrase.

Database representation of:

NEIL ARMSTRONG FOCUS
Rank: colonel in U.S. Army
Father: 3 children
Achievement: moon walker

The specification of *his status* will be ambiguous because Neil Armstrong has several roles in which he has the status of father, colonel and moon walker. But *his status* in D2-2 is not ambiguous for human speakers. To avoid unnecessary ambiguity, the expression with its syntax and semantics must be included as part of the focussing process. The syntax and semantics of the noun phrase may be thought of as focus constructs. The constructs are used to further highlight that part of the representation which is most relevant in the database of representations. In the case above, the constructs may be thought of as pointing to the representation of Neil Armstrong through a window which is his role as moon walker. Hereafter, focus will be spoken of as being on a particular noun phrase. This is an informal means of referring to the enocding of focus by the focussing process, that is, focus is encoded as a representation built by syntactic and semantic constructs of the noun phrase which points to another representation, its specification.

A second reason for providing the focus mechanism with the syntax and semantics of the expression in the focus motivated by the quantificational aspects of noun phrases as discussed in Webber [1978]. In numerous examples she illustrates the need for scope of quantification and the need to distinguish main clauses from subordinates. For example, Webber claims that to resolve the use of *it* and *she* in D3-2, the scope of quantifiers in the preceding sentence must be identified.

D3-1 Mary gave each girl Bruce knows a crayon.
2 She used it to draw a Christmas card for her mother.

While Van Lehn [1978] has suggested that people do not interpret scope in understanding single sentences, it is clear that to interpret the anaphora in D3-2, people must choose a scope reading for D3-1. Chapter 4 will discuss how Webber's formalism can be interfaced with the focussing process in examples like this.

In addition to the discourse focus, focussing must take into accout the actors of the discourse. An actor of the discourse plays a special role in anaphora disambiguation and is separate from the discourse focus. Actors can be the discourse focus only when no other focus is available, a matter discussed further in the chapter 4. Actors must be specified separately because (1) the focus of the discourse often does not make a direct association to the actor (see example below), and (2) actors can be spoken of anaphorically in the presence of a discourse focus. As a result, different rules for governing mention of actors are needed.

A typical example of an actor focus can be found in D4.

D4-1 Jerome took his pigeon out on a leash.
2 Since he was trying to train it,
3 he hollered "heel" and "run" at it,
4 as they sauntered along.

The actor focus is just whoever is currently the agent in the sentence. When the agent of the next sentence is a pronoun, the actor focus is chosen for co-specification. Jerome is the actor focus in the first three sentences of D4. Using this focus, the co-specification of *he* can be established as Jerome. However, the discourse focus is needed as well since *they* in D4-4 co-specifies both Jerome and his pigeon. This use of both the actor focus and discourse focus is quite common, as shown below, where the discourse focus is needed to establish who actually went to the movies from among the three actors. Anaphora involving actors will be discussed further in chapter 4.

D5-1 I wanted to go to the movies on Saturday.
2 John said he'd come too, but Bill decided to stay home and study.
3 So we went and afterwards had a beer.

## 2.4 The Focussing Process: Finding The Discourse Focus

At the beginning of the discourse, a focus must be determined. The means of finding a discourse focus which will be described here can be seen as a kind of bootstrapping operation. The focus mechanism needs a way to get a starting focus, so

this section will motivate and present an algorithm which chooses a focus for use after the initial sentence of the discourse. Since there are a number of phenomena to consider, all of these will be presented. After they are reviewed, the algorithm will be stated.

On the basis of one sentence, it is not always possible to predict what the focus will be. There are a few indicators of focus, which will be discussed in a moment, which are a highly reliable means of marking focus. When these indicators are not present, the only criteria which remain select a noun phrase or verb phrase as preferred for focus. The rule for focus recognition is stated in terms of preferences, or defaults, chiefly on semantic categories. From the semantic categories the focus recognition algorithm predicts an expected focus. However, predictions about the focus can later be wrong. Luckily, language is constrained in such a way that false predictions are easily recognized, and a better focus can be chosen. This method is an effective means for computing because once a false prediction is recognized, the true focus can be found easily. Thus either the second sentence of the discourse can confirm the expected focus and thereby establish the discourse focus, or the expected focus may be rejected in favor of some other noun phrase or verb phrase from the first sentence which is mentioned again in the second sentence.

Before the different cases are considered, the methodology used for them must be stated explicitly. Pronouns reflect the discourse element in focus; since the pronoun contains little lexical information, whatever it takes as antecedent must be the focus of the previous sentence. Therefore, in each of several examples which follow, a pronoun occurs in the second sentence. Its antecedent must be one of the phrases in the first sentence. That phrase will be the focus. Using this intuitive method, the reader can judge what the focus must be. What remains then is to analyze many cases using this intuitive method. The general form of the argument will be as follows. An expected focus will be suggested for the initial sentence. If it is the antecedent of the pronoun which occurs in the second sentence, based on intuitive judgments of native speakers, the expected focus will be taken as the discourse focus. When the expected focus is unacceptable, it will be disregarded. An unacceptable focus therefore indicates how that syntactic or semantic category fares as a preference for focus.

There are a few syntactic sentence types which make recognition of focus easy

since these sentence types have the purpose of singling out one element from others. These types are cleft, pseudo-cleft and there-insertion sentences as shown below:

(6) (pseudo-cleft agent) The one who ate the rutabagas was Henrietta.
(7) (pseudo-cleft object) What Henrietta ate was the rutabagas.
(8) (cleft agent) It was Henrietta who ate the rutabagas.
(9) (cleft object) It was the rutabagas that Henrietta ate.
(10) (agent) There once was a prince who was changed into a frog.
(11) (object) There was a tree which Sanchez had planted.

As the introductory sentence of a discourse, sentences (10) and (11) provide a means of introducing a new object or agent for further discussion. Sentences (6)-(9) rarely occur as initial sentences in a discourse since they assume there is some object already under discussion about which they provide new information; for example, (6) tells who ate the rutabagas, the rutabagas already being known about. As I will show in depth later on, sentences like those of (6)-(9) move the focus from one element to a new one. These examples suggest that there-insertion sentences mark an initial expected focus.

Semantic preferences for expected focus can be explored by first considering simple syntactic positions like subject and object. These will provide some insight into the nature of expected focus, but will fail to provide a basis for all possible situations. A set of semantic categories will then be considered. A discussion will be given of the kind of semantics which must underlie focussing. This discussion will be followed by examples which show the order of preference for the semantic categories.

To explore syntactic preferences as a prediction of expected focus, let us consider first some very simple sentences. In (12), an intransitive sentence without any prepositional phrase modifiers, the expected focus could be *Ben* or the walking which he is doing. Since walking is hard to discuss without a walker, *Ben* can be considered as a preferred expected focus.

(12) Ben ran.

(12) can easily be followed by (13) where the co-specification of *he* with the expected focus is confirmation of *Ben* as focus.

(13) He was thinking about his toothache.

The verb phrase can also be focussed upon as in (14) following (12):

(14) To do so, Ben had put on his new tennis shoes.

However, the verb phrase is much less frequently a source of focus. It is not a preferred position for focus since a bare *it* co-specifies with the verb phrase only when

other noun phrases are semantically unacceptable (evidence for this claim will be given shortly). The use of *do-it* or *do-so* anaphoric forms is a syntactic way of indicating that none of the sentential noun phrases are the source of the antecedent of *it*; this syntactic means of indicating that the verb phrase is the focus requires the *do* part. In the example above, "Ben" is chosen as expected focus and when the *do-so* anaphora is encountered, the expected focus is dropped in favor of the verb phrase as discourse focus.

The object of an action is the default position among all verb positions for expected focus. Two examples are given below.

D6-1 Mary took a nickel from her toy bank yesterday.
2 She put it on the table near Bob.

D7-1 Sandy walked her dog near a bull one day.
2 He walked quietly along.

In these cases the noun phrase in a prepositional phrase following the object of the action cannot be the focus of the discourse unless the expected focus is explicitly overridden by a full definite noun phrase (defnp hereafter) co-specifying with some other phrase of the initial sentence. In D6, *it* co-specifies a nickel. While it is inferentially acceptable for *it* to co-specify Mary's toy bank (since toy banks can be put on tables), on first reading, people understand the nickel to be the antecedent of *it*. A similar behavior occurs with D7. An example of expicitly overridding the expected focus is given in the alternate form of D6 below.

D8-1 Mary took a nickel from her toy bank.
2 She put the bank on the table near Bob.

The object-of-an-action as focus rule applies to sentences with verb complements. In these cases, the direct object of the verb complement is the expected focus, if a direct object exists. Thus in D9, the meeting is the expected focus.

D9-1 I want to schedule a meeting with Ira.
2 The time should be 3 p.m.
3 We can get together in his office.
4 Invite John to come, too.

To focus on *schedule*, explicitly marked sentences using *do it* or *do so* types of anaphora are needed:

(15) I want it done today.

(16) Do it as soon as possible.

A simple, but revealing question may be asked: What does it mean to speak of the object of an action? Is this a syntactic or semantic category? Consideration of sentences with prepositional phrase modifiers provides an answer. Many sentences do not have an object of an action, in the syntactic sense of direct object. For example,

(17) Please focus on the star of India in the case on the left.

(17) has an object of an action, *the star of India*, but this is not the direct object because of the preposition *on*. If focussing is to predict a focus for verbs with this kind of semantics, then the alternative of a semantic foundation of this verb/noun phrase relation is warranted. This foundation must account for direct object cases as well as sentences with prepositional phrases.

One possible semantic foundation is given in Gruber [1976]. He considers motional and non-motional verbs as part of a theory of lexical relations. He uses the term *theme* to indicate the relation to the verb of a noun phrase which undergoes motion in motional verbs, which plays the role of stative objects in sentences like "The chest is standing in the corner," and the object undergoing change in *buy-sell* pairs. Gruber's lexical relations are developed as part of transformational theory; they are used to indicate underlying similarities in verbs like *buy* and *sell*, and classes of verbs like *go, come, roll*, and *float*.

(18) Bernard rolled the log down the hill.
(19) The chest is standing in the corner.
(20) Mortimer sold the book for 10 cents.

Using thematic relations, the theme of (18) is log, of (19) is chest and of (20) is book.[1] Celcia-Murcia uses theme in a similar way in a semantic network representation in Simmon's [1973] natural language system.

Gruber's concept of theme is problematic because it is not well defined for all verbs. In fact in Gruber's work, he often uses rather complex arguments for deciding what is the theme for a given verb. This ought to give the reader pause; for the following argument might be given against using theme. Perhaps a semantic foundation

---

1. The reader interested in thematic relations is directed to Gruber [1976] and Jackendoff [1972]. Jackendoff presents a good overview of Gruber's work.

should be dropped in favor of syntactic cues only. For many examples, those where the sentence contains a direct object, the direct object serves the same function as theme, and direct object is at least a well established syntactic concept. This argument misses the point of (17). Some account must be given for just those verbs which require a preposition to mark the semantic relation between verb and noun phrase that is similar to the semantic relation of direct object for other verbs. The semantics of theme has the virtue of making possible the determination of underlying focus for these verbs. These verbs include not only *focus on*, but also *talk about, sit on* and *pick up* to name a few. What is needed is a clear definition of theme. That relation which appears most commonly as direct object, but appears in other positions, is theme.

The previous discussion shows that focussing needs an adequate theory of thematic relations. While thematic relations have been discussed in the linguistic literature for reasons independent of discourse anaphora, the argument for a theory of thematic relations due to focussing indicates that there are discourse reasons for a lexicon with certain semantics and for a theory of semantics in addition to the sentential reasons which have been expressed for some years.

However, since a complete account of thematic relations remains to be given, for the purposes of this section, the theme can best be generalized as the verb relation that indicates the property of being affected by the action of the verb. This description is a quick heuristic for use by this author and the reader; to build a full lexicon for English, more complex arguments may be needed for some verbs, but this description will be a rule of thumb for the discussion in this report. The relationship of "being affected" is one of general interest in human communication and seems to be the likely relationship for further communication. The expected focus is simply a means of noting from the hearer's viewpoint that likely relationship. Thus the choice of expected focus requires that for each verb, the case relation which specifies being affected must be stipulated in the verb semantics for inspection by the focus mechanism. In chapter 6, a review of two systems is given in which the expected focus process relied on a specification of the theme for each verb. Hereafter for each verb, the verb/noun phrase position which represents the relationship of being affected will be referred to as the theme.

An example of the theme as expected focus is given by D10 below where the

theme is the vase. In this example the expected focus is the vase and, the *it* of D10-2 (a) co-specifies the vase. In the alternate case of D10-2 (b), *it* co-specifies *wall*, and while the resulting sentence is grammatical, the discourse is odd to native speakers. For some speakers, the preference of vase is so strong that they see D10-2 (b) as co-specifying with the vase and imagine a photograph covered vase. This behavior is explained by the focussing process: the choice of *wall* as focus is unexpected; some hearers will use both the semantics of the verb and prepositional phrase to de-select the expected focus and select the other noun phrase while other speakers will create an image that justifies the expected focus of D10-1.

D10-1 The vase broke against the $\underline{\text{wall}}_i$.
2 (a) $\underline{\text{It}}_i$ shattered into many pieces.
(b) $\underline{\text{It}}_i$ was covered with many photographs.

In a sentence without a theme, i.e., where only prepositional phrases are present and none is the affected object, there does not appear to be a preference for expected focus. Most other thematic positions (instrument, goal and locatives) do not offer a strong preference for focus although some weak preferences sometimes appear. These weak preferences are for goal and any position in which an indefinite occurs.[1] However, it is difficult to know how reliable these preferences are without some means

---

1. Any of D11-1a though g can be followed by D11-2, Each sentence pair has a different behavior for the co-specification of the pronoun *it*. Semantic selectional restrictions can be ruled out for the pronoun because D11-2 is semantically neutral about whether it makes more sense relative to bridges, rivers or lakes.

D11-1 (a) I walked across the bridge over the river.
(b) I walked on the bridge over the river.
(c) I walked from the lake across the river.
(d) I walked over the river on the bridge.
(e) I walked across the bridge to the river.
(f) I walked on a bridge over the river.
(g) I walked over the river on a bridge.
2 I was surprised that it was wider than I had remembered.

In sentence pairs consisting of D11-1a through d and D11-2 (these contain neither a goal nor an indefinite), the use of *it* is ambiguous. There is no means of reading these sentences, other than with additional stress, which indicates whether the second sentence is mentioning bridges, lakes or rivers. For pair D11-1e and 2 (goal only), *the river* is the preferred antecedent. For pairs D11-1f and 2 and D11-1g and 2 (indefinites only), the preference is for *a bridge* as antecedent of *it*.

of determining the role of stress and prosodics in these cases. Therefore, no claims will be made about preference for expected focus for these positions. Instead the algorithm for computing expected focus below will rely on a simple scheme of sentence surface order for these thematic positions.

One thematic position which occurs last in terms of focus preference is agent. In nearly all of the examples in this section (except D10), there is an agent in the initial sentence of the discourse. In several examples, (19) and (13), D6 and D8, the agent is pronominalized in the next sentence in addition to the discourse focus. Only in the sentences (13) and (19) is the discourse focus the agent of the first sentence. In general the agent is the center of discussion in the discourse only when no other discourse element has been mentioned. Hence in the choice of expected focus, the agent is ordered last among possible noun phrase choices. Because of the use of the actor focus, in some discourses, the actor focus and discourse focus will be the same. This behavior is well motivated as will be shown in chapter 4.

The order[1] for expected focus which has been shown in previous examples is:

--subject in there-insertion clauses
--theme
--all other thematic positions with agent last

The previous discussion has mostly considered noun phrase positions as the focus of discourse. But it has been shown that the verb phrase can be the focus as well since *do-so* and *do-it* anaphora can co-specify with the verb phrase. Another type of anaphora, the sentential *it* anaphora, can also co-specify with the verb phrase.

D12-1 Last week, we went out to the lake near my cottage.
2 It was a lot of fun.

The preferred ordering of verb phrase for expected focus is not revealed by the *do-it* and *do-so* anaphora because they are marked syntactic cases. Sentential *it* anaphora

---

1. One comment must be made about the order of expected focus. It is exactly the inverse of the thematic hierarchy used by Jackendoff [1972] to express conditions on formation of passives and of reflexives. Why this is so is not clear and suggests itself as a topic of further research.

examples, however, show that theme and agent are preferred before the verb phrase. Examples such as D13 indicate that sentential *it* anaphora are not preferred as focus when a theme is present since the two uses of *it* co-specify bear, and not the capturing.

D13-1 Mike captured a bear.
2 Everyone said it made a lot of noise,
3 but I was asleep and didn't hear it.

The agent is preferred over sentential *it* as well. In D14, the *it* co-specifies the bear although D14-2 is semantically neutral[1] between bear and the entire first sentence.

D14-1 One of the bears was loose in the park.
2 It frightened all the campers and generally caused panic.

Examples like these support the conclusion that the verb phrase should be ordered last in the expected focus list.

Examples like D15 indicate that noun phrases which indicate manner and the entire sentence as antecedent are ambiguous.

D15-1 Jose walked in a brisk, haughty manner down the street everyday.
2 It annoyed his neighbors,
3 and they often commented among themselves about his snobbery.

D15 is ambiguous between the reading of Jose's manner annoying his neighbors and his walking in a particular manner annoying his neighbors. One conclusion which can be drawn from this observation is that manner is not thematic relation.

One syntactic sentence form is not covered by the use of thematic/case relations for expected focus. In sentences with *is-a* verbs, the expected focus is the subject of the sentence.

D16-1 The Personal Assistant group$_i$ is a research group that is designing pieces of a personal assistant program$_j$.
2 (a) Several graduate students and research faculty are members of it$_i$.
(b) * Several graduate students and research faculty are members of it$_j$.

---

1. By "semantically neutral," I mean that the case position in question does not rule out the use of a particular noun phrase. Thus the agent position of *make* accepts noun phrases which describe both animate beings and events. In essence I am taking advantage of the semantic selectional restriction information on verbs. This restriction was discussed in chapter 1.

While the predicate nominative is being associated with the subject in *is-a* sentences, it does not co-specify the subject. Instead the subject is being described as having some particular properties. Put another way, the subject is being described from a different point of view. However, that description can apply to other things; the *is-a* sentence applies that description to its subject. Since the description is a description of the subject, the expected focus is the subject of the sentence.[1]

To summarize, the choice of expected focus has been shown to require a semantic foundation although a few sentence types can be judged on the basis of syntax. The semantic foundation is an alternative to the approach of Baranovsky [1973] who uses a list of discourse "topics" ordered by recency. As the previous section has indicated, recognizing a focus of discourse is a complex process. The preceding examples have indicated the various special cases. However, the discussion of expected focus would not be complete without specifying algorithmically how the expected focus is found. The algorithm is presented below, with comments concerning the types of data structures and information needed for the process.

**The Expected Focus Algorithm:**

Choose an expected focus as:

1. The subject of a sentence if the sentence is an *is-a* or a there-insertion sentence.
   This step presumes information from a parse tree about what the subject, and verb are and about whether the sentence is there-insertion.
2. The first element of the default expected focus list (DEF list), computed from the thematic relations of the verb, as follows:
   Order the set of phrases in the sentence using the following preference schema:
   --theme unless the theme is a verb complement in which case the theme from the complement is used.

---

1. There are to-be nominal forms which do not contain the focus in subject position:
   (21) A woman with great ideas is Amelia Michels. She is inspiring and works incredibly hard.

As far as I can tell, these forms are a kind of topicalization that is well marked (to the point of being grammatically odd for some speakers). In these cases, the subject is inverted from predicate nominal position and hence the focus is in nominal position instead of in the subject.

--all other thematic positions with the agent last
--the verb phrase

This step requires a list of the surface order of the noun phrases, and a data structure which indicates which noun phrases fill which thematic slots in the verb. Such a data structure must be computed by a case frame mechanism such as the one reported in Marcus [1977].

## 2.5 Examples of the Use of the Expected Focus Algorithm

Now that the expected focus algorithm has been presented, its use on three sample sentences will assist the reader in understanding the algorithm.

(22) I took my sister to the zoo today.

(22) is not an *is-a* or there-insertion sentence. So using step 2 of the expected focus algorithm, a DEF list can be constructed. It consists of:

(theme) my sister
the zoo
today
(agent) I
(full verb phrase) took (I, my sister, the zoo, today)

By step 2, the expected focus of (22) is the first member of the DEF list, which is *my sister.*

Using the expected focus algorithm, since (23) is a there-insertion sentence, its subject, *an old man*, is expected focus.

(23) There once was an old man who lived in the woods.

(24) is a sample sentence where no theme is present; the theme of *talk* is what is talked about, which is not given in the sentence.

(24) Linda talked with her dog all day long.

The DEF list includes:

her dog
all day long
(agent) Linda
(full verb phrase) talked (Linda, her dog, all day long)

By step 2, the expected focus is *her dog.*

These three sentences exemplify different parts of the expected focus algorithm. They show its use on sentences with special syntactic forms, sentences with a theme, and sentences where no theme but other prepositional phrases are present. The expected focus algorithm is significant because it operates with many different kinds of sentences, and chooses from among several criteria which are relevant to the choice of expected focus.

## 2.6 Rejecting the Default Focus

According to the list of the previous section, the theme of the verb phrase in the initial sentence is the preferred default for expected focus. In D17, the theme of *tell* is what is told about, so the expected focus is *new discovery*, which is confirmed by the antecedent of *it* in D17-2.

D17-1 Prof. Salamander will tell his students about his new discovery tomorrow.
2 They have been waiting to hear about it for a long time.

However, the expected focus can be overridden. I have pointed out that full noun phrases can cause this behavior as well as do-anaphora. Another means, nominalization, requires a brief diversion for a full explanation.

A verb focus can be specified by means of a noun phrase which is a nominalization of the verb. The simplest form of these are noun phrases which are morphologically related to the verb, for example *meet* and *meeting*. Other noun phrases are either more general or more specific than the morphologically related verb phrases. For example, if Wilfred buys a book from Alfred, one can speak of the *purchase* or of the *transaction*; each is a nominalization of *buy*. In D18, the lecture is a more specific term than *telling* which is related to *tell*.

D18-1 Prof. Salamander will tell his students about his new discovery tomorrow.
2 They have been waiting for this lecture for a long time.
3 It will begin at 3 o'clock.

How can nominalizations be computed? When the expected focus is rejected, among the other default foci is the verb phrase. Each verb phrase must have some noun phrase associated with it that is the standard nominalization; morphologically related noun phrases, if they exist in the language, are such standard nominals. For verbs which do not have morphologically related noun phrases, a noun phrase which can

serve this role must be chosen and associated with the verb as part of the model of the hearer's knowledge. Because noun phrases are linked in an inheritance hierarchy, generalizations, and instances (more specific terms) can easily be computed.[1] Using such a method, in D18 the focus of the discourse is the lecture, not the discovery.

What happens to an expected focus that is rejected by a following sentence? Is it lost or is there some means of retrieving it? Another version of D17 is given below.

D19-1 Prof. Salamander will tell his students about his new discovery tomorrow.
2 They have been waiting for this lecture for a long time.
3 It will be given at 3 tomorrow.
4 The Board of Regents at the university sent a letter of commendation to the professor.
5 They have been waiting for the results for a long time.

Both *a letter* in D19-4 and *the results* in D19-5 relate to the focus of discovery: the letter commends the discovery, *the results* co-specifies with the results of the discovery. However, if the verb phrase is the only discourse focus, it would force either one of the following claims: 1) the two phrases in question relate to the telling, not to the discovery, or 2) the two phrases relate to the discovery, and this is somehow determined by searching the verb phrase structure when in focus. The first of these claims runs counter to the antecedence intuitions of English native speakers. The second claim is inconsistent since a letter of commendation could be given for one's students as well as a discovery, and there is no method for deciding which is intended when the verb phrase is in focus. How can the anaphoric phrases in D19 be resolved?

Suppose that when the discovery in D19 is rejected as expected focus, it is remembered as having been the primary default. It can be retrieved later if necessary. Then there are two sources for the anaphoric *the results*: the discovery and the telling. Discovery can be chosen because letters of commendation are often given by university governing bodies for discoveries, which have been long awaited. This last fact may seem objectionable because it implies use of general knowledge inferences about discoveries,

---

1. Fahlman [1977] presents a means of computing such inheritance without stepping up the tree link by link, thereby making the computation extremely fast.

letters and universities. While that implication is intended, it is not so objectionable as first supposed: any system, human or machine, which tries to understand discourse must make use of such inferences in order to knit the sentences of the discourse into a coherent whole since not every piece of knowledge can be linked in the association net. The advantage of the schema presented here is that general knowledge is used to confirm a choice of specification, rather than to choose the specification in the first place. Confirmation is simpler because the inferencing has a clearly stated end point. This advantage must not be overlooked since earlier research in computational theories of natural language have been plagued by the problem of when to terminate inferencing (see chapter 1, section 1 for discussion). The inference schema required here will be discussed further in the next section.

As these examples illustrate, the expected focus of the discourse can be foreshadowed by the verb phrase which appeared in the introductory sentence. If rejected, the expected focus must be kept available for possible later use. The expected focus is confirmed or rejected as focus on the basis of the syntactic form of the noun phrase in the sentence following the introductory sentence: full noun phrases show directly whether the expected focus, another noun phrase, or the verb phrase is in focus by the head of the noun phrase; pronominal anaphora confirm the expected focus unless contradictory knowledge exists; and inferential specifications following a foreshadowed expected focus depend on general knowledge.

Just as a choice can be confirmed, general knowledge can also reject a potential choice for co-specification. Expected focus violations involving pronouns are frequent as D20 indicates. There the expected focus is the graduation party, but the following sentence focuses on house.

D20-1 Cathy wants to have a big graduation party at her house.
2 She cleaned it up
3 so that there will be room for everyone who's coming.

Two questions come to mind: How can one recognize that the expected focus is not the focus? How can an alternative noun phrase be chosen as the focus? The use of knowledge inferences is necessary here. In D20 the choice of party for *it* can be rejected since having cleaned up an event would be rejected as incompatible with other knowledge about cleaning. The correct co-specification is available from the previous sentence. To find it, each alternate default focus must be considered in turn, again

using inferences from general knowledge to confirm or reject it. The first such alternate focus is the co-specification of *it*.

The default expected focus can be rejected only when the inference mechanism clearly indicates that the predicted co-specification is unacceptable. That is, the inference must contradict given knowledge from the discourse or be incompatible with other general knowledge. The fact that a noun phrase besides the expected focus might be acceptable as a co-specification is irrelevant as long as the expected focus is acceptable. For example, in D7, repeated below, while the bull might be an acceptable co-specification for the *he*, it is not considered since the expected focus is acceptable.

D7-1 Sandy walked her dog near a bull one day.
2 He walked quietly along.

A matter which is related to the problem of rejecting a focus is how speakers recover from co-specification failures. Consider the following variation on D7:

D21-1 Sandy walked her dog near a bull.
2 She saw how he threw back his great menacing horns.
3 He certainly was an unusual looking dog and the name "Little Bull" fit him well.

After D21-2 the co-specification of *he* seems to be the bull mentioned in D21-1. After the third sentence, the reader is likely to have discovered that the dog has been the focus all along and that this paragraph is a bit bizarre. Virginia Woolfe [1957] points out that literature is interesting for the ways in which authors break rules. This report will only point out what rules can be violated. Why those rules are violated, and how native speakers recover from violations of those rules remains to be explained.

In summary, the expected focus can be rejected in favor of another phrase in the discourse. Rejection is possible only when the predicted co-specification between a definite anaphor and the expected focus is unacceptable. The rejected phrase must be retained for possible re-introduction later in the discourse.

## 2.7 Inferencing in Focussing

Confirming the expected focus often requires inferencing, which can be quite complex. Winograd cites the sentence "The city council refused the demonstrators a permit because they feared violence," and he indicates some of the knowledge needed to

determine the antecedent of *they*. Charniak presents numerous examples of general knowledge, and Isner [1975] presents one approach to handling inference for Winograd's sentences. The crucial difference between these theories and the one presented here is that the focus mechanism predicts what the co-specification is and then inferencing confirms the prediction. A contradiction may be reached, which indicates that the expected focus must be rejected. The inferencing may be trivial: for D21, *he* as co-specifying Sandy's dog is rejected because dogs do not have horns. When inferencing is complex, focussing is advantageous.

Focussing simplifies the inference process because it indicates what the beginning and end points of the inferencing are, and which inference can be taken back if a contradiction results. Schemes such as Isner's depend upon unification to bind the pronoun *they* to a "constant" noun phrase. While Rieger [1974] never states how pronouns are to be resolved, his methodology for inferencing suggests use of unification in a manner similar to Isner. Focus techniques "bind" the pronoun to the specification of the focus and then look for an inference chain that supports the resulting sentence.

Consider cases such as Winograd's sentence and its dual, both given below. The use of actor focus techniques predicts that *they* co-specifies the city council in both sentences.

D22-1 (a) The city council refused to give the women a permit because they feared violence.
(b) The city council refused to give the women a permit because they advocated revolution.

For D22-1a, the inference chain from "city council fears violence" to "city council refuses to give the permit" would be established by antecedent or consequent reasoning of the following form:

Form of reasoning:
find chain of inference from (fear CC violence) to (refuse cc (give CC PT W)).
If (X refuse (give X Y Z)) is defined as caused by either:
Selfish (X)
Want (X Y)
Dislike (X Z)
there is event (W) and Undesirable (X W) and ((have Z Y) --> Occur W)
then the chain of inference must be found between (fear CC violence) and one of

the above.

The first three disjuncts cannot be proven, so a chain between (fear CC violence) and some event, which is undesirable to the council and which will occur if the women have a permit, must be found. If violence is taken as the event, then one can easily deduce that (fear CC violence) --> (Not (want CC Violence)), and (Not (want CC Violence)) --> (Undesirable CC violence). The third conjunct, that is, (have W Permit) --> (occur violence), cannot be established as true, although it is consistent with other information (no contradiction is reached). For focussing, consistency is sufficient, while for traditional schemes, establishing of the third conjunct is necessary. Furthermore, for traditional schemes, the simple chain of inference above would not occur, because it is not known when inferencing begins that it is the city council who fears violence.

For D22-1b, focussing predicts incorrectly that *they* co-specifies with the city council. Since a traditional scheme might choose this co-specification as well on its way to the correct solution, the significance of focussing follows from the control which occurs in inferencing. This claim can be illustrated by the details of the inferencing process. During the process of chaining from the city council advocating revolution to the council refusing to give a permit, a contradiction would be reached about the event of revolution being both advocated by the council and undesirable to the council (using the definition of refusing above). Traditional schemes might look for another event W to inference about, while for focussing, the contradiction reached follows from (advocate CC revolution) which is then retracted in favor of another focus choice, i.e., the women advocating revolution. Hence the search is considerably reduced. Once the choice of women for *they* is made, inferencing is also simpler in the same manner which was shown for D22-1a.

Sometimes there is no other focus choice. If no other groups have been discussed before (25) is uttered, the resulting situation for the focus mechanism is similar to what hearers encounter: the choice of co-specification seems incorrect.

(25) The council refused a permit because they advocated revolution.

The only possible antecedent for *they* is the council, a very odd reason for refusing a permit. However, either the speaker intended to say such an odd thing, or the speaker did not mean to say what was actually said. Thus focussing does not eliminate the need for inferencing; it offers a constraint on how it proceeds. The complexity of the

inferencing is constrained to asking for confirmation of the sentence predication, thereby eliminating combinatorial search for antecedents and non-terminating inferencing.

## 2.8 Focus Sets

Expected focus, though useful in anaphora disambiguation, is limited as a means of anaphor recognition in certain discourses. Consider D23 below:

D23-1 John and Mary sat on the sofa and played cards.
2 Henry read a book.
3 At 10 p.m. they went to Joey's Bar to hear a new rock group.

After D23-1 and 2, the focus of this discourse is not *sofa, cards* or *book.* It appears that D23 is about John, Mary and Henry and what they did for an evening. In other words, the focus in D23 is collected over several sentences. The sentences of D23 give an indication of the focus collection. The expected focus of D23-1 is John and Mary since *sit* is a stative verb which takes the sitters as theme. D23-2 does not mention any of the entities in D23-1, so none can be confirmed as focus. Instead D23-2 introduces a new set of objects, so that the two sentences have no common focus.

The focussing mechanism can be used to choose items on the basis of theme, actor and verb phrase in each new sentence in the discourse and to place these items in a focus set. It can be used to predict a focus once an anaphoric expression occurs. For D23 this method makes sets of (1) cards and book, (2) John, Mary and Henry, and (3) sitting, playing and reading actions. When an anaphoric *they* appearing in agent position occurs, the set of humans agents can be chosen for the focus.

While using this technique, any of the sets could be in focus in subsequent sentences, pronominalization is limited because of potential ambiguity. In D24-4b, *them* is unacceptable perhaps because it can co-specify ambiguously with the places mentioned, as well as with George and Joel.

D24-1 I am going to the MFA this Monday.
2 George is going to see the new market on Tuesday.
3 On Wednesday, Joel is touring the Chinatown section.
4 (a) We are going to have a wonderful time.
(b) ?* Everyone ought to see them when they visit Boston. (where them co-specifies with the places mentioned in 1-3)
(c) Everyone ought to see these places when they visit Boston.
(d) ?* This week is going to be very hectic.
(e) ? It is going to be very hectic. (where it co-specifies with the week under discussion)

These two examples allow the conclusion that collecting focus sets provides the set of phrases and interpretations from which the focus will be drawn, although the actual choice of focus is dependent on which anaphoric phrases occur in the sentence following collection of the focus sets. However, D24-4d indicates that focus sets alone do not account for all the discourse phenomena. For some speakers, D24-4d is odd as the last sentence in the discourse, while if it had been first, the discourse would be acceptable. No focus phenomenon seems to be able to explain this fact.

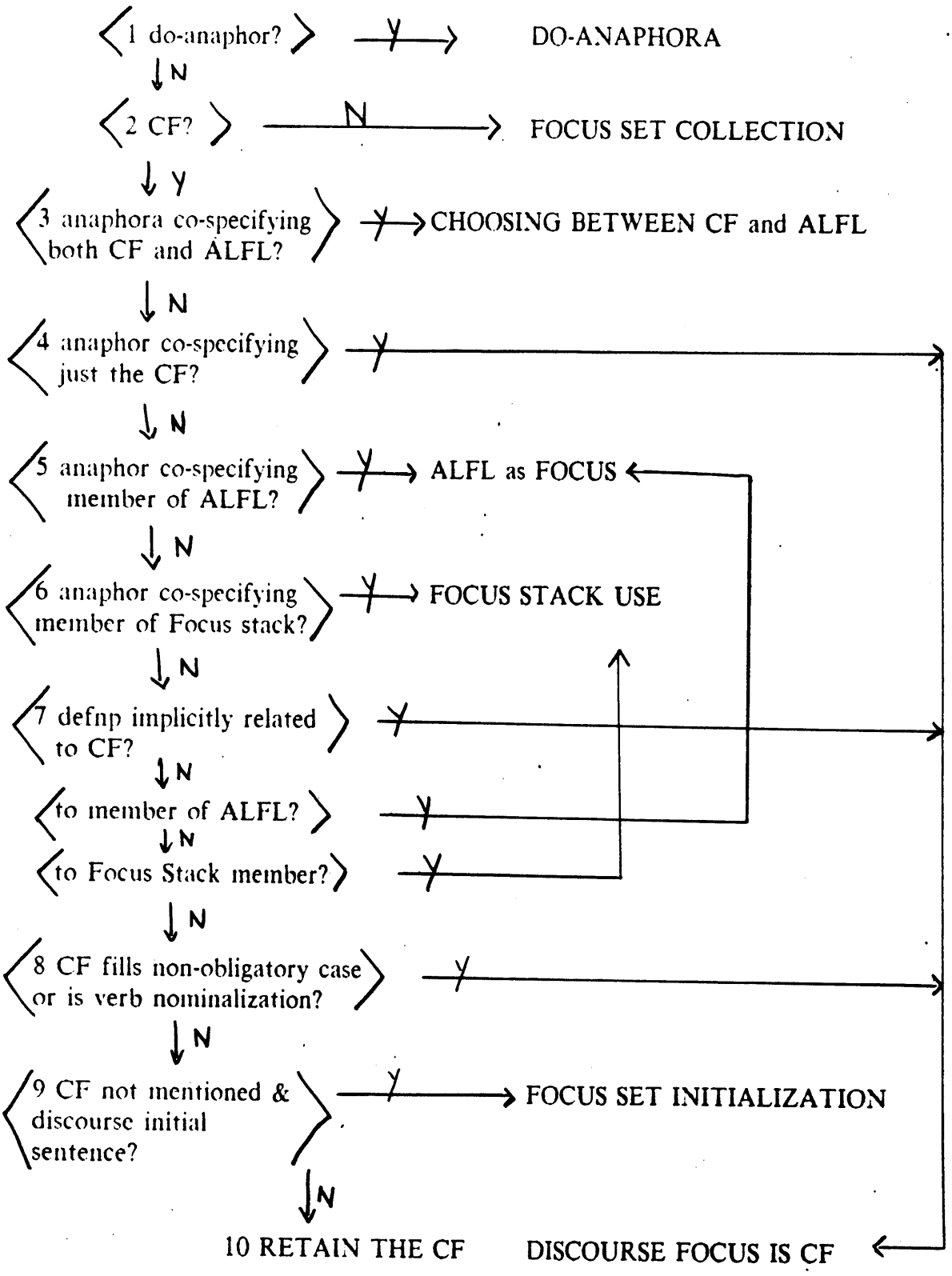
## 2.9 An Algorithm For Focussing

As the previous section has indicated, the expected focus must be confirmed as focus or rejected in favor of some other focus. The algorithm given below performs this task. This algorithm, called the focussing algorithm, uses several data structures: the CF (current focus), the DEF (default expected focus list), the PFL (potential focus list) and the focus stack; the latter two have not yet been introduced, and their purpose will be stated in a coming section. Step 7 of the algorithm makes use of implicit specification, a concept which will be explained in chapter 3. It should also be noted that a discourse is initial when the first and second sentences are under consideration, and in progress otherwise. The statement of the algorithm is preceded by figure 2 which illustrates the control flow.

**The Focussing Algorithm**

A. Steps 1-9 use an ALFL (alternative focus list). It is initialized to be either the DEF or PFL depending on whether the discourse is initial or in progress. A stack called the focus stack is globally available to this algorithm. On first use of this algorithm, the stack is empty.

**figure 2.2. Control Flow for Focussing Algorithm**



B. Set the current focus (CF) to either the expected focus found from the expected focus algorithm or the focus of the discourse when discourse is in progress.

To confirm the current focus as focus or to reject the current focus for another focus in the next sentence of the discourse:

1. **DO-ANAPHORA**: If the sentence contains do-anaphora, take the last member of the ALFL as the focus. Ignore steps 2 through 6. Stack the current focus in the focus stack.

2. **FOCUS SET COLLECTION**: If there is no CF by the initialization in B above, there is an occurrence of focus sets. When no definite anaphora have appeared in the current sentence, continue collecting focus sets. If an anaphor appeared and it is not in agent position, take its co-specification as focus.

3. **CHOOSING BETWEEN CF and ALFL**: If there are anaphora which co-specify both the CF and some member of ALFL, take as focus whichever is not in agent position. If both are non-agents, retain the CF as focus unless only the ALFL member is mentioned by a pronoun. In that case, move the focus to the ALFL member. (Focus is moved by stacking the CF, setting the CF to the co-specification of the anaphoric term, and then stacking any flagged implicit specs as long as that spec is not the spec to which focus moves.)

4. **RETAINING THE CF as FOCUS**: If there are anaphora which co-specify only the CF, retain the CF as focus.

5. **ALFL as FOCUS**: If the anaphora only co-specify a member of ALFL, move the focus to it. If several members of the ALFL are co-specified, choose the focus in the manner suggested by the expected focus algorithm.

6. **FOCUS STACK USE**: If the anaphora only co-specify a member of the focus stack, move the focus to the stack member by popping the stack.

7. **IMPLICIT SPECIFICATION**: If a defnp implicitly specifies an element associated with the focus, retain the CF and flag the defnp as implicit spec. If specification is associated with member of ALFL, move focus to that member and flag the defnp as implicit spec.

8. **LACK OF ANAPHORA**: If there are no anaphora co-specifying any of CF, ALFL or focus stack, but the CF can fill a non-obligatory case[1] in the sentence or if the verb phrase is related to the CF by nominalization, retain the CF.

9. **FOCUS SET INITIALIZATION**: If there are no foci mentioned and the sentence is discourse initial, collect focus sets.

10. **NO FOCUS USED**: Otherwise if there are no foci mentioned, retain the CF as focus. For any unspecified pronouns, the non-antecedent pronoun condition holds.[2]

Several parts of this algorithm will be discussed in greater detail in the chapters to come. Furthermore, the algorithm for actor focus is not included above. It will be delayed until chapter 4. To understand how the focussing algorithm works, its operation in a few examples will be given.

---

1. Obligatory relations are cases of a verb that must be filled or the sentence is odd as in "John sold." Non-obligatory cases need not be filled; e.g. in "John sold a book," one non-obligatory case is the person to whom the book was sold.
2. see chapter 4 for a discussion of non-antecedent pronoun uses.

## 2.10 The Focussing Algorithm for Focus Recognition: Example Dialogues

To illustrate the focussing algorithm in action, its behavior will be traced during the recognition of the initial focus of D1 and D18. Both discourses are repeated below:

D25-1 Last week there were some nice strawberries in the refrigerator.
2 They came from our food co-op and were unusually fresh.

D26-1 Prof. Salamander will tell his students about his new discovery tomorrow.
2 They have been waiting for this lecture for a long time.
3 It will begin at 3 o'clock.

D25-1 is a there-insertion sentence, step so 1 of the expected focus algorithm indicates that the expected focus is the subject of the sentence, i.e., *some strawberries.* The focussing algorithm states that the current focus be set to the expected focus, and the ALFL be set to the other noun phrases in D25-1, i.e., *last week* and *the refrigerator* and the verb phrase. The state of the entire focus mechanism at the point in which D25-2 is encountered, is illustrated in figure 3. The focussing algorithm causes step 4 to be applied to retain the CF of *some strawberries* as focus.

In the example above, rules governing the choice of anaphora given a focus (in this case an expected focus) determine that *they* co-specifies the CF. These rules are discussed in chapter 4. In chapter 3, three kinds of definite noun phrases are discussed,

---

**figure 2.3. Action of Focus Mechanism for the Start of D25-2**

CF: *some strawberries* --> database representation of phrase

ALFL: *last week, the refrigerator,* verb phrase of D25-1

Sentence: D25-2

Anaphora: *they* co-specifies with CF

Processor: skips through steps 1-3
At step 4: CF is taken as focus

those that are anaphoric, those that specify some element given a focus and those that specify an element which has not been mentioned in the discourse. The definite noun phrase rules distinguish the three kinds and provide a specification for the first two. The focussing algorithm is useful for the last kind in a way which will be discussed in chapter 3. *Our food co-op*, an example of the third kind of definite noun phrase, specifies an element outside the discourse context of D25.

The use of the focussing algorithm in D26 requires more testing for conditions. The expected focus is the theme of D26-1, "his new discovery." The state of the focussing mechanism at D26-2 is given in figure 4. Although there are two definite anaphora, *they* and *this lecture*, since *they* is an agent anaphor, it is not considered by the focussing algorithm. *This lecture* will fail to be judged the same as the expected focus; instead it will be judged as a nominalization of the verb phrase. Since a member of the ALFL is co-specified, while the CF is not, the focus will be moved to the verb phase. Note that *his new discovery* was stacked upon rejection.

The focus confirmation algorithm contains some data structures and control decisions which should be noted: an ordered alternate foci list (ALFL), a means of choosing between retention of the focus and movement to a new element of the discourse, the use of the focus stack for rejected current focus, and a means for detecting focus sets. The focussing process occurs after a sentence has been interpreted.

---

**figure 2.4. Focussing At D26-2**

CF: *his new discovery* --> database representation of the discovery

ALFL: *his students, tomorrow, Prof. Salamander*, verb phrase

Sentence: D26-2

Anaphora: defnp *this lecture* co-specifies with verb phrase in ALFL

Processor for Steps 1 - 2: Not applicable so skips past

at Step 3: CF is stacked in Focus stack.
Focus becomes *this lecture* plus specification.

While the algorithm is a simple one, the next section will show that the related data structures and control are central in the use of focus beyond the initial sentences of the discourse.

## 2.11 Focus Movement

Earlier in this chapter it was stated that the focus of the discourse can change. An example of this was given for D1 where the discourse changed from strawberries to the person who ate them. Discourses do discuss different entities. An element can be elaborated for a few sentences; then the discourse can move to a related element or dropped in favor of a new, unrelated element. Accounting for this focus movement is significant for anaphora disambiguation because the new focus has its own set of related database elements which can be mentioned in discourse. How can a focus move? Each additional sentence of a discourse introduces new phrases which can become the focus of the discourse. These new phrases must be associated with the current focus to keep the discourse connected. However, at some point, the associated phrase becomes the focus, because the speaker wishes to say more about it. This process is called focus movement. The discourse below illustrates focus movement from *meeting* to *office* (in D27-3a) and back to *meeting* (in D27-3b).

D27-1 I want to schedule a meeting$_j$ with Harry, Willie and Edwina.
2 We can use my office$_i$.
3 a) It's$_i$ kind of small,
b) but the meeting$_j$ won't last long anyway.

How can one tell if focus movement has occurred? In D27, the only indication of focus movement is that *it* co-specifies with the speaker's office rather than the meeting. Based on this example, one might guess that focus movement works in the following way: whenever a new term is introduced, the focus moves to it. This explanation will account for D27, but in general it yields incorrect explanations. For example, D27 could have been:

D28-1 I want to schedule a meeting$_j$ with Harry, Willie and Edwina.
2 We can use my office$_i$.
3 The meeting will be kind of short,
4 so we could have it in the conference room.

In D28-3, there is no focus movement. If the focus were moved to office and back to meeting, the moves would be unnecessary since all the sentences are about the meeting.

Focus movement is recognized in a manner which is akin to initial focus recognition. Any new term in the discourse is a potential focus. The sentence following its introduction may contain an anaphor which can be confirmed as co-specifying with the potential focus or anaphor which co-specifies with the element already in focus. If the anaphor co-specifies with the potential focus, the co-specification causes the potential focus to become the discourse focus.

Focus movement allows both versions of D27 to be explained: the initial focus is the meeting in D28-1 and D27-1. At that point, offices are a potential object of further discussion, as are any elements associated with meetings. However, by explicitly using *my office*, D27-2 and D28-2 introduce it as a potential antecedent for anaphoric terms which occur in the next sentence. Other elements related to *meeting* cannot be spoken of in this way; for example, the pronoun *it* could not be used to co-specify with the time of the meeting, until a specific time or the phrase *the time* is used. Therefore, the potential focus of D27-2 and D28-2 is *my office*. The third sentences of D27 and D28 behave differently. D27-3 uses anaphoric *it* which can be confirmed as co-specifying the potential focus, and so the potential focus becomes the focus, once the anaphor is interpreted. However, *it* in D28-3 does not co-specify with the potential focus; instead it co-specifies the meeting. Since D28-2 does not move the focus, the anaphors in D28-3 and 4 can be interpreted as co-specifying the focus of meeting. D28-4 by comparison moves the focus again, this time back to the meeting. The defnp form must be used because the phrase "but it won't last long" could be said of an office, and so the antecedent of *it* would be misunderstood. The method of noting the potential focus without moving the focus keeps attention on the focus until it is clear that the focus has moved.

The two versions of D27 indicate the basic point of focus movement: when the sentence following a potential focus contains no co-specification with the potential focus, then the focus movement does not occur. Since a potential focus may not become the discourse focus, focus movement is like focus recognition; if one tried to predict a focus movement before the anaphor occurred, the prediction could be wrong just as predictions of expected focus can be. Rather than consider focus movement a matter of prediction, it is best to think of it as a matter of recognition based on the anaphoric terms that follow in the discourse.

Potential foci have a short lifetime. If a potential focus does not become the focus after the interpretion of the sentence following the one in which the potential is seen, it is dropped as a potential focus. For example, at D28-3 *my office* is dropped. Hereafter if *office* is discussed, it cannot be referred to using *it* until some sentence re-introduces *office* as a potential anaphoric term.

As with co-specification with focus, general knowledge is needed to confirm the co-specification of an anaphor to a potential focus. Example D27 also provides some insight into the nature of the confirmation of *it* as the co-specification of the potential focus rather than the focus. The nature of the assertion on which *it* centers is significant; it must be possible for an assertion of smallness to be made about *my office* before an anaphoric connection can be assumed, and the focus moved.

Since any new term in a sentence can be a potential focus, when several terms occur in one sentence, some means of choosing a potential focus is needed. One would expect that the theme should be the preferred position for a potential focus. In fact the same order used for default expected foci can be used for potential foci, except that the phrase which confirms the focus; it is not included in the list because it cannot be a potential focus. In English sentences, phrases which mention new information tend to occur towards the end of the sentence, while old information occurs at the beginning; this method of choosing potential focus captures that behavior.

How strong is the focus/potential focus expectation? An indication of the strength of focus is given in the example below.

D29-1 Expert: Take off the bolts.
2 Apprentice: I am loosening them with the pliers that used to be in one of the tool boxes. Where are they?

The use of *they* in the last sentence is bizarre. The expected focus is bolts and is confirmed by the use of *them* in the apprentice's first statement. The apprentice also introduces the pliers as a potential focus, but *they* cannot co-specify to either the pliers or the bolts (because the apprentice knows where they are in order to do the task). Hence the use of *they* is strange. Some informants say that *they* could co-specify the tool boxes but that such a choice is forced upon them only as a last resort to find something that makes sense. Bizarre readings can result from failure to use the focus or potential focus as the co-specification of an anaphoric pronoun.

In general, focus movement occurs only when the definite anaphora co-specify something besides the focus. This behavior is due to the fact that what is currently in focus remains the focus until there is reason to assume a movement has occurred. However, this behavior, as one might expect, is not hard and fast. Two conditions influence what remains in focus. The first condition is part of the rules for pronouns, a recency rule which is applied when a pronoun is in a particular position. The rule will be discussed in chapter 4, but briefly stated it is this:

> (Recency Rule) If the pronoun under consideration occurs in the subject position, and there is an ALFL noun phrase which occurs as the last constituent in the previous sentence, test that ALFL phrase for co-specification before testing the current focus. If acceptable both syntactically and inferentially, choose the ALFL phrase as the co-specification of the pronoun.

This rule indicates that the focus is not always focussed on because there is a circumstance where it is not the default source for pronoun co-specification.

The second condition influencing focus concerns how the focussing algorithm proceeds when two different anaphora co-specify the focus and a potential focus respectively and only one of the anaphora is a pronoun. There is no certainty about which will be focussed on in the next sentence. Yet, there seems to be a preference for the focus to be marked by the pronoun use. In other words, pronouns seem to support a focus more strongly than defnps do. Consider the case below:

D30-1 I got a new hat
2 and I decorated it with a big red bow.
3 (a) I think the bow will brighten it up a lot.
(b) I think it will brighten up the hat a lot.
4 If not, I guess I'll use it anyway.

After D30-2, the focus is the hat, co-specified by *it*, and the potential focus list includes *a big red bow*. Either form of D30-3 uses anaphora which co-specify the hat and the bow. D30-4 is syntactically and semantically neutral on the choice of hat or bow as antecedent of *it*. However, if D30-3a proceeds, the *it* co-specifies hat, while if D30-3b is used, the bow is slightly preferred. This example suggests that unlike the general case, the element co-specified by the pronoun should become the focus. The second condition for focussing then requires that the focus move whenever both the current focus and a potential focus are co-specified but only one of them is co-specified by a pronoun; the

condition then requires that the focus move to the pronoun co-specification. This condition appears in the focussing algorithm as step 3.

This section has given an informal description of focus movement and shown the similar behavior between focus recognition and focus movement. This informal description is supported by a formal one, the focussing algorithm. That algorithm provides a process description of focus movement in discourse as well as focus recognition. The algorithm has a simple flow of control which serves to distinguish the several different focussing behaviors.

The significance of a single algorithm for both processes must not be overlooked: the one algorithm provides a uniform treatment of two phenomena which seem to be directly related, namely, focus recognition and focus movement. Furthermore, the one algorithm indicates just how the two processes are similar. The similarity in focus recognition and focus movement can be extended beyond the parallel between default expected foci and potential foci since there are syntactic structures which mark focus movement just as there are syntactic forms which mark initial focus. More importantly, the focussing algorithm shows that focus recognition and focus movement are both processes which require additional mention of the element in focus to confirm the movement.

However, there is one behavior which is unique to focus movement, backwards focus movement, which is the topic of the next section. Once it is described, and syntactic markings of focus movement are presented, the use of the focussing algorithm in focus movement can be considered with examples.

## 2.12 Backwards Focus Movement

Discourses sometimes return discussion to a previous focus. In focussing terms, a focus may eventually shift back to a noun phrase previously in focus. This process is called focus popping. In D27-3b the phrase *the meeting* co-specifies a meeting previously in focus. To retain previous foci, a stack offers a reasonable computational metaphor for the behavior. Generally whenever an expression mentions a phrase represented as a focus in the stack, called the stacked focus, the focus is popped, and the stacked focus becomes the focus again. In terms of the focus movement algorithm, a stacked focus is

considered as a possible focus following the discourse focus and potential foci list.

To claim that focus popping is in fact a stack behavior requires criteria for explaining why other behaviors cannot and do not occur. The basis for such claims requires further investigation which has not been undertaken here. Focus popping is described as a stack behavior because dialogues do return back to a previous focus without concern for intervening foci, and because once a focus pop occurs, the intervening foci are not mentioned without focus movement similar to the regular focus movement.[1]

A sample discourse will indicate why a stack notion seems to be the right one. In the discourse below, the focus moves from Wilbur to the book, to relativity theory, to quarks, and to elementary field theory. Then a pop back to the book occurs. Once the pop is made, Wilbur can be co-specified by *he* easily. A stack representing the foci at the time that D31-8 is processed is given in figure 5.

D31-1 Wilbur is a fine scientist and a thoughtful guy.
2 He gave me a book a while back which I really liked.
3 It was on relativity theory.
4 It talks about quarks.
5 They are hard to imagine,
6 because they indicate the need for elementary field theories of a complex nature.
7 These theories are absolutely essential to all relativity research.
8 Anyway, I got it while I was working on the initial part of my research.
9 He's really a helpful colleague to have thought of giving it to me.

In the more common cases of focus movement, popping back to an old focus is accompanied by the use of a definite noun phrase to specify the old focus. The definite noun phrase is a clear signal of what is being talked about because of its distinguishing noun phrase head. Anaphors can be used as well, though their use must be distinguished on different grounds. In general, a constraint, which will be called the stacked focus constraint, holds: since anaphors may co-specify the focus or a potential focus, an anaphor which is intended to co-specify a stacked focus must not be

---

1. Current implementations of the focus mechanism use a simple last-in first-out stack to pop foci, since this type of stack reflects the passage of foci through a discourse.

**figure 2.5. Stack of foci at D31-8**

FOCUS: elementary field theories

FOCUS STACK:

| quarks |
|---|
| relativity theory |
| book |
| Wilbur |

---

acceptable as co-specifying either the focus or potential focus. If, for example, the focus is a noun phrase which can be mentioned with an *it* anaphor, then *it* cannot be used to co-specify with a stacked focus. An example which does not violate this constraint is D32. There the focus is a career in law. The focus moves to the friends of the speaker with a potential focus of *interesting cases*. The *it* in the last sentences refers back to law career and re-establishes career as focus.

D32-1 A: Have you ever thought of a career in law?
2 B: I have some friends who are lawyers. They really work on interesting cases, but I don't think it's for me.

The D31 example given previously suggests that pronouns may not be readily understood in popping to an old focus. Some readers find D31-8 an abrupt transition, especially if *anyway* is deleted from the sentence. Though hearers can discern the co-specification of *it* in D31-8 because of the stacking behavior, the stacked focus constraint may need to be extended to include a constraint which prevents pronouns from being used to co-specify more than one position back in the stack. In addition transition words may be needed to move back more than one position. Reichman [1978] lists an entire collection of words such as *anyway* which she calls clue words. She uses these to account for the manner in which speakers shift from a focus space with one focus to another with a different focus.

The use of anaphors which violate the stacked focus constraint will confuse a computer using the model described here as well as human speakers. Violations of this rule which result in peculiar sentence interpretations often cause "recovery" procedures to be used; hearers do try to find co-specifications for anaphors that violate focus rules by testing out old foci. In these cases, as in other recovery cases, general knowledge is used to confirm that a focus stack element is the intended co-specification.

However, a pronoun can be used another way; namely, to move the discourse back to a focus in the stack, even when that pronoun can co-specify the focus or a potential focus. This apparent violation of the stacked focus constraint is not a violation at all; the focus moves due to other levels of structure which exist in discourse. These structures make it possible to use a pronoun to move the focus back without confusion. These cases will be discussed in chapter 4 where pronoun anaphora are treated in detail, and in chapter 6 where an implementation of focussing and of discourse control is presented which interprets examples like the one below.

D33-1 E: Bolt the pump to the platform.
2 A: Where are the bolts?
3 E: They are in the tool box.
4 A: What tools should I use?
5 E: The ratchet wrench.
6 A: Where is it?
7 E: It's on the table.
8 A: I found it.
9 A: Okay. It is bolted. What should I do now?

## 2.13 Syntactic Constructions for Focus Movement

Clefts, pseudoclefts and there-insertion mark focus movement by a syntactic means, as discourse D1 repeated below indicates.

D34-1 Last week there were some nice strawberries in the refrigerator.
2 They came from our food co-op and were unusually fresh.
3 I went to use them for dinner, but someone had eaten them all.
4 Later I discovered it was Mark who had eaten them.
5 Mark has a hollow leg, and it's impossible to keep food around when his stomach needs filling.

A there-insertion is used to introduce the initial discourse focus of strawberries in D34. Then the cleft sentence, D34-4, is used to introduce a potential new discourse focus, Mark. (26) would have been equally acceptable in place of D34-4, since it uses a

pseudocleft agent. (27) and (28), the object forms of clefts and pseudoclefts, are unacceptable.

(26) The one who ate the strawberries was Mark.
(27) It was the strawberries that Mark ate.
(28) What Mark ate was the strawberries.

The object form of cleft and pseudoclefts above center attention on the already established focus, but since the focus is already clear, this re-emphasis results in an odd discourse. Clefts, pseudoclefts and there-insertions which focus attention on an item already in focus are unacceptable. One general characterization of discourse can now be stated: Emphasize new material and de-emphasize old. Low pitch and amplitude in speech (see Chafe [1976]), pronominalization and the focus-movement syntactic forms are all instances of this tendency.

To summarize this section, the algorithm for determining the potential focus list (PFL) is given:

> 1. If a cleft or pseudocleft sentence is used, the potential focus is the cleft item if and only if the element in non-clefting position co-specifies the focus. When is does not, the sentence is incoherent.
> 2. Otherwise order a potential focus list of all the noun phrases filling a thematic relation in the sentence, excluding a noun phrase in agent position and the noun phrase which co-specifies the focus if one exists. The last member of the PFL is the verb phrase of the sentence.

## 2.14 The Focussing Algorithm for Focus Movement: Example Dialogues

Since the focus stack and PFL have been defined, the focussing algorithm can be used to determine focus movement in a discourse. An example of focussing with focus movement is given below.

D35-1 Alfred and Zohar liked to play baseball.
2 They played it everyday after school before dinner.
3 After their game, Alfred and Zohar had ice cream cones.
4 They tasted really good.
5 Alfred always had the vanilla super scooper,
6 while Zohar tried the flavor of the day cone.
7 After the cones had been eaten,
8 the boys went home to study.

Using the expected focus algorithm, the expected focus for D35-1 is baseball (it is the theme of the verb complement). There are two pronouns in D35-2, but only one is considered by the focussing algorithm because *they* is in agent position. As shown below, the use of *it* confirms the expected focus as focus.

---

CF: *baseball* --> database representation of phrase

ALFL: verb phrase

Sentence: D35-2

Anaphora: *it* co-specifies with *baseball*

Processor at Steps 1-2: Not applicable

at Step 3: CF taken as discourse focus since
anaphor co-specifies with CF

---

D35-3 also mentions baseball, by means of the defnp *their game.* This use is a case of lexical generalization of focus which will be explained in chapter 3; such uses are important because there are constraints on what defnps can be used to co-specify with the focus. D35-4 shows a movement of the focus.

---

CF: *baseball* --> database representation of phrase

ALFL: *ice cream cones*, verb phrase

Sentence: D35-4

Anaphora: *they* co-specifies with *ice cream cones* in ALFL

Processor at Steps 1 - 4: Not applicable

at Step 5: Since no anaphor co-specifies with CF, stack CF and take CF as *ice cream cones* plus its specification. Discourse focus is CF.

---

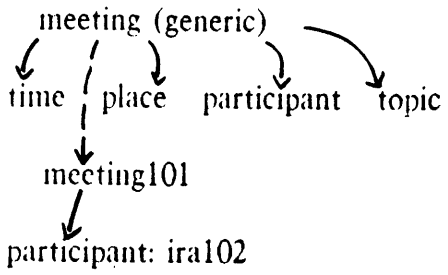
D35-4 contains the anaphor *they*. Since it does not co-specify with the focus, the ALFL (set to the PFL by Part A of the focussing algorithm) contains *ice cream cones*, an acceptable co-specification for *they*, so *ice cream cones* is confirmed as focus. The old focus of baseball is stacked in the focus stack. This is how the focus moves. The remaining two sentences talk about individual ice cream cones. This phenomenon, called co-present foci, is the topic of chapter 5.

Using the focussing algorithm to establish focus, the example discourse given in chapter 1 can be analyzed for its use of focussing for focus establishment and anaphora interpretation. In this example, an expanded representation of the focus is given, so that the reader may understand how the representation of focus is used in focus recognition.

D9-1 I want to schedule a meeting with Ira.
2 The time should be 3 p.m.
3 We can get together in his office.
4 Invite John to come, too.

For each sentence, the focus must be re-confirmed by the focussing algorithm. D9-2 re-confirms focus by step 7: the defnp *the time* is associated with a meeting by means of the associative network hierarchy of elements depicted for meeting in figure 6; associations are explained fully in chapter 3. D9-3 re-confirms focus by step 8; meeting is a nominalization of the verb *get-together*. D9-4 re-confirms by step 8 as well since meeting fills the non-obligatory case of *invite* which takes the event one is invited to as

**figure 2.6. Focus mechanism with full database representation of focus**




ALFL: verb phrase

Sentence: D9-2

Anaphora: (none) Related defnps: *the time*

Processor at Step 1 - 6: Not applicable

Step 7: Defnps: *the time* specifies by means of
the focus, so CF is retained as focus.
Defnp is flagged is implicit spec.

---

its filler.

Once the focus is established, it can be used to provide co-specification for those definite anaphora which are uninterpreted and to indicate which anaphora are ambiguous. In D9, pronoun rules using focus will establish *his* as co-specifying with Ira, who is one of the participants of the meeting which is in focus. There are cases where more than one possible co-specification will be found using this rule, as D36 shows:

D36-1 We are going to have a big dinner.
2 John will be the cook, and Wilbur the baker.
3 I will be the wine master.
4 We will eat at his house.

However, in such cases the co-specification for *his* is ambiguous; the rules using focus predict results similar to those where human speakers and hearers have difficulties in interpreting the co-specification. One of the purposes of this report is to establish rules for acceptable co-specifying phrases in a discourse. The next three chapters will

consider different kinds of anaphora and rules for each.

## 2.15 Conclusions

In this chapter many new concepts have been introduced and observations made. First the informal concept of focussing in a discourse was observed as a part of language behavior. Focussing was then formalized as a computational process which defines how a focus is found and moved. It was observed that many phenomena influence the choice of antecedent for an anaphoric expression: syntax, phonological stress and prosody, semantic knowledge and general inference. Focussing has been proposed as a method for tying together these phenomena. In particular, an algorithm has been constructed which determines how to set the focus and how to move it, with supplementary algorithms for expected focus, potential new foci and default foci. Focussing controls the inferencing associated with interpretation of the antecedents of definite anaphora because inferencing is used only to confirm the antecedent which has been chosen for a definite anaphor by focus based rules. Use of a database associative network hierarchy with a representation of the in-focus phrase allows the focus mechanism to make use of the implicit connections between phrases in speakers' discourses. The algorithms and the related data structures provide a detailed explanation of focussing and focus; they are both an extended definition of the focus concept and an explanation of the focus in language behavior.

It was stressed in the beginning of this chapter that the focussing mechanism must perform two tasks for anaphora disambiguation: it must simulate the hearer's behavior in understanding anaphora, and it must simulate the hearer's lack of understanding of certain bizarre cases. Using the focussing algorithm several examples of focussing in discourse have been considered which simulate the hearer's understanding as well as indicate which discourses are bizarre. More remains to be specified in the chapters ahead, where rules for definite anaphora using the focus are given.

In presenting the focussing algorithm and its use on discourse examples, a number of phenomena have been mentioned, which are listed below. Each of these phenomena will be discussed in the chapter indicated. These phenomena enrich the use of focussing and provide for new insight into the nature of definite anaphors in English

discourses.

1. Constraints on defnps: chapter 3

2. Defnps which mention elements outside the discourse: chapter 3

3. Actor focus: chapter 4

4. Rules for pronoun anaphora: chapter 4

5. Discourse structures which allow pronouns to "violate" the stacked focus constraint: chapter 4 and chapter 6

6. Co-present foci: chapter 5

7. Use of *this* and *that* for focussing: chapter 5

## 3. Comprehension of Anaphoric Definite Noun Phrases

### 3.1 Co-specification and Definite Noun Phrases

Definite noun phrases have posed difficulties in language research for a long time. This chapter has several goals for explaining the behavior of such phrases. First, definite noun phrases used to refer will be considered in some detail, for the purpose of explaining some of these difficulties. Second, a subset of the definite noun phrases, that is, definite noun phrases that are used anaphorically will be considered.

The greater part of this chapter will develop an explanation of anaphoric definite noun phrases; in particular, this chapter will provide two innovations. The first is a set of rules and an explanation of the variety of ways in which definite noun phrases (hereafter called defnps) may be used to co-specify with another noun phrase. The second is a presentation of what must be part of the hierarchic network which acts as a database of knowledge of the world for the hearer. With this network, rules will be given for the use of defnps which specify parts of the network and which are related to defnps (and therefore knowledge specifications) already present in the discourse. These innovations allow an explanation of the following discourse and its definite noun phrases.

D1-1 I want to have a big party$_i$ with lots of guests.
2 The party$_i$ ought to be on Saturday so everyone can come.
3 The food will be catered by the A & B Munchie Makers,
4 and the champagne will come from BLM.

In D1, *the party* is an anaphoric defnp, while *the food* and *the champagne* are not. Yet the two latter phrases, when interpreted, are related to the party in a consistent way. It is the behavior of these kinds of defnps which may be explained using the rules and representation discussed in this chapter. Given these rules, it is also possible to distinguish when a defnp is not being used anaphorically. The completion of the goals of this chapter will provide an approach to some of the difficulties concerning defnps which have been so puzzling.

The behavior of defnps will be clearer if the use of these "referential terms" is understood informally. To begin, some questions must be asked about referring terms, questions for which a theory of definite anaphora comprehension should provide an

account.

(1) When does a noun phrase refer to someone or something?
(2) What conceptual element in the memory or the database of the hearer's knowledge, if any, is specified by a non-anaphoric definite noun phrase?
(3) When does a given anaphoric expression co-specify with the same element as another expression?

The first question may seem a trivial one: don't all expressions refer? In fact, as chapter 1 suggested, many do not, and there are a variety of ways expressions may be used. This chapter will indicate which there are because determining how an expression is used is important for co-specification comprehension.

The second question will not be answered in a general way. A general answer would tell us how to find the specification of any definite noun phrase used by human speakers. The general answer cannot be given here because it requires a theory of knowledge representation semantics which has not yet been provided, though Smith [1978] has investigated some aspects of it. Knowledge representation must determine a structure which reflects one to which people seem to have access in their own minds. Such a theory may provide the framework for an explanation of how people know when an expression specifies something they already have some knowledge about. This chapter will present some parts of that structure which seem necessary, given the use of anaphoric definite expressions in English.

This report seeks to answer question (3). The answer depends upon (1) and (2) because the answer to (3) requires knowing first that a given expression specifies something. A definite noun phrase used anaphorically co-specifies with the first use of the referring expression, which itself specifies an element of the database. So, for example, if a speaker mentions Wilbur in one sentence and then speaks of him as "the guy" in the next sentence, the first use may be modelled as a specification in a database, and the second use may be modelled by a co-specification with the first element.

In some cases, the first expression used in a discourse does not refer, but bears a special relationship to another expression, often an anaphoric definite noun phrase. The anaphor may actually be used to refer to something, and the first expression helps the hearer decide what that entity is. Thus if someone is told,

D2-1 Hernando bought some cookies at the store.
2 The cookies were rotten.

the hearer knows what *the cookies* is intended to refer to even though the first expression *some cookies* may not specify any existing structure in the database.[1] In the technical sense defined in chapter 1, *some cookies* has no specification which exists in the database prior to the discourse. Instead a specification is constructed in the comprehension process. The expression *the cookies* also specifies a database element, and this element is the one described by *some cookies.* In other words, the first use says "I'm going to introduce some cookies that Hernando bought," while the second says something about those particular cookies. Thus the definite noun phrase use in D2 is also a co-specification. By describing database specifications, we are not committing ourselves to their existence in the world, a philosophic position of problematic nature. Instead we are committed to the idea that the use of a noun phrase creates an element in a person's mind (or by analogy in a computer database) which may then be talked about. Establishing the co-specifications is of interest computationally because it makes possible the use of that representation for machine to human communication.

In the remainder of this chapter I will consider anaphor comprehension only from the hearer's point of view. Thus in discussing anaphora and other referring expressions, I will be concerned with a model for how the hearer disambiguates these expressions used in discourse. By symmetry, one might suppose that the generation of referring expressions by a speaker could make use of a similar model. Such a supposition remains untested in this report and is to be verified by later work. Furthermore, I will not be concerned with *comprehension as intended*, defined in the next section, since this process requires the additional information of what the hearer believes that the speaker knows about. However, I will point out at various times how the theory under discussion would need modification if the hearer's beliefs were included.

---

1. In the philosophical literature, *some cookies* is held to be an expression which does not denote anything. This notion of referring may seem counter-intuitive to the reader because the phrase creates such a strong image. It is exactly the sort of mental activity which the reader experiences which is captured by talking of specifications.

## 3.2 The Nature of Non-Anaphoric Definite Noun Phrases

### 3.2.1 The Indended Referent of an Expression

First, the nature of question (2) must be explained further, since the nature of the element specified affects how defnps may be used anaphorically. To make the discussion of specifications perfectly clear, first consider how the specification of names takes place. Suppose the expression *Julius Caesar* is used to refer, and to the person Julius Caesar for whom there is some representation *X* in a hearer's memory.[1] To answer the first question above, the hearer must decide that names are referring expressions. To answer the second question, the hearer must decide 1) whether *Julius Caesar* refers uniquely and 2) what conceptual element in the hearer's memory represents the hearer's real-world referent, namely *X*. These two decisions together with the initial assumptions appear to be necessary and sufficient conditions for comprehension; by deciding that *Julius Caesar* refers uniquely and choosing a conceptual element, the hearer has decoded the entity to which the speaker was referring.

There are, however, situations where the hearer's decisions about the referent and the speaker's intended referent do not coincide. Suppose the hearer decides that *Julius Caesar* refers uniquely to Julius Caesar, who was a Roman emperor. The speaker may also have intended it to refer uniquely, but to this author's deceased cat, whose name was Julius Caesar. Now there are three possibilities: either the hearer knew about Julius Caesar the cat, but decided the expression referred to Julius the emperor; or the hearer only knew about the emperor; or the hearer knew of neither, but instead just connected the name and some other information by chance. The last possibility does not fit a description of reference comprehension. Randomly connecting information between some representation in one's memory and a referring expression may be a cognitive act, but intuitively one would probably call it confusion.

---

1. I will use the terms *memory* or *database* or *knowledge* to refer to the collection of information somehow stored internally which a person uses to function in the everyday world. *Element* will be taken to be a piece of that memory while *entity* will be used for objects in the real world.

Comprehension of referring expressions must make some account of the semantic connection between words and a representation of those words.

In the case where the hearer only knew about the emperor, it seems safe to conclude that the reference may have been comprehended, but *incompletely and incorrectly*. There are many other clues in communication about the referent of terms, in addition to those given by referring expressions in isolation. Without these clues, comprehension of referring expressions is incomplete because the hearer has no means of knowing whether s/he may have the wrong representation. Even with the best set of clues, the hearer may still choose Julius the emperor. Here we will say that comprehension has taken place, completely but incorrectly, because the hearer has used all the relevant communication knowledge to decode the speaker's message. We may conclude that the speaker's rules for generation of referring expression and/or the speaker's knowledge of the hearer, are faulty (thereby contradicting the speaker's assumption given in chapter 1).

In the case where the hearer knows of both possibilities and chooses the incorrect one, the hearer may have erred due to failure to follow other communication clues, or again because the speaker's rules and knowledge were lacking. In conclusion, a referring expression is *comprehended as intended*, if and only if the same specification as that intended by the speaker is chosen from the elements in memory. The expression is otherwise just *comprehended* when the hearer chooses an element from memory which is specified by the referring expression using all the available communication clues, but does not choose the same element as intended by the speaker. An expression will be considered *incompletely comprehended* if the hearer fails to use all the communication clues available at the time the communication occurs. A discussion of some of these clues are will be covered later in this chapter.

Previously I have not considered the possibility of error on the part of the hearer because of the hearer's beliefs. Suppose, for example, that the hearer believes the speaker hates to even speak of cats. Then the hearer may conclude that *Julius Caesar* is most likely a reference to the emperor of Rome. I am not going to account for this possibility in the forthcoming discussion; instead I will restrict the discussion to clues from the communication process. Hearer beliefs raise a separate set of philosophical and computational problems, and would extend the scope of this study too broadly.

However, the issues are significant in the total picture of comprehension of referring expressions.

### 3.2.2 Difficulties with Definite Noun Phrases

Definite noun phrases, like names, may be used to refer to entities in the real world. Russell [1905] says of the expression *the author of Waverly* that it denotes Sir Walter Scott, and that when it is strictly used, a defnp denotes uniquely.[1] Thus by using a definite article, a speaker is often saying in effect "there is one object in the world denoted by the phrase that follows and I mean that one." Of course a defnp may be *used* to denote someone without *actually* denoting anyone, as is the case with *the woman who wrote Waverly*. This defnp is used to refer to someone, but there is no entity in the world, assuming the current state of the world[2] which corresponds to that description. Nothing in the syntactic or semantic form of the expression itself suggests that the expression has no denotation. How can the hearer determine whether the defnp refers to someone or not? Of course, if the memory element which is the specification of the author of Waverly has the property of being male, the hearer may decide that the expression does not refer to anyone. But if no memory element exists, the hearer cannot decide whether the expression refers to anyone. Looking at defnps in this way, however, fails to account for all the phenomena of defnps, because it involves a false assumption.

The Russellian analysis is inadequate for comprehension because defnps, unlike names, are not always used to refer! A particular defnp may not only fail to denote a real world object, but also the defnp may be used without the speaker intending to refer. Even more surprising, a defnp may be used to refer, but the speaker may not expect the hearer to know the referent of the defnp; the defnp form is used to indicate that the referent is knowable, but possibly not significant for the communication at

---

1. By *strictly used* Russell means used without ambiguity.
2. Possible world semantics will not be discussed here. Issues of transworld identity for counterfactuals may require extensions to the representational theory and model given here to deal with non-existing entities which are presumed to exist. These cases have not been investigated in this work.

hand. Donnellan [1977] shows that some defnps are used attributively. If we happen upon Smith who lies dead with foul wounds, one can say "Smith's murderer is insane." Used attributively, *Smith's murderer* does not refer to anyone, and the phrase does not describe a particular existing person. It is as if to say, Smith was murdered and the murderer, whoever that may be, is insane. Thus the speaker using an attributive defnp does not assume that someone fits the description; the speaker is describing for the hearer a set of properties which will help the hearer determine the referent if one actually occurs later. Using a referential defnp, the speaker expects the hearer to realize who is the referent.

The speaker who uses a defnp may be simply indicating that the referent is knowable. Thus if one says:

(1) Larry read a lot of linguistics in the hospital.
(2) Larry read a lot of linguistics in a hospital.

the (2) use is not the same as the (1). While the hearer does not know which hospital *the hospital* refers to, it is clear it refers to some particular one. In this case, comprehension of the referential term does not involve finding a memory element which represents the real world entity to which the expression refers. Instead comprehension requires deciding that some unique hospital was meant. For general comprehension, this concern is considerable. A hearer may be able to interpret the speaker's purpose in saying (3) (that is, "I'm telling you about John") without determining the real world referents of the noun phrases, but to interpret the request of (4), disambiguation is needed. The disambiguation is problematic because the expression *the Chinese government* can refer to more than one thing.

(3) John got help from the Chinese government in adopting an Oriental child.
(4) Get a visa for your trip from the Chinese government.

Defnps are sometimes used neither to refer, nor to describe specific individuals or objects, but to characterize a class of entities, the properties of which are specified in the head noun phrase and its modifiers. Thus (e1) used in this way does not refer to an individual. It characterizes any member of the class of individuals who are men and book writers. (e1) is similar to attribution except that the description applies to a class.

(e1) the man who writes books

The following classes of defnps may be defined. A defnp that is used to refer uniquely

to one entity, whether or not such an entity exists in the real world, is a specific defnp. A defnp that characterizes a class of entities by means of an individual whose properties are delineated by the properties of the head noun phrase and its modifiers, is a generic defnp. A defnp is attributive if it describes an entity without referring to one. A defnp may be ambiguous in use (u-ambiguous hereafter) if its use as a specific, attributive or generic is not identifiable, while a defnp is ambiguous in reference (r-ambiguous hereafter) if it is used specifically and there is more than one entity fitting the description of the defnp.

### 3.2.3 Context in Defnp Disambiguation

Few utterances, if any at all, are spoken without some surrounding contexts of information. For example, most conversations happen in a location where there are other objects present. Most stories have at least the context of a story teller, a hearer and the story being told. There are contexts with more presumed common knowledge, such as what the hearer knows of the speaker's own identity, or some shared additional information between them (e.g. they have children or parents in common).

Contexts are needed to determine the referent of a defnp. If I say, when I am standing in my kitchen with a friend,

(5) Get me the hot dish holder.

The defnp, if specific, must refer to some unique object in the world. There may be lots of hot dish holders represented by conceptual elements in my friend's mind, but I am referring to a specific holder. Since nothing in (5) distinguishes the one I mean from the whole collection, either I have misused the language, or there is a context which contains only one such hot dish holder, and my friend is aware of that context at the time of my saying (5). In this case, the necessary context is the kitchen, and the referent is probably an item in the kitchen. Reference made to an object external to the conversation is called exophoric reference. It is made relative to a class of contexts which help disambiguate defnps. Contexts of reference which exist in addition to the one created by the linguistic expressions in the discourse are implicit contexts. Exophoric references refer to entities in one or more of these implicit contexts. Thus when I say (5), one of the implicit contexts is the kitchen and its contents since *the hot dish holder* refers to an object in that context.

There are many implicit contexts to which a defnp may be related. Rather than an implicit context consisting of objects near the speaker, the implicit context may be events that the speaker believes are common to the hearer. The speaker who opens a dialogue with (6) below is assuming some previous context (a discussion with the hearer or some other situation) where the reference of *the A.I. Lab Language Group* was first established. In (7), the speaker is again assuming a pre-established referent, but since the hearer may know of several different dogs, some specific context must be chosen that will distinguish a single dog.

(6) The A.I. Lab Language Group wants to meet next week.
(7) The dog is sick again.

Contextual information of yet another kind appears in story telling. At the beginning of a story, the hearer expects characters to be introduced. Sometimes this is done with indefinite noun phrases, which are commonly used in discourse to introduce new items, but often a story-teller uses names or defnps as in the sentences below.

(8) The heiress lived the life of a recluse. She died under mysterious circumstances, but the murderer was never found.

(8) is not a case of cataphoric referencing (referring forward in a text) since the phrase *the heiress* may fully specify an object itself. However, hearers of (8) do not have to search their memories for a referent to *the heiress* in (8). They use the context of story beginning to guide them in deciding that the referent is not knowable but, like *the hospital* in (2) exists uniquely.

The work of Grosz [1977] shows how a focus space may be used to create an implicit context of database elements. When the specification of a defnp is sought, the focus space delimits a part of the database in which to look, so that search of the entire database is avoided. Focus spaces are generally computed from the items mentioned in discourse, given a database in which every item the hearer knows is represented as a database element. The focus space computation may represent other implicit contexts as well. The particular structures needed to represent the kitchen context for (5) are unclear. Instead, I will consider the discourse context further, and in particular, the behavior of anaphoric defnps. As more is learned about how focus moves in it, and the structures it must contain, the choice of representation for other implicit contexts may be more evident.

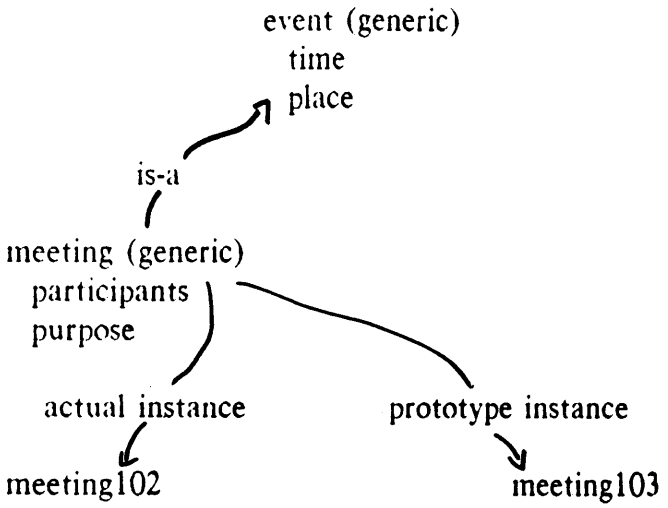
Now let us turn our attention to a particular discourse context. As has been shown, the speaker initially defines it by naming or referring to entities from some already existing context (the hot dish holder), or the speaker creates entities unique to the conversation (some cookies). Once some entities are mentioned, a discourse context is created, which may be modelled as a focus space. Elements from the discourse context may be co-specified by anaphoric expressions; also items in some other context (such as the kitchen) may be mentioned. The anaphoric expressions which are defnps will be the main concern of the next sections; the task is to explain their use. It will also be necessary to show how to determine which defnps are used non-anaphorically, that is, to specify some element in an implicit context other than the discourse context. I will also explore some approaches to finding specifications for non-anaphoric defnps and the computational difficulties which result. Lastly, I will discuss and give rules for a class of defnps which neither refer anaphorically, i.e. co-specify, nor refer to items in some other context. First, I will turn to the co-specifying defnps.

### 3.3 How Anaphoric Defnps Co-specify

Focussing provides a means for determining the co-specification of an anaphoric defnp. In D2, shown above, the initial focus is *some cookies* and its database specification. The defnp *the cookies* co-specifies with the focus. Yet the focussing process is richer than the cookies co-specification suggests; the manner in which *some cookies* is represented is significant to focussing, and the structure of the anaphoric phrase affects co-specification with a focus as well. This section will present a set of rules for the use of focus and illustrate how these rules are used. Rather than present the rules initially, each rule will be shown by illustration and the rule set will be summarized at the end of the section.

To observe the manner in which defnps are related to a focus, a potential focus, or stacked focus, data structure encoding knowledge people have of various relations among objects is needed. A sample knowledge network is shown in figure 1. Suppose that the focus of discussion is on meeting102 in this network. How do speakers talk about meeting102? How do they talk about meeting103? What are the differences in the two? Meeting102 represents a use of a phrase like *a meeting* in "I want to have a meeting with Bill." That is, meeting102 is a specific meeting which the speaker indicates as a focus of discussion. Meeting103, on the other hand, is a prototypic usage.

**figure 3.1. Contents of Sample Knowledge Network**




---

In (9), *a meeting* is used to indicate a typical meeting and (9) says something about meetings as a whole.

(9) A meeting ought to be held early enough in the day that everyone is not too tired to attend.

Meeting102 and meeting103 are both instances of the conceptual element of meeting, an event which has participants and a purpose as well as the time and place characteristic of all events. The network depicted in figure 1 shows the relations which speakers use in discussing meetings.

These observations may be used to state a rule for the use of anaphoric defnps. Rather than state it fully at the start, a simple formuation will be given that will then be "debugged" by example. Once an actual or prototypic instance of a conceptual element is introduced in a conversation, a simple defnp, consisting of the definite article and a head noun, may be used to co-specify that instance. This rule is called the explicit backwards co-specification rule. Common forms of explicit backwards co-specification (EBC) are given in D3 and D4 below.

D3-1 I want to have a big party$_i$ with lots of guests.

2 The party$_i$ ought to be on Saturday so everyone can come.

D4-1 I'm going to tell you about the elephant$_j$.
2 The elephant$_j$ is the largest of the jungle mammals.
3 It$_j$ weighs over 3000 pounds.
4 At one point in its$_j$ existence, the elephant$_j$ had to protect itself from the lion,
5 but now its$_j$ herds are so large, that most lions won't even venture near.

What the reader will notice about D4 is not only the co-specification with focus for the second and third uses of *the elephant*, but also that the co-specification implies these uses are generic. Whereas defnps in isolation are often ambiguous on the generic-specific classification, this occurs less often in discourse context since the focus can provide the generic/specific distinction for the defnp. As stated, the EBC rule makes a prediction about defnps, u-ambiguous in isolation, which occur in sentences following the focus; they co-specify with the focus, and hence are disambiguated as generic or specific. D5-2 below contains a defnp which is u-ambiguous in isolation, but in the discourse context, since it co-specifies with George's elephant (the focus), it is specific.

D5-1 I sent George an elephant last year for a birthday present.
2 The elephant likes potatoes for breakfast.

When a defnp, occurring after a focus of the same head noun, is not u-ambiguous in isolation, it does not necessarily co-specify with the focus. Instead, the sentences may be related only because they mention the same general class of elements. D6 illustrates the problem. D6-2 is a generic sentence in isolation, and the defnp is generic. Even in the context of D6-2, where the focus is Mary's ferret, hearers interpret the underlined defnp as generic.

D6-1 Mary got a ferret for Christmas last year.
2 The ferret is a very rare animal.

The discourse context does not override a strongly generic reading of a defnp. In order for this to be so, sentential level processing must have occurred without consideration of the demands of the context. Since the EBC rule as stated predicts co-specification in cases like D6, it must be revised: specific and u-ambiguous defnps which contain the same noun phrase head as the focus, and which occur after the focus in the discourse, co-specify with the focus.

Once a speaker is focussed on a particular element of the knowledge network, there are restrictions on how that element may be co-specified with. Defnps used to

co-specify cannot contain more description than is known about the focussed element. If the time of meeting102 in figure 1 has not been established, the speaker may not use a defnp like *the meeting at 3 o'clock* to co-specify with meeting102.[1] In D7 below, D7-2 seems odd in the discourse to English speaker-hearers because it is not clear what *the fairy tale book* has to do with the first sentence.

D7-1 I bought a book today.
2 The fairy tale book is by the Brothers Grimm.
3 It is really well illustrated.

One could say following D7-1, "The book I bought is a fairy tale book by the Brothers Grimm" (since D7-1 states that the speaker bought the book), but one cannot say D7-2. Why can't a defnp that contains more information than the focus co-specify with the focus? The discussion of focus movement in chapter 2 shows that an anaphoric expression following the focus either co-specifies with the focus or introduces an element which is the potential new focus of the discourse. The difficulty with phrases like *the fairy tale book* is that one cannot tell if it is intended to co-specify, or because it is somewhat different from the focus, is intended to be used as a potential new focus. The EBC rule may be reformulated:

> The Explicit Backwards Co-specification rule: specific and u-ambiguous defnps which contain the same noun phrase head as the focus, which follow the focus in the discourse, and which do not contain more information than is included in the focus, co-specify with the focus.

The EBC rule explains why a negative existential cannot be referred to using a defnp. A sample case, from Karttunen [1968], is given in D8. D8-2 is generally regarded as an unacceptable sentence following D8-1. The sentence is certainly grammatical, so the assumption by Karttunen is that the referential term *the car* is being used in some inappropriate manner.

D8-1 I don't have a car.
2 * The car is black.

Sentences such as D8-1 without the negation are ambiguous between a specific and non-specific reading. D8-1 is ambiguous as well. It says "it's not the case that I own

---

1. Vendler [1967] pointed out that over-specified noun phrases do not refer back.

any car," or "I own something which is not a car." There are two possible entities which could be described by this sentence: any car which the speaker does not own, and something which is not a car. If the focus of D8-1 is any car that the speaker does not own (which seems an unlikely reading although possibly the speaker's intention), then the EBC rule predicts that *the car* co-specifies with the focus. But the focus is the collection of things which speaker does not own, so it is odd to talk about its color. If the focus of D8-1 is anything which is not a car, the defnp cannot be used since the description suggests it is a car. Thus for either reading D8-1 is odd. This example points to an issue which will be discussed further in chapter 4: scope of quantification for an element in focus is often decided by the sentences which follow in the discourse.

A similar case, also from Karttunen, does not involve negative existentials, but entities within modal contexts:

(10) * Mary expected a present from John although the present was expensive.

(11) Mary expected a present from John although the present wasn't the thing that worried her.

The defnp in (10) according to the EBC rule must co-specify with the focus. What is significant is that the co-specification is acceptable, as (11) shows. What is odd about the second clause of (10) is the predication. This chapter cannot give an account of such semantics, but intuitively, it seems odd to predicate a property such as expense of an entity which exists only in some hypothetical situation such as expectation. Thus as long as there is a co-specification specified by the focus, a defnp may be used, but the predication about the defnp must be semantically meaningful.

One of the ways in which a speaker may talk about a focus is to use a defnp which names the focus in terms of more general elements in the network, that is, by terms which may be reached through the is-a links in figure 1. This use will be called lexical generalization of the focus; Grosz [1977] categorized it in her description of defnps. In D9, *the poor old beast* is a lexical generalization of *the dog*, that is, its head noun is a term which is a class generalization of the focus.[1] Determining the class

---

1. This term comes from the observation of Halliday and Hasan [1976] that lexical cohesion includes the use of reiteration of four types: same word; synonym, superordinate, and general word.

generalization of the focus is possible when the focus is represented in the way that is assumed in this report: as an association network with an is-a hierarchical structure. Using that hierarchy, it is possible to determine whether a phrase such as *beast* is hierarchically related to *Malamut*.

D9-1 Harold took his Malamut$_i$ to the vet yesterday.
2 The poor old beast$_i$ was quite lame.

One might expect that some constraint on the amount of information in the lexical generalization of the focus is needed. This is the case, since the underlined defnp in D10-2 (a) is unacceptable[1] following D10-1 as a co-specification with the focus while D10-2 (b) is fine.

D10-1 Harold took his Malamut$_i$ to the vet yesterday.
2 (a) * The beast who is old was quite lame.
(b) The mangy, snarling, unfriendly beast was quite lame.

In general, any restrictive relative post-nominal modifiers on a noun which are a lexical generalization of the focus force the defnp to fail to co-specify with the focus. Pre-nominal modifiers, regardless of their complexity, preserve co-specification. They may be used to provide new information, but unlike D7, the information is not marked as already known to the hearer. Hence lexical generalization seems to follow the same pattern as the simple form of EBC: defnps that co-specify with the focus cannot contain any more information than is associated with the focus. To summarize, a second EBC rule for lexical generalization is:

> The EBC Lexical Generalization Rule: A defnp which lexically generalizes the focus co-specifies with the focus, as long as it does not contain restrictive post-nominal modifiers.

The two EBC rules presented here may be used to judge co-specification of a defnp with the focus, and also with a potential focus or stacked focus. A defnp in question must be compared with the focus using the EBC rule first. If the focus fails to

---

1. The reader is cautioned about reading D10-2 (a); the non-restrictive form of the defnp (read, "the beast" (pause) "who is old") is entirely acceptable. The rule for restrictive relatives stated here is so strong that readers tend to give 2 (a) the non-restrictive reading and miss the unacceptablility being discussed.

be applicable, the rule may be applied using potential foci and stacked foci viewed in terms of the network hierarchy as well.

In Grosz' [1977] work, several kinds of defnp resolution are considered. Grosz points out that more general descriptions than a specific item in focus may be used to refer to the focussed item, e.g., if the focus includes "novel", then reference to "the book" is a reference to novel. A second class of general decriptions are references involving selection from a set when the set is in focus, e.g., if a high school class is in focus, reference to "the brightest student" is a selection from the set. It is the hierarchic net which allows us to explain Grosz' observations, to describe general descriptions more fully, and to specify the limits of their use.

### 3.4 Implicit Backwards Specification

Many definite noun phrases which occur in discourse are not cases of backwards co-specification with the focus. Such defnps are related to the focus in one of several ways which may be characterized using the hierarchical network. Since the focus is defined in terms of the network, these relationships can be easily determined. The focus acts as an anchor point for finding specifications for such defnps; the general class of such specifications, I will call implicit backwards specification. Such cases are to be distinguished from explicit backwards co,-specification because the defnp does not co-specify with the focus itself; the defnp can be said to specify an element that is closely associated with the focus. From the viewpoint of definite anaphora comprehension, these cases are different because they depend on the focus for associations rather than direct co-specification. Importantly, implicit specification accounts for one of the ways in which discourses are distinguished from unconnected text.

There are four kinds of implicit specifications: associated, inferential, set-element, and computed. Examples of each will be given below. A defnp may be related in one of these four ways to either the focus, or if the focus has no such implicit specification, to one of the potential foci. To determine the specification of such a defnp, the defnp may be compared to the focus for one of the focus relations, or that failing, to one of the potential foci. Interestingly, the stacked foci do not seem to be used in this way. Perhaps because there is additional processing time associated with

these judgments, it is not possible to extend the judgments to the focus stack.

In D11 below, the defnp *the time* specifies the time of the discourse focus, the meeting. D11 is an example of associated specification; it allows the speaker to talk about an object closely associated with the focus without mentioning the focus explicitly. Glancing back to figure 1, the reader will observe that the associations of meeting include participants and purpose, and by inheritance from higher nodes in the net, place and time. The phenomenon of association between two noun phrases has been cited by Norman and Rumelhart [1975]. Grosz [1977] suggested that association is a means of bringing other items implicitly into focus.

D11-1 The pa group wants to have a meeting.
2 The time will be 3 p.m. on Tuesday.

Implicit backwards specification is constrained by the association network surrounding the focus. Any element closely associated with the element which is represented by the focus can be mentioned using a simple defnp. Thus in D12, sentences with acceptable defnps, as well as ones with unacceptable defnps, are given.

D12-1 I went to a new restaurant with Sam.
2 The waitress was nasty.
3 The food was great.
4 The soup was salty, but the wine was good.
5 * The rug was ugly.
6 I like the band that plays there.
7 ? The elephant with the green tutu danced an impressive jig.

Simple defnps signal a close association between themselves and the focus. The simple defnp may be used because the hearer has a pre-established link between the two representations. In figure 1, the associations for meeting102 are represented in two places, those such as *participant* are linked with generic meetings, while those such as *time* are linked with events and "inherited" to other elements lower in the tree. Non-simple defnps have infinitely more variety because the modifiers can specify the relation of the defnp to the focus at hand as with *the band that plays there* in D12-6. Non-simple defnps which do not state some connection are less acceptable. D12-7 seems odd, but hearers will attempt a connection by assuming that the defnp is somehow related to the focus; that is, that the elephant is part of the floor show at the restaurant. This is an example of Grice's perspecuity maxim which was assumed in

chapter 1.

Another use of focus is as an inference point for inferred specification. Inferred specifications, such as *the murderer* in (8), presented here as D13, are not mentioned explicitly in the previous discourse, nor can they be considered on general principles as closely associated with the focus. Their use reflects an inference about the focus on the part of the speaker.

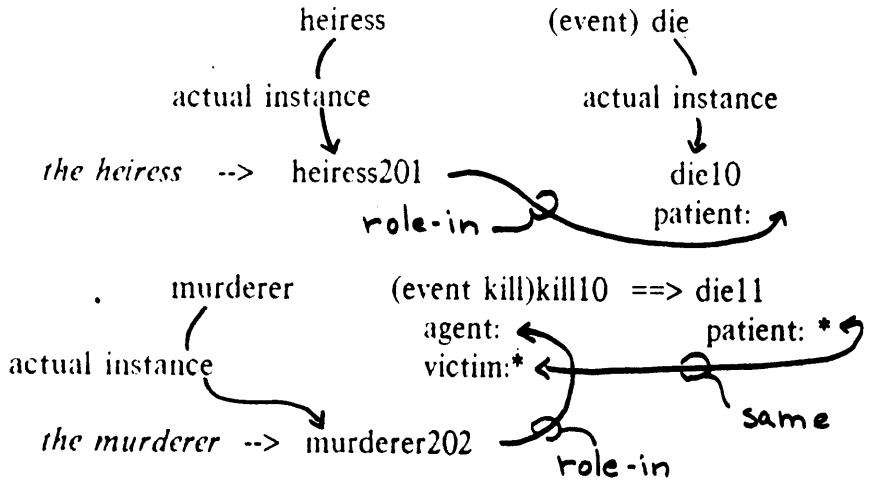
D13-1 The heiress lived the life of a recluse.
2 She died under mysterious circumstances,
3 but the murderer was never found.

In D13, *the murderer* represents an inference that the heiress' death was due to a specific type of circumstance, a murder. I consider the description of exactly how such inferences occur to be beyond the scope of this work. However, given a net of the type described in Fahlman [1977], with two inference points like heiress and murderer (and the information associated with heiress from the context thus far), the relation of the murderer to the heiress may be inferred. Such an inference does not produce a representation of a real-world entity to which *the murderer* refers. Instead, the inferred relation of murderer and heiress provides sufficient information to produce the representation if it exists. When a representation does not exist in the database, the inference between the murderer and the heiress suggests that the speaker is attributing of some individual that s/he is a murderer.

Figure 2 depicts the network for D13. The inferences represented in the net are those linked by pointers to the representation of *the murderer*. The inferences from the representation of murderer202 include that murderer202 was the agent in some killing, which is represented as kill10 (with parent nodes of kill and event); furthermore, a killing results in the death of the victim of the killing. In the net, the identity of the victim of kill10 and patient of die11 is represented by a star. In existing AI knowledge representation languages such as FRL, the identity can be guaranteed by the use of procedures which forward the value associated with the victim of kill10 to the value of the actor in die11, and the other inferences can be done "automatically." This net represents the same relations which one would use for the representation of complex relations in Fahlman's NETL.

A concrete example will explain how attribution may occur in an example such

**figure 3.2. Representation of network for discourse D13**




---

as D13. Suppose the hearer knows that the heiress was killed by Jones. Then on hearing D13, the hearer not only concludes that *the murderer* refers to the murderer of the heiress, but also that Jones is denoted by the referring expression. However, another hearer upon hearing D13 and not knowing what the first hearer knew, could only conclude that *murderer* is attributed of a person who is assumed to have murdered the heiress. The denotation is not known to the second hearer, but if someone were to tell him/her that Jones murdered the heiress, the hearer could conclude to whom *the murderer* refers. In effect, the defnp used in this way stresses the attributional use of expressions which Donnellen [1977] has observed. The argument presented here is not only concerned with the nature of focus; it is a statement of what information is sufficient to make up a description which may denote a unique entity. Viewing inferred specification defnps as attribution has an implication for a computational model which disambiguates such defnps. This model must be able to use an expression without knowing its specification; if some knowledge makes that specification available at a later point, the model must be able to link the specification to the referring expression.

At first glance associated and inferred specification appear to be one in the same thing. The discourse below, from Karttunen [1968], will indicate just how the two differ.

D14-1 I was driving on the freeway the other day.
2 Suddenly the engine began to make a funny noise.
3 I stopped the car,
4 When I opened the hood, I saw that the radiator was boiling over.

With the focus of *driving* in D14-1, the specification of the engine may be found, since vehicles are driven on freeways and vehicles have engines. The association chain here suggests that the connection between D14-1 and *the engine* involves a few inferences. These inferences are part of a hearer's general knowledge and are generally true; they are part of the knowledge in the association network. With D13, however, the inference about the murderer involves a supposition which is not generally true since dying under mysterious circumstances does not always imply murder.

The distinction between associated and inferred specification may thus be stated: associated specification involves common sense inferences which are true about the world, while inferred specification involves a supposition which the speaker has made which is not necessarily true. This distinction has computational significance. In either case, the implicit specification says, "looking at this focus and this defnp represented as an element of the database, find a semantic connection between these two items." In the case of D14, the connection results from the following: people drive vehicles on freeways, engines are part of vehicles, so *the engine* is the engine in the vehicle the speaker was driving. For D13, the inferences must include that the hieress has died, that something or someone caused her death, and that a murderer is an agent who causes a death. If a supposition is made that the something causing the hieress' death was a person, then *the murderer* is that person. In figure 2, the supposition amounts to deciding that die10 is the same event as die11. Supposition is essentially an abductive inference, defined by Pierce (see Goudge [1969]) and used in artificial intelligence work in medical diagnosis (see Pople *et al* [1977]) and in language related tasks by Bullwinkle [1975]. Abduction is not yet a general technique in inference programs, although these examples suggest that it will be needed in order to understand certain anaphoric defnp uses.

To summarize, the associated and inferred specificiation rules may be stated:

> Associated Specification: A defnp will be judged as a specification of an element associated with the focus if the defnp names an element associated with the focus directly, or via inheritance on the network hierarchy. The association may involve some common sense deductions about the defnp.
>
> Inferred Specification: If the defnp can be associated with the focus by supposition, which may involve common sense deductions as well, the defnp specifies an element associated with the focus by means of the supposition.

These two "rules" are more informal than the others stated thus far. Given a hierarchical net, direct association and inheritance may be computed; that part of the rules involves computational machinery which is well defined. However, the inference uses are only informally stated. Common sense deductions have been studied extensively in the artificial intelligence literature, but supposition much less. An inference machine for performing these inferences will not be given here. It has been sketched out in the discussion above, but its exact content remains to be discovered.

Another kind of implicit focus relation is illustrated by D15. I call this relation the set-element specification since *the clown with the unicycle* is an element of the set of clowns which the focus specifies.

D15-1 I went downtown today,
2 and there were clowns performing in the square.
3 The clown with the unicycle did this really fantastic stunt.

Set-element specification may be stated in the following rule:

> When the focus is a set, if the defnp is a singular containing the same defnp head as the focus, and additional modifiers, then the defnp specifies one member of the set specified in the focus; the representation associated with the defnp will be marked as an element of the focus. When the defnp lacks additional modifiers, the specification is odd.

As with inferred specifications, the focus does not make it possible to choose the specification of the referring expression. Instead the focus points at a representation of a set, of which the specification of that phrase is a member. These cases are easier to distinguish than other specifications because the head noun is the singular of the noun phrase represented in the focus. Unlike defnps using the EBC rule, set-element

specification demands a modifier that distinguishes it from the focus. Without the modifier, there is no means of determining which member of the set is being discussed. Grosz [1977] says of cases such as the set-element relation that an inference is needed to establish additional properties of an object in focus; the set-element specification rule encodes this inference.

Another kind of focus relation, which I call computed specification, is shown in D16. Here *the last meeting* does not refer to the meeting mentioned in the previous sentence, but that meeting may be used as a point for determining a last meeting, if one is represented in the database. Otherwise, it is a description of the element required, as with inferred specifications and set-element specifications. Several modifiers - *first, last, next, second* and the other ordinals - are used in this way.

D16-1 Aunt Het's Sewing Bee wants to have a meeting this week.
2 The meeting should be on Tuesday.
3 The last meeting, which was at 5, was too late, so schedule this one earlier.

The computed specification rule is difficult to state without assuming a semantics for ordinal terms which indicates how the ordinal is used to choose the associated term. The semantics of an ordinal term will be called an ordinal sequence computation. For example, *last* computes the last relation either relative to a whole group, taking the last one in the group, or relative to a member of the sequence, taking the last one relative to that member. If the sequence were meetings sequenced by time, *the last meeting* points to either the last meeting in the whole sequence, or the meeting which occurred just previously to the member of the sequence which is in focus. The computed specification rule is:

> If a defnp contains an ordinal modifier, the same head as the focus, and no relative clause modifiers, it specifies an element which may be chosen in the database by the ordinal sequence computation relative to the focus; defnps containing full relatives use the relative rather than the focus for the ordinal sequence computation.

By these examples, the nature of focus in discourse can be re-emphasized. It is the focus which connects sentences of the discourse. In the process of determining the focus relation between a defnp and the focus, the link in the discourse is created. The

form of a defnp which follows a given focus must stand in one of these relations to the focus. When it does not, two situations occur. In the case of set-element or computed specifications, a simple defnp with a different noun phrase head is an odd use, and for set-element specification, no modifier is also odd. This prediction is accurate, as D17 shows.

D17-1 I went downtown today,
2 and there were clowns performing in the square.
3 * I saw the chair.

If the focus of D17-2 is the clowns, then the chair is odd on set-element grounds.

Using the maxim of perspicuity, a defnp which fails to be related by the rules above may be judged to be related as an inferred or associated specification. This judement requires some reasoning about speaker beliefs. Specifically, the hearer must determine that the defnp is related by association or inference, by making a supposition of the form "the speaker wants me to believe these things are related." If the hearer finds no support for that supposition, the defnp will again be odd. What constitutes support? In D13, support was given by the information that the heiress died under mysterious circumstances. In D12, no support was evidenced in the text for assuming the restaurant had a dancing elephant, but by assuming dancing elephants could be the entertainment provided by a restaurant, a reasonable supposition may be found for linking the elephant to the restaurant. Suppositions seem to work if the hearer is required to suppose only one fact, and that fact is related to some facts stated in the text. When one must suppose a fact without some explicit reason for doing so, the defnp becomes odd. Thus for D17, on association criteria, a chair is difficult to associate with a bunch of clowns without any further information, and the simple defnp does not lend itself to supposing only one fact. Hence the odd usage.

To summarize, the closure rule for implicit specifications is:

> When the focus is a set, a simple singular defnp with no modifier is an odd implicit specification. When the defnp has ordinal modifiers and a different noun phrase head than the focus, it is an odd implicit specification. Other defnps which are not related by direct or inherited association must be judged related. If they cannot be, the defnps with no modifiers are odd uses, while those which have modifiers specify outside the discourse context.

What may be said about a defnp which specifies "outside the discourse context?" Consider a defnp which does not bear any relation to the focus such as *the chairman of the math department* in D18 below.

D18-1 George wants to have a seminar to discuss representation in frame-like languages.
2 He wants to invite the chairman of the math department.

The focus in D18 is the seminar of D18-1. *The chairman of the math department* does not co-specify with any noun phrase in the discourse, nor does it specify any element associated with the discourse. Instead it specifies some element which was not previously mentioned, that is, an element outside the context so far presented. The math department chairman is not directly related to the focus of seminar, nor is it related by association since math department chairmen are not associated with the general notion of seminars. However, the focus does provide an important piece of information; it is the source of the elided event to which the chairman is invited. Hence D18 is different from D12-5 or D17 where there is no link between the focus and the sentence in question. What may be concluded is that the focus is not useful in determining the specification of this defnp. The sentence is not odd because of focus links; yet something besides the focus supplies the specification. This conclusion points to the limitations of the focus: it captures only those specification and co-specification relations which are internal to the discourse. In a sense, the focus is a partial history of the discourse context and what the hearer knows about it.

How is the specification of *the chairman of the math department* found? Since the specification of *the chairman of the math department* lies outside the discourse context, it is like the exophoric reference of hot dish holder discussed in an earlier section. A more global context such as that of the speaker's situation in time and space must be used to determine a context of possible specifications. This context must be limited because there are potentially many math department chairmen in the speaker's and hearer's memories. However, global focussing, discussed by Grosz [1977, 1978], provides the basic data structures which may be used in conjuction with the discourse focus. I do not intend to describe just what such a context might look like, but I do want to indicate that it may be "grown" from a search through the associative network, to other elements which are related to any of the objects in the discourse, including the speaker and hearer. The association net includes not only abstract representations of general classes of real world entities, but also representations of real world objects. The

associations between real world objects may be gathered by a search method which collects associations close to the elements in global focus and then extends for other associations until one is found that matches the defnp in question. Context growing is essentially a Quillian [1968] net operation; this operation is made possible by the presence of the focus. It also suffers from the combinatorial explosion of relations which is inherent in the Quillian technique.

Focus makes context growing possible in the following way. Suppose the hearer is trying to determine the specification of some highly ambiguous name like *John* who is to be invited to a meeting which certain other people will attend. The global focus includes the meeting, and these people. From the specifications of these elements, context growing produces more specifications because the associations of the participants include who their collegues are, who normally attends meetings and similar information; *John* will specify whichever representation is found first. When several are found in the same expansion, the network will reflect a possible ambiguity in the speaker's discourse. Of course, it is possible that no specification will be found. Such a circumstance is yet another example of the hearer knowing that a defnp refers without being able to tell who the speaker had intended as the reference. The implications for computational models regarding defnps which specify outside the discourse context are clear. Specfically, the models must express what the speaker believes the hearer knows about, so that the speaker does not produce referring expressions which the hearer cannot disambiguate, and the models must express what the hearer believes the speaker knows about so that the hearer can decide what to do with referring expressions which are ambiguous.

In general, determining the specification of a noun phrase from among the elements of a database must be described with a complex process, and the above method of context growing is inadequate. More difficult to interpret than *John* are noun phrases such as *the oldest structure in the world.* Suppose that it is intended to have the great wall of China as its network specification. Finding the specification requires a means for finding the information that the great wall of China is indeed the oldest structure in the world (if that information is available somewhere in the database). Presumably what is required is a semantics which provides the meaning of each of the words in the string, as well as their composite meaning. This task has perplexed

linguists and natural language researchers for a long time. No solution is offered here. Nor is it clear how to represent or reason about such problems in AI languages which are used to describe language use and access databases of the kind assumed here. Creators of various frame languages (Minsky [1975]), especially FRL and KRL (see Roberts and Goldstein [1977] and Bobrow and Winograd [1977]) have addressed themselves to the importance of a language in which is it possible to express the semantics of English. Martin [in preparation] provides a new framework for interpreting meaning in English.

## 3.5 An Algorithm for Interpreting Defnps in Discourse

Now that a set of rules for defnps have been described, these rules may be summarized and an algorithm indicating their use may be given. Of the rules discussed in the previous sections, the closure rule is not given below. Instead it serves as a step in the algorithm which follows:

> The Explicit Backwards Co-specification rule: specific and u-ambiguous defnps which contain the same noun phrase head as the focus, which follow the focus in the discourse, and which do not contain more information than is included in the focus, co-specify with the focus.
>
> The Explicit Backwards Co-specification Lexical Generalization Rule: A defnp which lexically generalizes the focus co-specifies with the focus, as long as it does not contain restrictive post-nominal modifiers.
>
> The Associated Specification Rule: A defnp will be judged as a specification of an element associated with the focus if the defnp names an element associated with the focus directly, or via inheritance on the network hierarchy. The association may involve some common sense deductions about the defnp.
>
> The Inferred Specification Rule: If the defnp may be associated with the focus by supposition, which may involve common sense deductions as well, the defnp specifies an element associated with the focus by means of the supposition.

The Set-Element Specification Rule: When the focus is a set, if a defnp is a singular containing the same defnp head as the focus, and additional modifiers, then the defnp specifies one member of the set specified in the focus; the representation associated with the defnp is marked as an element of the focus. When the defnp lacks additional modifiers, the specification is odd.

The Computed Specification Rule is: if a defnp contains an ordinal modifier, the same head as the focus, and no relative clause modifiers, it specifies an element which is chosen from the database by the ordinal sequence computation relative to the focus; defnps containing full relatives use the relative rather than the focus for the ordinal sequence computation.

The first two of these rules will be referred to as the Explicit Backwards Co-specification rules while the latter four are the Implicit Backwards Specification rules.

The algorithm below determines the function of definite noun phrases which occur in discourse. For those which are used anaphorically, the co-specifications are determined in the function of the algorithm. For those which specify by association, the kind of association is determined. This algorithm relies on a data structure like that shown in figure 1, that is, one which encodes associations of a database element, provides for links expressing the general class of the element and provides for inheritance of associations from more general class elements.

An algorithm for determining defnp function in discourse:

1) (Focus Relation) Given a defnp, it co-specifies with the focus if either of the explicit backwards co-specification rules hold. Otherwise the implicit backwards specification rules (except inferred specification) may be applied to determine its specification.

2) (Potential Foci Relation) If no rules apply, the explicit backwards co-specification and implicit backwards specification rules (except inferred specification) may be re-applied to the defnp using each of the available potential foci in place of the focus in the rules.

3) (Stacked Foci Relation) If no rules apply, the explicit backwards co-specification rules may be re-applied for the available stacked foci in place of the focus in the rules.

4) (Closure and Outside Specification) If no rules apply, for a defnp without modifiers, the inferred specification rule may be applied; if it fails to hold, the use is odd. For a defnp with ordinal modifiers, the use is odd while for other defnps containing modifiers, the defnp specifies outside the discourse.

5) (Focus Set Use) When no focus has been established, a defnp which lexically generalizes one of the focus sets co-specifies with that set.

Steps 1 and 4 of the algorithm have been exhibited in the examples presented in the last several sections. The need for step 2 is indicated by the example below.

D19-1 We are going to have a staff meeting tomorrow.
2 We can have it in my office.
3 There are blackboards on the walls, and lots of tables.
4 I think the room will be quite adequate.

While the focus of D19-2 is staff meeting, *the walls* cannot be associated with it so that step 1 of the algorithm does not apply. Step 2 however does; in D19-3, *the walls* is an implicit backwards specification by association to the potential focus of *my office*. D19-3 should also move the focus to the office. As stated in chapter 2, the focussing algorithm moves focus only by direct co-specification. However, the backwards implicit specification rules give evidence for changing the relation between a defnp and focus to include any of the implicit specification relations.

When the focus is moved from staff meeting to office in D19, the staff meeting is stacked in the discourse focus stack. Bearing this process behavior in mind, suppose D19 were to continue as follows.

D19-5 The meeting can be held anytime after 3 pm.
6 if you need anything for it,
7 tell my secretary,
8 and he'll arrange it.

*The meeting* co-specifies with the stacked focus of staff meeting. This co-specification is found using step 3 of the algorithm above. Step 3 does not allow for implicit backwards specfication rules because D19 would be odd with the following continuation instead of D19-5.

(12) I think we can set the time to follow the mid-morning coffee break.

*The time* specifies the time of the stacked staff meeting. While most hearers are able to interpret this defnp correctly, the disjointedness of the discourse indicates that such a defnp use is not very acceptable, perhaps because it requires extra processing effort on the part of the hearer. Because of the marginal acceptability .of such defnps, the implicit backwards specification rules are not used in step 3.

The algorithm above does not predict correctly the use of certain generic defnps, including those discussed as part of the explicit backwards co-specification rules. Generic defnps may be incorporated in the rules and algorithm above, but to do so, the nature of generic noun phrases must be better understood. The following simple question clarifies the difficulty with generics. How are generic noun phrases, definite or indefinite, distinguished from non-generics? This question has plagued language researchers for a long time and remains unanswered. The next section will present some rules distinguishing generics in some instances. The rules are descriptive; they do not exist within a theoretical framework which explains why they work and how consistent they are. Nor are the rules presented in terms of a computational process. The purpose of these rules is to indicate that generic noun phrases may be distinguished in some cases and are ambiguous in others. However, their role in a computational model will also be considered. Based on the rules for generics, the function of generics relative to focussing will be discussed.

## 3.6 Distinguishing Generics

### 3.6.1 Basic Rules

Attributive and specific defnps both indicate a relationship to individuals, while generics indicate a relation to classes. One would like to distinguish these two uses in discourse. When looking at single sentences as in (13) and (14), containing an expression such as *the television*, it is reasonable to ask: is this defnp a generic or not?

(13) The television broke.

(14) The television has extinguished reading as a childhood pastime.

Hearers are able to distinguish generics in sentence pairs such as those above. In the worst case, however, the generic usage is not distinguishable from the specific/attributive without the assistance of a context. Yet generics do occur in the first sentence of a

discourse where there is no context that will help to discriminate, so I will consider rules which may be used to determine a generic usage in isolation. Generics are considered as anaphoric phenomena because distinguishing a generic is significant for the interpretation of defnps: the relation between a focus and a defnp varies, depending on whether the focus is generic. In presenting these rules, I will also show why, in the worst case, the generic cannot be distinguished on the basis of a single sentence. The rules presented here for generics are a purely descriptive account of the generic defnp. An explanatory theory of generics will provide an account of some of the observations made below. In the discussion that follows, I will proceed from consideration of simple cases to those that are more complex.

There are two sentential characteristics which reliably mark defnps in a sentence as generic.

> **Rule 1**: Sentences containing *is-a* verbs with defnps as subjects mark generics when no specific defnps occur elsewhere in the sentence.

> **Rule 2**: Sentences that are announcements are generic unless the defnp is one of the default cases which will be discussed below.

(15) is an instance of rule 1 while (17) is an instance of rule 2. (16) indicates why rule one must take into account defnps which are specific.

(15) The elephant is a large mammal.
(16) The elephant is a big hit with the kids.
(17) I want to tell you about the orangutan.

If *the kids* is taken to refer to some set of kids known to the speaker and hearer, then *the elephant* in (16) cannot be used generically if the speaker wishes to utter a meaningful sentence. As will be indicated below, (16) is ambiguous because *the kids* may also be taken as a generic. Rule 1 excludes sentences like (16) because they are not unambiguously generic. Other indicators of generic sentences, such as tense and periodic time phrases, are not reliable by themselves. Smith [1969] presents several examples of sentences with defnps with a variety of tenses that are ambiguously generic or non-generic.

**Rule 3:** Defnps such as *the sun*, or *the moon* make up a class of default specifics.

These defnps always specify a specific element, when no discourse precedes. In a database, these cases are governed by a default mechanism which marks their default specification accordingly. Other specifics such as *the President* or *the constitution* require some assumptions about the cultural heritage of the speaker and hearer. They are also specific, but the specification requires a means of indicating some non-linguisitic context.

The rules presented so far must necessarily depend on the criterion of applying to sentences which occur initially in a discourse. None of the defnps has the same behavior when a discourse context or an external context preceeds it. Thus if the speaker is standing next to an orangutan, (17) could be used to refer to the present orangutan (though, a deictic *this* is preferable because it marks the speaker's pointing behavior). Similarly, if the speaker has begun an essay about the moons of Jupiter, the expression, *the moon* might refer to one of the Jupiter moons.

For the remaining rules, which deal with various defnp uses, defnps will be classified according to their syntactic structure.

**Rule 4:** The following two cases of simple[1] defnps are u-ambiguous:
(a) Simple defnps in sentences whic only contain other simple defnps.
(b) Simple defnps in sentences with indefinites or possessive adjective defnps.

(18), like (16), exemplifies the ambiguity of 4a, while (19) shows the ambiguity of 4b.

(18) The robot is replacing the car.
(19) Steve's go-cart ran well against the motorcycle.

When simple defnps occur in sentences which contain referring expressions, the relationship between the referring expression and the defnp helps to determine whether the defnp is generic. If there is a relationship, then one must consider the set of memory elements described by that relationship. In (20) the relationship of dissident

---

1. A simple defnp is a noun phrase consisting of just one noun and a definite article.

and the Rhodesian government is at least one of existence (there exist dissidents to the Rhodesian government), and the set of dissidents is large.

(20) The Rhodesian government prevents the dissident from leading a normal life.
(21) The Supreme Court bases its rulings on the constitution.
(22) The Chinese worship Chairman Mao.

**Rule 5a:** Whenever the set of memory elements described by the relation between a simple defnp and a referring expression is larger than 3 elements, a generic reading is assumed.

**Rule 5b:** In cases where only 1 or possibly 2 members comprise the set, the reading is specific, and in the case of two members, r-ambiguous.

(21) contains an example of a specific, unambiguous use of the defnp, *the constitution*, while in (22), *the Chinese* is r-ambiguous.

When a defnp appears to be generic, as predicted by the above rule, its generic behavior sometimes results in sentences which are odd to English speaker-hearers. A generic defnp which occurs in a predicate argument position of a verb which cannot accept generic terms causes this behavior. In (23), *the dissident* is still generic, but the resulting sentence has an odd reading as a generic.

(23) The Rhodesian government has caught the dissident.

The precise verb classes which consistently reject the generic are unclear although many physical actions do not accept generic objects.

**Rule 6:** A verb which does not accept a generic in a given predicate argument position will force u-ambiguous defnps to be specific, while those which are generic will result in odd sentence readings.

Adjectives attached to defnps act either to indicate an object with a certain property or to indicate that the entire noun phrase describes a class. For some adjectives, only the property relation seems to be defined, e.g., *old* as in *the old man* or *the old dog*. In fact, the adjective alone can rarely be used to distinguish. Thus *black* may establish a property as in *the black hat*, but it may create a class description as in *the black man*. A more productive means of considering adjectival defnps is to ask

whether the adjective and the noun phrase together suggest a class as in (e2) or an individual element as in (e3).

(e2) the American household
(e3) the current American government

To answer this question, a knowledge base with the following information is needed. First, class categories for some adjective-noun phrase pairs must be given explicitly. Color words such as *black, red, yellow* and *white* applied to *man* mark a class description while the other color words in general are just property indicators. This observation has been important in the specializer-generalizer categorization used by Martin [Hawkinson, 1975] in the OWL representation language. A second type of information, especially important for adjectives that specify classes and property features, is whether the defnp distinguishes a single individual as in (e3) or specifies a group as in (e2).

The behavior suggested above may be summarized in the following questions about an adjective-noun phrase pair:

1.) Is there only one database specification of the phrase? Example: e3
2.) Is there a class use acceptable for this pair? Example: e2
3.) Are there many individuals described but none outstanding? Example: *the brown dog*

Rule 7 indicates the purpose of these questions. Examples of each case are given below.

**Rule 7:** A *yes* answer to 1) above suggests a specific reading, while a *no* answer to 1) and a *yes* to 2) suggests a generic reading. A *no* answer to 1) and 2) and a *yes* to 3) indicates r-ambiguity while a *no* answer to all three is impossible.

(24) The current American government is moving towards a socialistic economy.
(25) The black man has come to a new view of society.
(26) The brown dog chases rabbits everyday.

In addition to these questions, the sentence in which the defnp appears provides a context for interpretation; rule 6 applies to noun phrases containing adjectives. As a result, in (27) the generic reading must be rejected because the predicate-argument relations demand a non-generic subject.

(27) The black man was moving towards the window.

For defnps with restricted relatives, two characteristics cause these defnps to appear generic: simple present tense marking (as in e4a-c) and periodic time phrases (as in e4d-e). Simple past tense causes the defnp to appear non-generic (as in e4d-e without the time phrases).

(e4a) the man who reads *The Grapes of Wrath*
(e4b) the lion that doesn't have enough to eat
(e4c) the child who goes to bed early
(e4d) the professor who wrote literary work in the 19th century.
(e4e) the man who practiced medicine in the Middle Ages

However, these cases are only suggestive because two other factors must be considered. As with other defnps, the hearer's general knowledge may suggest one or two individuals who fit the description provided by the defnp. Thus questions 1) and 3) used for adjectival defnps are useful for restricted relative defnps.

A problem unique to restricted relative defnps centers on a way to formulate question 2) above. One cannot simply ask if there is a class specified by the noun phrase plus restricted relative. The answer could require that each description be stored in memory; an infinite number of descriptions would result without any explanation for how new ones are created or the old ones learned. Instead, the formulation of generics for restricted relative defnps must distinguish the uniqueness of the noun phrase. Uniqueness results from a semantics rich enough to define a set of properties which may be used to specify elements of the database; when only one element is so specified, the noun phrase is a unique description. The interpretation of restricted relatives is similar to the interpretation of simple defnps stated in rule 5 since both depend on the hearer's knowledge of how many memory elements fit the description. In (28), the defnp seems too common to describe or refer to one individual, while in (29) the defnp may be either generic or specific.

(28) The man who reads *The Grapes of Wrath* will like it.
(29) The man who performs acrobatic stunts with a triple scoop ice cream cone will give an exciting show.

It seems that a knowledge base must be able to characterize certain information as descriptive of many persons (for example, reading a well-known book). If a network hierarchy such as the one discussed earlier is used, the characterization would require

statements true of people in general to be placed in the network hierarchy in such a way that they could be inherited as true by any sub-network below a given point in the hierarchy. Any other descriptions ambiguously distinguish between a characterization of a class and some individual.

As with other defnps, the sentence context either confirms, shifts or denies a generic reading. A predicate that accepts a generic reading will confirm a defnp which by itself appears generic as in (28). A predicate that prefers a specific argument will force a specific interpretation for a defnp which may be read either way, as (30) shows. Some predicates do not distinguish, and when the defnp alone is u-ambiguous, the context provided by the resulting sentence does not distinguish either, as in (31). An unacceptable sentence results from the combination of a generic defnp and a predicate which requires a specific argument, as (32) indicates.

(30) The child who goes to bed early is my bratty little sister.
(31) The woman who teaches drama at Franklin High gets a good salary.
(32) * The man who reads *The Grapes of Wrath* will come to dinner on Saturday.

Defnps with attached prepositional phrases function differently than the other defnps. They differ because the embedded prepositional phrases function as a context in which the head noun phrase of the defnp is specified. Thus in (e5) below, *baseball game* specifies a context for *pressbox* which specifies a context for *reporter*. The context allows the hearer to infer the relationship of the head noun phrase to the embedded phrases and to interpret the entire phrase generically or non-generically.

> **Rule 8**: A defnp with several attached prepositional phrases "inherits" the specific/generic feature from the most embedded noun phrase. If the noun phrase is definite and u-ambiguous in isolation, the full phrase will be as well, while when the most embedded noun phrase is indefinite, the full phrase is generic.

Rule 8 is exemplified by (e5) which is specific, (e6) which is u-ambiguous and (e7) which is generic.

(e5) the reporter in the pressbox at the baseball game
(e6) the child in the monkey house at the zoo
(e7) the reporter in the pressbox at a baseball game

Why are such cases different from other defnps? The distinction is determined by the prepositional phrase. Some prepositional phrases specify a context in which the head noun phrase is characterized. In (e8), *pressbox* specifies a context in which certain kinds of people are likely to be found, so that *reporter* in that context indicates a particular instantiation of the people in the pressbox. The result is a chracterization of the kind of reporter, i.e. one who is in the pressbox. The examples below give similar characterizations.

(e8) the reporter in the pressbox
(e9) the reporter on the field
(e10) the leader of the platoon
(e11) the member of the platoon

In the initial sentence of a discourse, these defnps may be generic. However, they may also be specific since the noun phrase may be a description of an individual. The defnp in the prepositional phrase is used to indicate that the entity so defined is knowable. On the other hand, certain noun phrases may be clearly distinguished as charactizations because the noun phrase in the attached prepositional phrase is unambiguously generic. The phrase *the average school* is only a characterization; as a result, (e12) is generic.

(e12) the teenager in the average school

The defnps above may be contrasted with defnp-prepositional phrase pairs when the prepositional phrase specifies properties of the individual. The noun phrase therefore is a description and non-generic.

(e13) the man with the gun
(e14) the lady at Mortimer's funeral
(e15) the woman in the hat

As with other defnps, distinguishing property specifications is difficult. Worse yet, because a language must be productive in allowing new analogies, new characterizations may be created from noun phrases that were once property specifications. For example, a head noun phrase may specify a class and the prepositional phrase provide a description of a sub-class. These cases are different from (e8) and (e9) where the prepositional phrase specifies the class. An example of such a defnp is given in (e16) which may be used as a generic term for a businessman by some speakers.

(e16) the man in the grey flannel suit

As with other defnps which are either generic or specific, the sentence predicate confirms, shifts or denies the generic reading. For defnps which have a strictly generic reading, use of a sentence predicate which demands a non-generic results in an odd sentence while u-ambiguous defnps either become generic in a sentence type which is generic, or non-generic in a sentence which requires non-generics.

(33) The leader of the platoon is responsible for all individuals under him. <generic>

(34) The leader of the platoon worked hard getting the platoon combat ready. <specific>

(35) * The teenager in the average school was invited to the White House for a special program.

This discussion confirms a suspicion which the reader may have had previously: many defnps are u-ambiguous in isolated sentences. However, since there are some sentence and noun phrase types where u-ambiguity does not occur, these cases should be detected as well. How might such detection be possible in a computational model? The rules presented so far fall into three classes; rules 1, 2 and 6 are sentence related, rules 3, 7, and 8 are single defnp centered, while rules 4 and 5 depend on relations between defnps. The sentence related rules may be detected by various kinds of sentence template information; for example, part of the case frame semantics for verbs may specify those case positions which require a specific or generic noun phrase filler. Rules 7 and 8 may be implemented as part of the semantic interpretation of the individual noun phrase. However, these rules require a knowledge system which would embody the questions which are part of rule 7. Such a system remains a topic of research in artificial intelligence. Rules 4 and 5 present a different problem; the computation for finding some relation between two noun phrases may be possible within a Fahlman net, but that process must occur after all other semantic interpretation is complete. It is unclear just how this computation should fit into an understanding system.

The next section will discuss the role of focussing in generic interpretation. The rules presented so far will be used to show how the defnp in isolation differs from the same defnp in context. Because some of the rules in this section seem computationally infeasible, the examples will draw on those rules which are less so, that is, on rules 1, 2, 3, 8 with indefinites in the prepositional phrase, and 4 for sentences of only one defnp. Rules 8 and 4 will be used in a restricted form which lessens the difficulties associated with the full rules. However, for a full theory of focussing with

generics, some theory of generic interpretation will be necessary.

Before generic specification may be considered, one last digression about generics is necessary. The meta-rule for generics given in the next sub-section will not be used in the discussion of generic specification, but it may help the reader understand something more about the curious nature of generic phrases in discourse.

### 3.6.2 A Meta-rule for Interpretation of Generic Defnps

The examples in the previous section show that in many cases a defnp is u-ambiguous, and that sentential information may not eliminate the ambiguity. Hearers rarely notice u-ambiguities, so the question remains: why do u-ambiguous sentences appear non-ambiguous?

An assumption made in chapter 1 is that speakers and hearers adhere to the Gricean maxim "Be perspicuous." Grice describes perspicuity by sub-maxims such as "avoid obscurity of expression, avoid ambiguity," and so on. Since the hearer is decoding a speaker's message without access to the speaker's mind, the hearer may assume a particular interpretation of a defnp because the alternative interpretation, if meant, is r-ambiguous. For example, (36) uses *the white man* generically although the entire sentence is u-ambiguous. If the speaker intends the hearer to interpret the defnp as specific, speaker and hearer must both know one individual so described.

(36) Bill considers the white man to be the source of Boston's social unrest.

When there are either too many or no individuals for a specific use, the message is not perspicuous, because the phrase is u-ambiguous, thereby violating the demand that the speaker "avoid ambiguity." If the speaker had intended a specific reference, an alternative formulation such as "Moses Bown" or "the man who is the city's number one administrator" would be an unambiguous use. By assuming the speaker is being perspicuous, the hearer is able to interpret u-ambiguous defnps such as *the white man* in (36) as unambiguous.

For the same reason, a defnp which seems to be generic may be read as a non-generic. A hearer should understand sentence (32), reproduced below, as referring to some person who is described as a reader of *Grapes of Wrath*. If the noun phrase is interpreted as a generic, the interpretation results in an odd sentence, so the hearer must

choose a more acceptable reading. The non-generic interpretation is acceptable because the resulting sentence maintains the communication maxims of perspicuity.

(32) The man who reads *The Grapes of Wrath* will come to dinner on Saturday.

(37) Invite the man who reads *The Grapes of Wrath* to dinner.

An objection to this explanation is that the hearer of (32) does not know the referent of the defnp, and hence the sentence is not perspicuous even on the non-generic reading. As was indicated earlier, not all defnps must have referents that are known to the hearer; the referent may be knowable but not known at the time of the utterance. (32) must be interpreted in this way. As was shown earlier, certain sentential forms, such as introductory assertions, do not necessarily require known referents of the defnps in the sentence but others, especially commands, do require that referents be known. Thus (37), which is a command, fails to be perspicuous on either reading.

Many initial sentences in a discourse are not in accordance with this maxim. In some cases the speaker or author may intend to disregard the maxim fo literary reasons. For example in D20, the reader following the maxim and rule 4a, will interpret the first sentence as generic and later will be forced to re-evaluate the interpretation, all of which lends a certain feeling to the story.

D20-1 The robot is replacing the car. Sandra pondered a while. Jack must be crazy getting rid of the car, she thought. We need it for transportation, and the Wirly's robot will never do all the things we use the car for.

Often, however, speakers may fail to notice their u-ambiguous noun phrases, so hearers may employ the strategy of assuming the perspicuous reading and looking for evidence in the rest of the discourse that indicates the speaker's intent to use the maxim. In an essay beginning with (38), the reader may assume that the underlined defnp is generic based on perspicuity, but s/he may look for further evidence[1] later in the essay. Such evidence may not be forthcoming (one may imagine an essay which gives no indication of which reading is intended beyond the continued use of the full noun phrase), and the resulting essay may be judged as somewhat vague.

---

1. Answers to the question "are several such individuals mentioned in the following discourse?" provides evidence since each individual is an example supporting the generic statement. If the defnp is used repeatedly, the reader has no reliable evidence.

(38) The Chinese government has always maintained a policy of isolationism.

The conclusion to be drawn is that the hearer's comprehension task involves not only using phrasal and sentential level rules to interpret the possible use of a defnp, but also relying fundamentally on three of the assumptions of communication specified in chapter 1. By assuming that the speaker intends to communicate about something, and to make communication requests, the hearer may assume the speaker has intended to refer to something or to depict a class of entities. By assuming that the speaker and hearer share some common knowledge of their environment, the hearer may resolve cases of u-ambiguity of defnps based on the presence or absence of particular entities from the hearer's representation of shared knowledge.

## 3.7 Focussing and Generic Specification

Focus provides a means for disambiguating generic defnps by use of the rules given above and use of the focus in explicit backwards co-specification rules. As discussed earlier, the EBC rule predicts that a defnp which is u-ambiguous will be generic or not based on the focus. By contrast, a defnp which is generic in a sentence in isolation will be unaffected by focussing. The strictly generic defnp case may move the focus to the generic interpretation. D21-2 is generic because the sentence is a case of rule 1.

D21-1 I got a new ASR 33 this week.
2 The ASR 33 is an old but reliable output device.

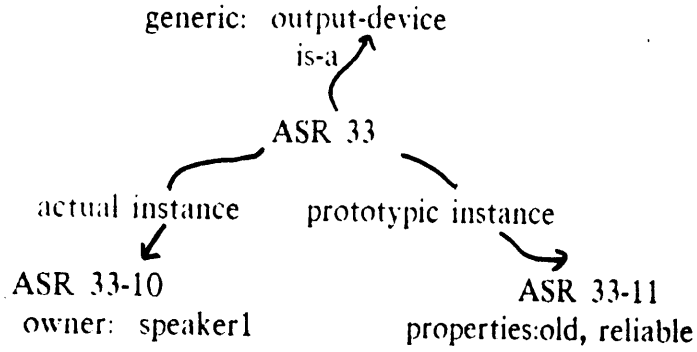
This example suggests a new rule governing generic defnps and their relation to focus. That is, a defnp which is generic and contains the same head as the focus when the focus is non-generic is associated with the focus by generic specification. Figure 3 depicts both the specific and generic uses of ASR 33 in D21. The hierarchy indicates why both a prototypic and actual instance of a generic element are needed.

Before this rule is added to the collection given previously, one other example must be considered.

D22-1 The cat is an independent and solitary hunter.
2 But not my cat;
3 she hunts in the company of my neighbor's cat.
4 They make quite a pair.

The initial focus of D22 is a generic defnp used as a prototype. The explicit backwards

**figure 3.3. Generic and Specific Representations of D21**




---

co-specification rule predicts that *my cat* should co-specify with the focus since *my cat* is specific. Yet the EBC rule should not hold here because *my cat* is distinctly specific while the focus is generic. D22 shows a symmetry with D21 for defnps which are different on the generic/specific reading from the focus. Given both examples, the generic association rule may be stated and the explicit backwards co-specification rule may be re-formulated:

> The Explicit Backwards Co-specification rule: A u-ambiguous defnp which contains the same noun phrase head as the focus, which follows the focus in the discourse, and which does not contain more information than is included in the focus, co-specifies with the focus. Generic and non-generic defnps with the same noun phrase head, etc., co-specify with the focus when it is the same generic/non-generic type.
>
> Generic and Non-Generic Specification: A defnp which contains the same noun phrase head as the focus, which follows the focus in the discourse and contains no more information than the focus and which is of a different type than the focus on generic/non-generic feature, is associated with the focus by generic specification when the defnp is generic, and by non-generic specification when the defnp is non-generic.

The strictly generic defnp used after a non-generic focus is just one case of implicit backwards identification using prototypes. Generics occur in associated specification as well. For associated terms, the u-ambiguous defnp is dependent on the

reading of the focus, unless a full modifying noun phrase is attached, as D23 shows.

D23-1 I want to have a party.
2 (a) The time of a party is hard to decide on.
(b) The time is hard to decide on.

*The time* as a simple defnp may be used only as an associated specification of the time of a party. Only the complex noun phrase has the syntactic and semantic distinctions which reflect the generic usage. For D23-2a, rule 8 predicts that the defnp is generic in isolation. Hence *the time of a party* is an associated specification not to the focus of D23-1 (*a party*), but to the prototypic party associated with the focus. The associated specification rule may be re-formulated as follows:

> A defnp will be judged as a specification of an element associated with the focus if the defnp names an element associated with the focus by direct or inherited associated, and when the defnp is either u-ambiguous or of the same generic/non-generic type as the focus. When the defnp is generic and the focus non-generic, the defnp is an associated specification of the prototypic instance of the parent element of the focus.

Set-element identification exists for generic foci as well as specific foci. A significant difference is that the focus for the generic case may be a singular defnp or it may be a plural noun phrase with either a definite article or no article. The set membership is indicated by a distinguishing modifier, just as with specific set-element implicit identification. An example of generic set-element identification is given below with both a singular defnp focus and a plural unspecified focus.

D24-1 (a) The Australian aborigine represents an almost extinct hunter-scavenger social group.
(b) Australian aborigines represent an almost extinct hunter-scavenger social group.
2 The aborigine in the southern sections of Australia sometimes gathers food, but the other aborigines do not.

Inferred generic identifications also occur. In D25, *the owner of a motorcycle* is a generic defnp:

D25-1 Alfonso was in an accident with a motorcycle last week.
2 I think the owner of a motorcycle ought to be required to take driving lessons.

*The owner of a motorcycle*, generic by rule 8, is generically related to the first sentence

by an inference of what happened to the agent. The same kind of machinery that is used for inferred specifications for non-generic defnps is used for making these inferences as well. How are generic inferred specifications to be distinguished from non-generic inferred specification? A strictly generic defnp as in D25-2 remains generic. Those defnps which are u-ambiguous at the sentential level, as with D13-2, are disambiguated as specific because of their relation to the focus.

This section has presented some modificiations needed in the explicit backwards co-specfication rule and in the implicit specification rules. These modifications are motivated by the difference in use between defnps which are generic and those which are not. The reformulations allow some speculation about the generic rules presented previously, namely, speculation about rules 4 and 5. These rules are somewhat unsatisfactory because they demand a kind of processing that seems out of place in normal sentence interpretation, processing which is not compositional but requires a view of the relations among syntactically disjoint parts of the sentence. Perhaps rules such as 4 and 5 are not used unless a judgment is forced. Instead, simple defnps are assumed u-ambiguous and the context, especially the focus, is used to disambiguate a reading. On this explanation the defnp in (19), repeated below, will be u-ambiguous without concern for other noun phrases; its disambiguation will depend on contexts such as D26.

(19) Steve's go-cart ran well against the motorcycle.

D26-1 (a) There were three different vehicles in the race today: a hand powered motorcycle, a scooter car, and a gyro controlled soap box cart.

(b) One of the most difficult vehicles to beat in any race is the motorcycle.

Clearly this speculation remains to be confirmed by further research in the understanding of generics.

## 3.8 Conclusions

This chapter has covered a wide range of phenomena and data. It has presented some general issues about non-anaphoric defnps and some rules for generic defnps. The main theme of this chapter has centered on an algorithm and rules for co-specification of anaphoric defnps, specification of defnps associated with focus and the network representations needed to use the rules.

The discussion of an algorithm for co-specification and specification using the focus has relied in a fundamental way on the type of network discussed by Fahlman and represented in figure 1. The purpose of this network is not only to encode the kinds of knowledge which speakers and hearers have when they speak (or decide not to speak) their language, but also the network shows how speakers actually use their language. Simply put, speakers encode representations, which are encodings of entities in the world, into phrases in their language. Because it is so general, this statement is not interesting in itself. However, the process of encoding pieces of network representations becomes interesting because it is combined with a reduction of redundancy in the surface form. Once a meeting is mentioned, the speaker need not tell the hearer that *the time* is the time of the meeting; because of the network of knowledge, the hearer knows the connection between the two without explicit marking in the utterances of the speaker. Thus the redundacy is not lost, it is simply embedded in a different level of comprehension.

The network of representations of knowledge is necessary for the hearer to keep up with the speaker. Consider what happens when a piece of it is removed. Either the hearer must decide that a defnp such as *the time* is disassociated with meeting, or s/he must use the perspicuity maxim and supposition to infer a connection. The former results in incorrect comprehension; the latter is successful only when the price of additional processing time needed to make the inference may be paid, for example when the hearer is reading a text. When a hearer must act in "inference mode" for many phrases, comprehension becomes difficult because confusion occurs when suppositions may not be made, and because in spoken situations, comprehension must occur in real-time.

Of the focussing rules discussed in this chapter, the inferred association rule is most problematic. There are two difficulties in detailing a computational framework for inferred association. The first is representing inferences and making suppositions. Certain kinds of inference, particularly those involving events occurring in time, are not well understood, while making suppositions is a largely unexplored area of computational processes for general reasoning. The focussing examples suggest that inferred association may not be possible without a focus mechanism. A second difficulty is the modelling of and reasoning about speaker-hearer beliefs. Suppositions about defnps involve the initial

supposition that "the speaker wants me to believe these two phrases are related" or at least "the speaker believes these two phrases are related." Such a supposition not only involves modelling beliefs of the speaker, but judging what these reasonable beliefs about the speaker might be. Some recent work by Cohen [1978], Doyle [1978] and Allen [1978] provide modelling, but further research is still needed.

The main thrust of this chapter has been to provide a computational account for how an anaphoric defnp co-specifies with another expression. The computational account has required us to address more than just co-specification because once a hierarchic network of associations and a focus of discourse are assumed, some defnps which are used to specify certain elements in memory may be considered as well. The account of co-specification includes constraints on the kind of information associated with the focus, the presence of is-a links in the hierarchy, and the role of generic phrases. The account of specification for certain non-anaphoric defnps indicates which are specified relative to the focus and which relation the phrase bears to the focus.

# 4. Focus in Pronoun Anaphora Resolution

## 4.1 Focussing and the Characteristics of Pragmatic Anaphora

Chapter 1 presented four characteristics of language which affect pronoun anaphora. These characteristics must be included in a theory of definite anaphora comprehension. Examples of each are given below.

1. Universal or existential interpretation of a noun phrase (from Webber):
The men who tried to lift a piano dropped {it, them}

2. Semantic selectional restrictions (from Katz and Fodor, Woods, and Walker):
I really like roast pig. It tastes great.
Pratt roasted a pig in his fireplace but no one realizes it.

3. Sentential syntactic restrictions on disjoint reference (from Lasnik and Reinhart):
Near Dan, he saw a snake.
John loves his mother.

4. Bound anaphora (from Webber, and Partee):
No child will admit he's sleepy.

This chapter will present rules which rely on the discourse and actor foci to govern the co-specifications of pronouns. The rules use of a representation which delimits universal/existential interpretation and use of information from sentence syntax and semantics for the other characteristics as part of the co-specification process. By explicitly taking advantage of syntactic and semantic sources of information, focussing rules for pronoun anaphora may be viewed as a kind of selection process. The rules select from the various pieces of information those which are relevant for the pronoun at hand. Selection behavior is further evidence for focussing as a control of definite anaphora interpretation.

This chapter will also substantiate two claims of chapter 1:

1. Focussing controls the inference mechanism needed to determine an antecedent relationship between a focus and an anaphoric noun phrase, because inferencing is used to confirm a hypothesized link between an anaphor and a focus.

2. Focus, together with a representation such as Webber's and a computation of c-command, provides a mean for disambiguating pragmatic anaphora.

These two claims, together with an explanation for how the focus mechanism may be used for pronouns, allow the statement of rules for pronoun interpretation. These rules provide an acount for (1), (2) and (3) below and some observations about the nature of (4).

1. (Inference--Charniak [1972], and Rieger [1974]) When Janet heard about the costume ball, she thought about what Mother could wear. Mother had to tell her that she had not been invited.

2. (Foregrounding--Chafe [1975]) Yesterday I saw a little girl get bitten by a dog. I tried to catch the dog, but it ran away. The little girl was scared, but she wasn't really hurt.

3. (Focus popping-Grosz [Deutsch, 1974]) Bolt the pump to the base plate. {intervening 1 minute of discussion about using the ratchet wrench, its location, how an extender is put on}. It is bolted, what should I do now?

4. (Parallel actions) Put the mud pack on your face for 5 minutes, and then pull it off.

This chapter has the following structure. First the basic rules for co-specification of a pronoun are presented informally and then a full set of rules for the pronouns are given. Examples of how the rules are used are presented as well. The chapter then turns to a discussion of actor foci so that the motivations for the actor focus in the pronoun rules are clarified. More examples of the rules using actor and discourse foci are given. Discussion also includes some further comments on the syntactic restrictions on disjoint reference. Rules for possessive pronouns are then shown, and examples of their use are given. The remainder of the chapter addresses issues concerning representation and the limitations of focussing in pronoun rules. The reader is cautioned that this chapter covers a variety of issues and anaphoric uses.

Therefore, the chapter is long and detailed. To balance the detail, the examples which occur throughout provide the reader with a touchstone for the issues which are discussed.

## 4.2 The Basic Rules

Chapter 2 mentioned briefly the rules for choice of co-specifier of a pronoun given a focus and an alternate focus list. Two fundamental rules, the recency rule and the basic focus rule, are stated below.

> (Recency Rule) If the pronoun under consideration occurs in the subject position, and there is an alternate focus list noun phrase which occurs as the last constituent in the previous sentence, test that alternate focus list phrase for co-specification before testing the current focus. If acceptable both syntactically and inferentially, choose the alternate focus list phrase as the co-specification of the pronoun.
>
> (Basic rule) Test the discourse focus for syntactic criteria and inference checks as the co-specifier of the pronoun. If acceptable, mark the pronoun as co-specifying with the current focus. If unacceptable, test the members of the alternate focus list.

A simple example, from Denes and Pinson [1973], will illustrate how the focus is used.

D1-1 If a neuron remained in its inactive condition indefinitely,
2 it would be of little use to the nervous system.
3 When it is stimulated strongly enough,
4 its delicate ionic balance is upset.

The expected focus after D1-1 is *a neuron*, since it is the affected entity. In D1-2, *it* may potentially co-specify with *a neuron* or *inactive condition*, if one considers all the noun phrases in the sentences. However, the focussing rules reduce the choice to the expected focus. This co-specification is acceptable, and so the anaphora are resolved, and the expected focus is confirmed as focus. D1-2 also introduces as potential focus *the nervous system*. In D1-3, *it* may potentially co-specify with the neuron (the focus), or *the nervous system* (the potential focus). The focus of neuron is chosen because the recency rule is not met.

The recency rule makes focussing seem somewhat ad hoc. What is the

justification for the recency rule as part of the focussing rules? An old grammar school heuristic states that the antecedent of a pronoun is the last noun phrase that agrees in gender, number and person with the pronoun. Below are three examples where the underlined noun phrase co-specifies not with the focus, but with the last noun phrase in the preceeding sentence. I have observed the focus recency rule to be consistently accurate. While I see no clear reason for this phenomenon, it may be the case that recency is used by speakers to change their focus immediately using the recency of speaking to override the power of focussing.

D2-1 Mary is giving a surprise party at Hilda's house.
2 It's at 340 Cherry St.

D3-1 Today was Jack's birthday.
2 Penny and Janet went to the store.
3 They were going to get presents.
4 Janet decided to get a kite.
5 "Don't do that," said Penny.
6 "Jack has a kite.
7 It's a real neat one."

D4-1 Fill the pan with the cake mixture.
2 It will be slightly lumpy.

As chapter 2 indicated, the focussing algorithm must be amended to include actor focussing. Actor and discourse foci are interrelated in the disambiguation of pronoun anaphora, but first the general behavior of actors must be considered. The next section describes how actor foci work. First the full set of rules for non-possessive pronouns will be given. There are a number of cases such as unreliable pronoun use, and non-backwards antecedent pronouns which will be discussed after actor focus has been explained. The use of these rules, complete with focus movement, will be shown in examples in the later sections of this chapter. The first two rules will be stated in English because of their simplicity. The other rules will be given by flowchart.

Rules for co-specification of non-possessive personal pronouns:

(1) First and second person singular pronouns co-specify with the speaker and hearer of the discourse unless the sentence in which they occur is in quotes.

(2) First and second person plural pronouns: use the actor focus if acceptable, followed by the focus for a class with the speaker (hearer) in it.

(3) A prediction of a focus as co-specification of a pronoun succeeds if the focus phrase meets syntactic anaphoric filters and is accepted by the inference mechanism.

Figures 1 and 2 are to be read using the following conventions. The circled parts of the algorithms are end states. Some of these reflect the dimensions of the focussing theory; for example, the state labelled "forward co-specification or backwards non-antecedent pronoun" indicates that one of these two conditions holds. The rules do not indicate how to proceed from these states since these conditions are not necessarily explained by focussing. In the figures, *yes* is abbreviated by Y, *no* by N, success (of the syntactic and semantic filters and inference mechanism) by S, and failure by F.

By carefully observing the rules given thus far, the reader will see that these rules depend fundamentally on the focus and its associations. The discourse focus is preferred for non-agent anaphors as the choice of a co-specification, while for agent anaphors, the actor focus is preferred. When that choice fails to meet certain syntactic, semantic and pragmatic constraints, potential discourse foci are considered in turn, followed by the actor focus for non-agent anaphors, and the discourse focus for agent anaphors. The rules for agent position anaphora also include discourse actor focus conditions. This special behavior will be delayed until the next section.

There are some caveats about the use of these rules. Rule 1 is somewhat incomplete, since in quoted reports of what someone has said, the speaker and hearer may be someone other than the default speaker and hearer of the discourse. A full treatment of these cases requires a context which indicates who the current speaker and hearer are. A complete rule may depend on more global focussing conditions such as those discussed by Grosz [1977] since the speaker and hearer are outside of the discourse in the same way as the referent of the first use of a name.

A second caveat concerns the situation where there is no focus. Suppose a discourse begins with:

**figure 4.1. Rule for Third Person Pronoun in Agent Position**

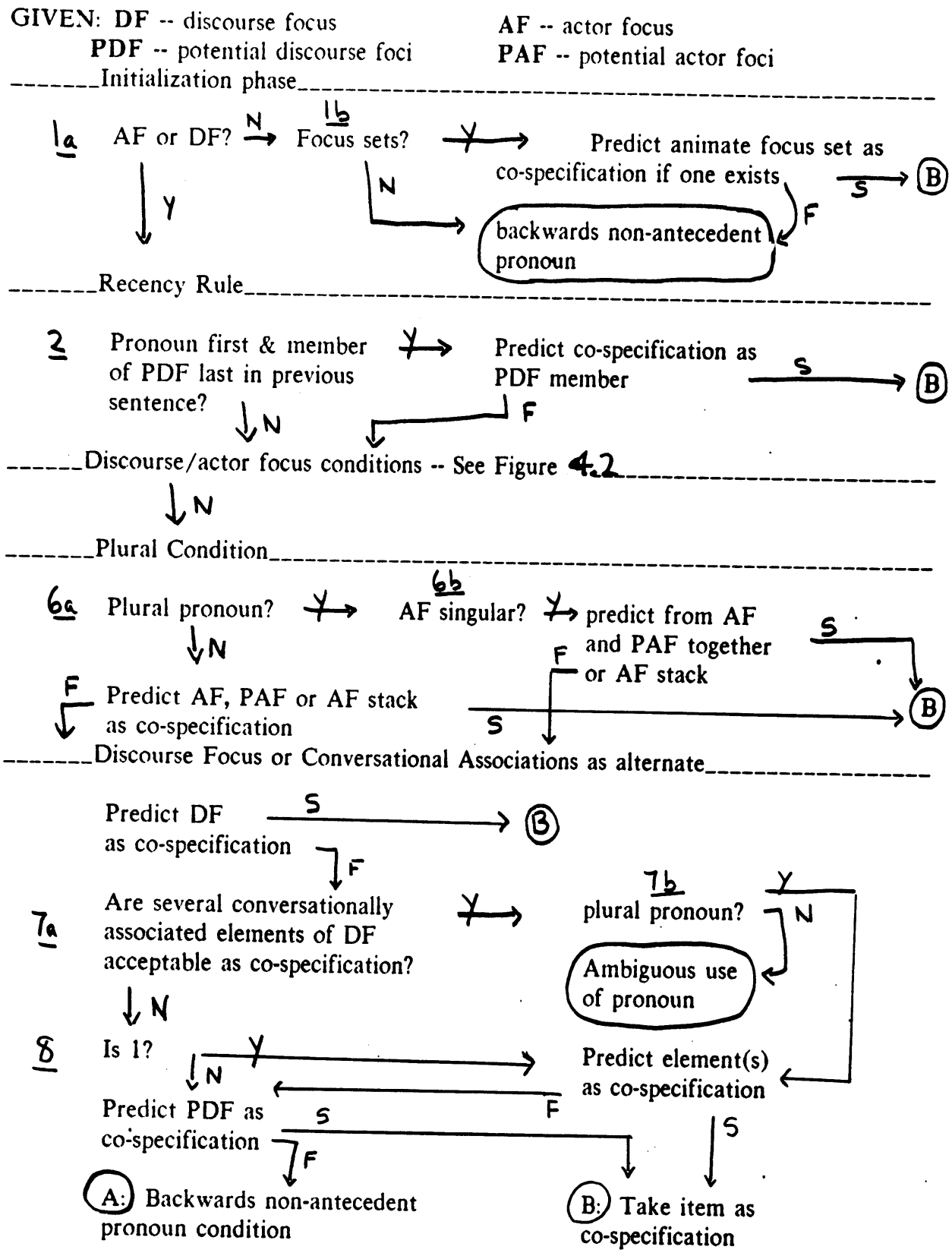
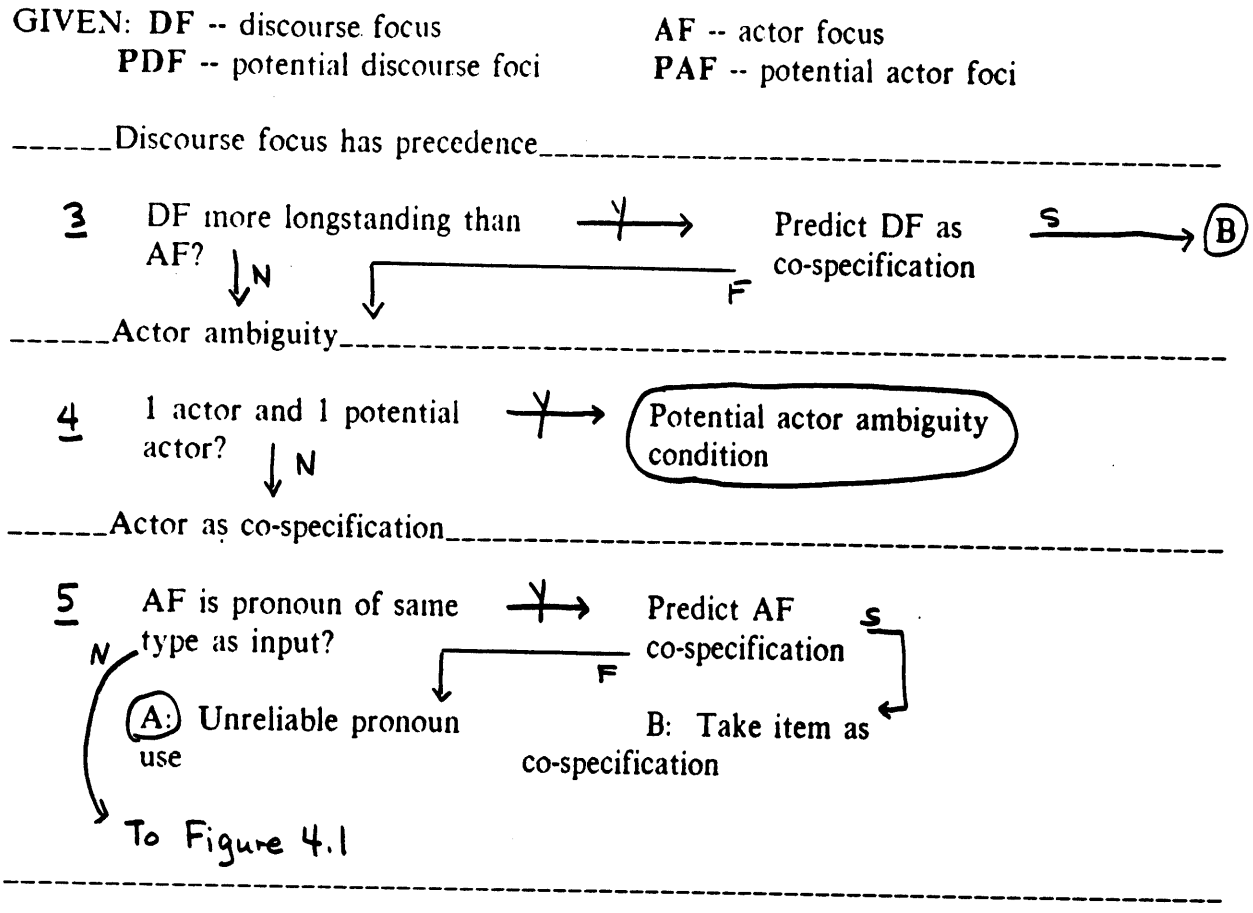


**figure 4.2. Discourse/Actor Conditions for Agent Third Person Pronoun**

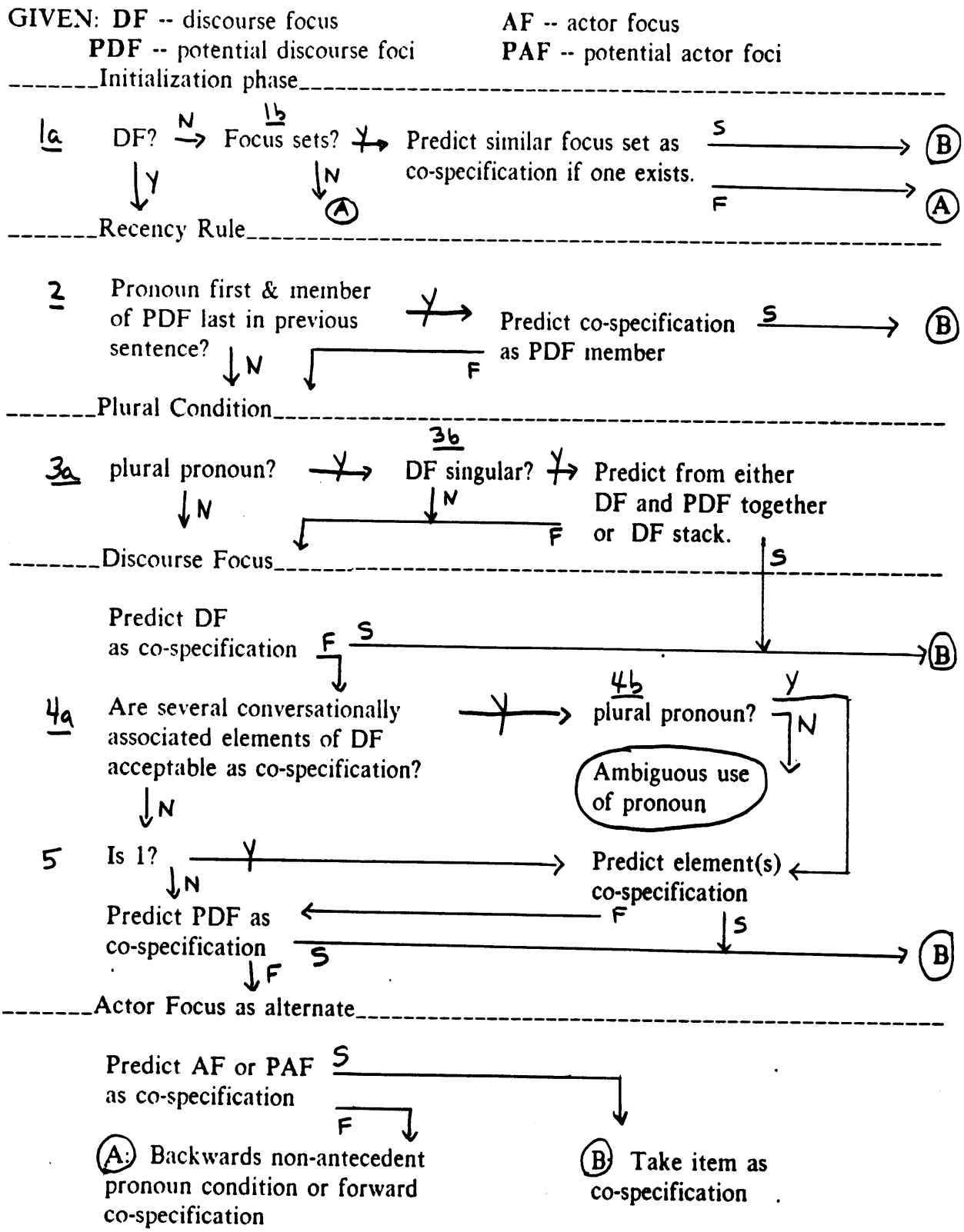



(1) Shem loves his sister.

The use of *his* does not rely on the focus since none is established. Therefore, the rules above make no predictions when a focus is not established. Co-specification within a sentence may work quite successfully by partitioning the noun phrases into semantic classes, but this heuristic is not a part of focussing. When (1) occurs in mid-discourse, the *his* must co-specify with the actor or discourse focus, that is, the focus rules apply to pronouns which may freely co-specify with either a noun phrase within a sentence or one outside of it. The focussing rules force the freely co-specifying pronoun to co-specify with a focus if one is established; the focus is therefore the source of co-specification rather than some other noun phrase within a sentence.

A third caveat concerns the phrase "conversationally associated elements of the discourse." Because of the nature of the hierarchical network which includes a representation of the focus, there are many elements associated with the focus. Thus if

**figure 4.3. Rule for Third Person Pronoun in Non-Agent position**



a discourse is about steak and kidney pies, the elements of the crust of the pie, the vegetables in it, and the temperature at which it is eaten are associated with it. However, for purposes of pronoun co-specification, none of these associations may be co-specifications for pronouns unless explicitly mentioned in the text. For example, the speaker may not use *it* to co-specify with the crust of the pie until the crust is mentioned in the discourse. Hence only elements associated with the focus by means of their explicit mention in the text are potential sources for pronoun co-specification.

The rules presented in figures 1, 2 and 3 mark predictions with S or F branches. An S branch is taken if the predicted item meets the syntactic, semantic and inference criteria relevant to the sentence. Syntactic criteria include gender, number and person as well as the disjoint reference computation of Lasnik and Reinhart. Semantic constraints include rules of scope; these constraints will be discussed further later in this chapter. The inference criteria are judged by the inference mechanism described in chapter 2 and further explained in chapter 3.

Briefly, the inference machine confirms the predictions of the focussing rules by proving that the sentence with the pronoun replaced by the co-specified noun phrase is consistent with other knowledge. Of course, what constitutes consistency with other knowledge is complex. Certainly it might include semantic selectional restrictions. Focussing rules will predict that *it* in D5 below co-specifies with a pig (the inital expected focus).

D5-1 Pratt roasted a pig in his fireplace,
2 but no one realizes it.

On semantic selectional restrictions, an inference machine could reject this choice. For other sentences deeper reasoning is required. For example, in the sentence pair below, to be discussed fully later, the initial prediction of focussing rules is that *him* co-specifies with *my dog*.

D6-1 I took my dog to the vet yesterday.
2 He bit him in the hand.

This choice must be rejected by the inference mechanism on the basis of information which states that dogs cannot be bitten in the hand because they do not have hands. This report will not detail the functioning of the inference mechanism; its function is an important aspect of ongoing research in artificial intelligence. Instead this report adds to our understanding of inference mechanisms which may be used in computational

systems because it indicates what the mechanism must necessarily contain for a processing model of definite anaphora comprehension.

To conclude this section, two simple examples of pronoun anaphora will be traced through the algorithms in the figures.

D7-1 I lost a necklace at the office yesterday.
2 I inherited it from my grandmother,
3 and it meant a lot to me.

D8-1 Yesterday Max went to Bloomingdales with Ned and Winston on a shopping trip.
2 While he was there, he bought some sneakers for his mother.

D7 shows the action of the focussing processor for pronouns in non-agent position. The discourse focus of D7-1 is *a necklace*, its sentence interpretation and discourse representation while the actor focus is the speaker. The rules in figure 3 cause the focussing processor first to check that there is a focus, that the recency rules do not hold, and whether the pronoun *it* of D7-2 is plural. The next step is to predict that *it* co-specifies with the focus of necklace. Since a necklace may be inherited from someone and that information is consistent with having lost it, and since the syntactic constraints do not rule out the co-specification, *it* is taken to co-specify with the necklace in focus. The reader will note that the alternate potential focus of office is never considered by the processor.

D8 gives a simple example of the use of the agent position rules. The actor focus of D8-1 is *Max*, its sentence representation and database specification (if one is found), while the discourse focus is Bloomingdales. The potential actor foci include Ned and Winston. To interpret *he* in D8-2, the focussing processor answers *yes* to decision 1a, *no* to decision 2, *no* to 3 (because the discourse focus and actor focus were established in the same sentence), *no* to 4 (because there are two potential actors), *no* to 5, *no* to 6a and finally reaches a prediction. It takes the actor focus as the co-specification of *he*. This choice involves checking syntactic and semantic constraints, which do not force rejection; since the actor focus of Max is consistent with buying sneakers, inference criteria are satisfied as well.

## 4.3 An Actor Focus

As chapter 2 indicated, the actor or actors in a discourse are distinguished from the discourse focus. An actor is an animate object which may function as the agent of a particular verb. Actors may co-specify with the focus of the discourse, as D9 shows.

D9-1 I haven't seen Jeff for several days.
2 Carl thinks he's studying for his exams.
3 Oscar says he is sick,
4 but I think he went to the Cape with Linda.

In D9, the speaker, Jeff, Carl and Oscar all can function as actors. Among these, the speaker, Carl and Oscar actually are the actors while Jeff is the focus of discussion. Actors are distinguished because in many discourses, the actors are mentioned and then pronominalized in addition to the discourse focus. In a sense then, there is a separate actor focus in many discourses which may be used to properly interpret anaphoric expressions which co-specify with the actor. When the discourse focus is inanimate, interpretation of anaphora is straightforward, but when the discourse focus is animate as well as the actor focus, the interpretation is more problematic.

Usually there is no need to confirm an expectation of the actor focus because the actor focus is simply the database element associated with the agent case of the verb. Additional potential actors may be specified elsewhere in the sentence, as with Jeff in D9-1. A rule for pronominalization of pronouns in agent position may be stated:

> Animate Discourse Focus Rule: A pronominal expression in agent position which meets person, number and gender agreement[1] with the actor focus co-specifies with the actor focus unless the pronoun may also co-specify the discourse focus. In the latter case, if the discourse focus was established in a sentence before the actor focus, and meets syntactic anaphoric filters, the discourse focus is the intended co-specification.

In D9, the discourse focus is Jeff while the actor focus moves from the speaker to Carl to Oscar. All occurrences of *he* co-specifies with Jeff since Jeff was established as

---

1. Henceforth these tests shall be referred to as syntactic anaphoric filters.

discourse focus before any of the other phrases which satisfy person, number and gender.

The animate discourse focus rule does not indicate what to do when the discourse focus and actor focus are both animate, have the same gender, number and person, and are established during the same sentence of the discourse. In this case, the co-specification of a pronoun is ambiguous. If the pronoun fills an agent case, the actor is preferred, but this preference is not a strong one. In D10-2a below, *he* co-specifies with *John* (the actor focus) but if D10-2b followed D10-1, *he* may co-specify either John or Mike (the discourse focus).

D10-1 John called up Mike yesterday.
2 (a) He wanted to discuss his physics homework.
(b) He was studying for his driver's test.

It appears that in such cases the ambiguity may not be easily resolved unless additional information is known about the two foci which stipulates that the sentence is true of only one.

Just as there are potential discourse foci, there are potential actor foci as well. A potential actor is a noun phrase which specifies an animate element of the database and does not occur in agent position. In most sentences, the noun phrase in agent position specifies a database element, and the actor focus determines the co-specification. But when the noun phrase in agent postion is a pronoun, it may co-specify with either the actor focus or a potential actor. Ambiguities for an anaphor in a sentence occur in the case where an actor and one potential actor are both present in a previous sentence and where the discourse focus is a non-actor entity. These circumstances are analogous to the ambiguity between the actor and animate discourse foci.

Suppose that sentence (2) below followed each of (3), (4), (5) and (6).

(2) He knows a lot about high energy physics.

(3) Prof. Darby will tell Monty about the neutron experiment.
(4) Prof. Darby will lecture Monty on the neutron experiment.
(5) Prof. Darby will help Monty with the neutron experiment.
(6) Prof. Darby will teach Monty about the neutron experiment.

Some native speakers find all of these sentence pairs ambiguous, that is, they cannot tell whether *he* is co-specifying with Monty or Prof. Darby. Some native speakers find only

the pair (3) followed by (2) ambiguous; in the other three cases, there is no ambiguity. These examples are surprisingly similar to D10. It appears that when a particular actor is in focus, and a potential actor appears as well, sometimes the actor focus is not chosen. Here again stress may be a deciding factor, but in order to consider other factors, suppose stress is not relevant to the Darby-Monty cases. How do some speakers decide that *he* co-specifies with Monty or Prof. Darby? It appears that they choose either the actor focus or potential actor and look for evidence for why their choice is preferred over the other possibility. When that evidence is not forthcoming, informants are confused. Such a judgment suggests that the inference mechanism postulated thus far be capable of a special judgment when given one actor and one potential actor; it must weigh its findings, and choose one of the two candidates as superior.

What kind of evidence can be used in such cases? Suppose that the hearer knows nothing else about Monty or Prof. Darby beyond that given in the initial sentence, i.e., the hearer knows that there is some person named Monty (presumably male, given the name), that there is a person who is a professor named Darby (presumably male in our culture!), and that Darby is giving information to Monty about some physics experiment (just how is significant), the experiment being marked by the definite article as known to the hearer. The question is: Who knows a lot about high energy physics? In (4) and (6), Darby has a particular role in information giving--he is characterized by virtue of the action of imparting knowledge to someone less informed, and (2) may be viewed as telling about his knowledge. In (5) Darby has a helping relation to Monty, but that seems to weak to indicate that Darby is the energy physics expert since helpers do not have to know as much as the ones they help. However, Darby has a particular position which indicates he knows things--he is a professor. This fact figured into two informants' choices of Darby as the antecedent. For (3) the act of telling indicates that the teller has some knowledge to impart, but it also indicates that the person being told has some means for understanding what is being told. So for (3), speakers sometimes are unable to judge which of the two people knows about high energy physics.

A computational system which makes such judgments must have a very rich knowledge base about the world (to know that Monty is a male name, that professors are by default male, that professors are experts, that neutron experiments are physics

experiments) and be able to draw inferences associated with actions which can be performed. None of this is surprising, as it has been discussed by artificial intelligence researchers before. However, a computational framework for carrying out such judgments is still beyond the state of the art.

While the knowledge and computations needed to weigh the evidence between two or more choices is not yet understood, the actor focus/potential actor rules have the advantage of pointing out the ambiguous situation, of reducing the number of candidates which need to be considered to two, and of suggesting a preferred candidate if supporting information is found. In previous AI systems, such as Charniak's, if the demon mechanism could be designed to discover these ambiguities, it would have no method for choosing among the alternatives. The rule for these cases may be stated as follows:

> POTENTIAL ACTOR AMBIGUITY CONDITION:
> Whenever a pronoun may co-specify the actor focus and a single potential actor exists, expect a possible ambiguity. To resolve, (1) look for evidence supporting the statement in which the pronoun occurs, evidence which is true of the actor focus as the co-specification, but not of the potential actor. If this is found, the actor focus is the co-specification. (2) However, if there is evidence true for both, choose the actor focus but indicate ambiguity. (3) Choose the potential actor when evidence exists for it but not for the focus.

The matter of weighing evidence to decide between two candidates is similar to the semantic choice mechanism postulated by Marcus [1977] for parsing certain kinds of structures such as prepositional phrases. This device, when used for prepositional phrase attachment, demands a judgment from the semantic process of the parsing system which indicates which clause the phrase is more naturally associated with.

The resolution of plurals such as *they* for actor foci is more complex. Not only must syntactic anaphoric filters be met, but also *they* may co-specify with something other than a conjunctive noun phrase such as *Gracie and George*. It may co-specify with several different actors mentioned in different sentences. Suppose D9 were augmented to include D9-4:

D9-4 They don't know him like I do.

*They* co-specifies with both Carl and Oscar together although the two are never mentioned together. The different actor foci are useful here because the set of them constitutes the co-specifications for plural pronouns. Sometimes the discourse focus provides a set relationship, and this set is preferred over the actor focus. In D11 below, Howard's secretary arranged the meeting but is not assumed to be in attendence. *They* co-specifies with the members of some set implicitly related to the focus of meeting, i.e., the set of participants which is Mortie and Moe.

D11-1 Howard's secretary arranged a meeting for Mortie and Moe.
2 I hear that she told them about the new secret ray gun.

Sufficient argument has been presented to allow for the actor focus to be incorporated in the focussing algorithm. The addition to the algorithm is given below. Given an actor focus, anaphora in agent position are not used to retain or to move the discourse focus, as long as other anaphora occur in the discourse.

> 10. The Actor focus is the agent in the current sentence (and its specification), if one exists, otherwise, the actor focus remains unchanged. If the actor focus takes on a new specification, the old actor focus is stacked in the actor focus stack.

Before the rules for discourse and actor foci may be shown by example, the role of syntactic constraints on anaphora must be considered further. This is undertaken in the next section. Following that discussion, an example of the use of actor and discourse foci will be given.

## 4.4 Focus and Sentence Syntactic Restrictions

The animate discourse focus rule must be ordered after rules which compute sentential disjoint anaphora, such as the c-command rules of Reinhart discussed in chapter 2. A computation of Reinhart's c-command relation will predict which phrases within a sentence are disjoint in reference; the possibly co-specifying phrases within a sentence may also be computed at this time. However, a rule which predicts co-specification within a sentence must follow use of the animate discourse focus rule. The sentential co-specification rules would predict that *he* in D9-3 shown previously possibly co-specifies with Oscar. However, by using the focus rule first, this possibility will be ignored since *he* will be found to co-specify Jeff. Similarly, ordering the focus

rule before the sentential co-specification rule predicts the correct co-specification of *he* in D12-2 below. This example is interesting because the Reinhart or Lasnik formulations fail to indicate that *the man* must be disjoint in reference from the pronoun.

D12-1 Whitimore isn't such a good thief.
2 The man whose watch he stole called the police.
3 They caught him.

What is the nature of disjoint reference rules? First these rules must be considered from a theoretical linguistic viewpoint, after which their computational formulation may be described. The syntactic rules specified by Lasnik and Reinhart are stated in terms of disjoint reference. The reason for this formulation is that it is well known that rules for predicting antecedence involve (in some way) semantic as well as pragmatic criteria. In an attempt to keep grammar rules from having to explain all these phenomena (an approach that seems implausible), one may postulate syntactic rules governing when a pronoun may replace a deleted full noun phrase and when it may not. From the viewpoint of Chomsky's [1965] theory, the principle of recoverability of deletion[1] assures that the antecedent of a pronoun can be recovered. What is needed then are rules that stipulate which noun phrases cannot be possible sources of recovery for the full noun phrase.[2] Once these rules are established, it is assumed that the choice of the antecedent involves some set of filters (semantic and pragmatic) which will choose among the possible antecedent noun phrases. From this view, focus is considered as one of these filters.

However, D12-2 cannot be accounted for on the basis of the syntactic disjoint reference rule. The Lasnik and Reinhart formulations use phrase structure trees which

---

1. This principle states that any transformation that deletes information must not delete information which cannot be recovered from elsewhere in the sentence. This principle constrains the kind of theory given by transformational theory. It prevents the theorist from proposing an underlying structure with any sort of extra data that "disappears" in the transformation process. Such a constraint is needed if deletion is allowed since without it, many special notations, which later get deleted, could be hidden in the underlying structures.
2. This explanation is essentially Lasnik's [1976] argument for rules of disjoint reference.

do not appear to capture the disjoint reference in D12-2. Because of this failure, one might want to argue that sentential rules are unnecessary since focus predicts the proper co-specification. However, a closer look at D12 will reveal a difficulty in the behavior of the focussing algorithm: how does the algorithm proceed when two definite anaphora both may co-specify the focus? As chapter 3 indicates, a defnp that is more general than the focus may be used to co-specify the focus. Thus in the alternative D12, *the guy* co-specifies with Whitimore.

D13-1 Whitimore isn't such a good thief.
2 The guy accidentally called the police while he was stealing some jewelry.

How then does one decide that *he* and *the man* in D12-2 do not both co-specify Whitimore? Unless there is a rule of disjoint reference, there is no principle on which to decide. One is forced to argue for some general inference techniques to decide in each case. What is needed is some criterion stipulating when, if at all, a focus cannot co-specify both *he* and *the man* in cases such as D12. Such a criterion is stipulated in Chomsky [1976]. Since his formulation also rectifies the incorrect prediction in the Lasnik-Reinhart rules, it will be presented in some detail. Even with such a criterion, one must choose which of the two phrases should be tested for co-specification by the focussing mechanisms. For that, a principle of discourse will be used: Choose a pronoun to co-specify with the focus before the choice of defnp. Because the pronoun contains less information than the defnp, it seems more likely that it is used to co-specify with the focus.

In order for *he* to be discovered as disjoint in reference with *the man* in D12-2, the logical form of D12-2 given below must be considered.

(7) The man x (such that he stole x's watch) called the police.

Logical form, Chomsky argues, is produced from surface structure by semantic interpretation rules, which include disjoint reference principles, replacement of *who* by its meaning, conventions on control and variable binding, and conditions on anaphora. In linguistic terminology, a pronoun is said to "co-index" with its antecedent.[1] It is a

---

1. The use of co-index is analogous to the use of co-specify but assumes less about the structure of knowledge than has been associated with "co-specification" in this report. To be consistent with the linguistic account, "co-index" will be used.

basic principle of anaphora in logical form which blocks co-indexing between x and *he*:

A variable cannot be the antecedent of a pronoun to its left.

This formulation of co-indexing for anaphora seems to be highly special purpose and entirely non-intuitive. For example, other pronouns within the scope of a variable are considered bound. In (8), the *he* is bound to *every man* so that it is interpreted as co-indexing with the variable in the quantified phrase.

(8) Every man should look both ways before he crosses the streets of Boston.

(9) (x) A man x is such that x should look both ways before x crosses the streets of Boston.

By analogy with (9), the pronoun *he* in (7) is within the scope of the variable x, so the pronoun should co-index with the variable. The anaphora principle is sensible for distinguishing this case from D12 only if it can be motivated. Even though the motivation is not given in Chomsky [1976], it will be offered here. Suppose that logical form and phrase structure trees are not totally unrelated structures. Among other constraints, disjoint classes of phrases created by c-command or precede and kommand rules are retained in logical form. Then the anaphora principle may be viewed as another means of viewing disjoint reference in logical form. In (7), x must be disjoint in reference with *he* since *he* precedes and kommands x, and x is not a pronoun while in (9), the disjoint reference condition does not hold. The use of logical form and syntactic constraints together provides the needed criteria for disjoint reference. Since both logical form and syntactic constraints such as c-command are needed, adequate sentential rules of disjoint reference require that syntactic constraints hold in all levels of sentence interpretation. Thus logical form is constrained by the surface syntactic form of a sentence. Such constraints are implicit in the discussion of logical form in Chomsky, but never stated explicitly.

On the computational side of disjoint reference rules, the use of Chomsky's formulation of logical form and anaphora principles suggests that interpretation of a sentence must include disjoint reference rules. In particular, the focussing rules for anaphora use the disjoint reference information that is computed during syntactic and semantic processing of the sentence. Another kind of disjoint reference information which is needed by the focussing rules is given by reflexivization rules. Such rules not only indicate that *John* must co-index with *himself* in "John washed himself," but also

indicate that in "John washed him", *John* and *him* cannot be co-indexed on the basis of syntactic constraints. It is reasonable to expect the focussing rules for anaphora to rely on processing of a sentence for disjoint reference classes because the criteria for disjointness are syntactic and semantic information "local" to the sentence. Such criteria do not rely on information outside the sentence, so the sentence processing may proceed without requiring information of the focus mechanism or some other contextually sensitive process. Furthermore, reliance of focussing rules on sentence syntactic and semantic criteria modularizes the kinds of information which must be utilized in interpreting the comprehension of anaphora. In summary, the syntactic criteria checks in the focussing rules include not only gender, number and person, but also contraints on disjoint reference from sentence syntactic and semantic criteria.

## 4.5 Examples of the Focussing Algorithm for Actor and Discourse Focus

To understand the focussing algorithm for actor focus and discourse focus, and the use of the focussing rules, the performance of the algorithm and the rules on two examples will be shown. The first, D14 below, is really two different examples. It consists of an initial sentence followed by either of the sentences D14-2a or D14-2b. These pairs are presented because the co-specifications of *he* and *him* in each of the second sentences are reversed. The behavior of the focussing algorithm in these pairs indicates clearly how focussing proceeds when both actor and discourse foci are present.

D14-1 I took my dog to the vet yesterday.
2 (a) He bit him in the hand.
(b) He injected him with a new medicine.

According to the expected focus algorithm of chapter 2, the expected focus of D14-1 is *my dog* and its database specification. Figure 4 shows the state of the focussing mechanism when D14-2a is processed. The members of the alternate focus lists are presented with their database representations only in figure 4 so that later figures may be smaller in size. First the pronoun rules are applied, and then their results determine how the focussing algorithm proceeds. Figure 4 shows that the co-specification of *he* is the potential actor list (PAF) member, *my dog*. This co-specification results from the prediction following decision 6a in figure 1. The current actor focus (CAF) of *I* corresponds to AF in the flow diagram. It fails to be acceptable on syntactic anaphoric filters, while the potential actor focus *my dog* is acceptable. The co-specification of the non-agent third person pronoun *him* is determined by decision 3a since recency does not

**figure 4.4. Anaphora Co-specifications for D14-1 and 2a**

CDF: *my dog* --> Dog201
owner: speaker310
specification: Woofie
syntax: np102

CAF: *I* --> speaker310

PDF: *the vet* --> *vet202*
*specification: ?*

*yesterday* --> yesterday203
is-a time

*I*

*took* --> took204
is-a: take
agent: *I* interpreted as speaker310
object: *my dog* interpreted as dog201
recipient: *the vet* intepreted as vet202
when: *yesterday* interpreted as yesterday203

PAF: *my dog, the vet*

Sentence: He bit him in the hand.

Anaphora: *he*: By decision 6a of agent rules, co-specifies with *my dog*
*him*: By decision 3a of non-agent rules, co-specifies with *the vet*

---

apply. The current discourse focus (CDF) is unacceptable as co-specification of *him* since dogs cannot be bitten in the hand. The first member of the potential discourse focus list (PDF) is acceptable; the co-specification of *him* is therefore *the vet.*

The use of the pronoun focussing rules does not indicate how they are implemented. The example above assumes a sequential processing in which one pronoun is resolved and then the next. Suppose, however, that focussing rules are implemented

as concurrent processes with one process for anaphora in agent position and one for anaphora in other positions. Then the inference rejection of *my dog* as discourse focus could be unnecessary if the pronoun co-specification of *he* had already been established as *my dog* by the use of focussing rules. Instead *my dog* could be rejected on the basis of syntactic filters based on rules for reflexives; the alternative is attractive because the decision is much simpler.

The use of the focussing algorithm for focus movement is depicted in figure 5. The expected focus of the speaker's dog is not confirmed as discourse focus because the non-agent anaphora do not co-specify with the expected focus. The first step of the focussing algorithm which applies is step 3. According to it, the focus becomes whichever anaphor is not in agent position; the focus becomes the potential focus member *the vet* by focus movement. The expected focus of *my dog* is stacked. For the actor focus, the speaker is stacked since *he*, the agent of D14-2a, co-specifies with *my dog*.

For sentence pair D14-1 and D14-2b, the initial state of the focussing mechanism is given in figure 6. It is similar to the one shown in figure 4, but the pronoun focussing rules have different results. By prediction at decision 3a, *him* may co-specify with the current focus because a dog can be injected. The current focus of *my dog* will be retained as focus as it is the only non-agent anaphor. Using decision 6a, the co-specification of *he* to the CAF is unacceptable as in the previous case. However, the first member of PAF must be rejected also because dogs do not give injections. The alternate member of PAF is *the vet* which is the co-specification of *he*. When choosing an actor focus, the old actor focus of the speaker is stacked, and *the vet* and its specification becomes the actor focus.

In the final example of this section, co-specifications for one of the pronouns produces incorrect results. This example shows how pronoun use leads the focussing rules astray. However, for most speakers this example is bad English, so it is not surprising that the focussing mechanism is lead astray. The advantage of the rules given thus far is that they predict the oddness of the behavior.

D15-1 Harriet paid $5.98 to Zelda for a pig.
2 Although she thought it was more than she expected,
3 she was glad to have the money.

**figure 4.5. Determining Actor and Discourse Foci of D14-1 and 2a**

CDF: *my dog* --> Dog201
owner: speaker310
specification: Woofie
syntax: np102

CAF: *I* --> speaker310

PDF (also ALFL): *the vet, yesterday, I,* verb phrase of D14-1

PAF: *my dog, the vet*

Sentence: He bit him in the hand.

Co-specification of anaphora: *he* --> *my dog*
*him* --> *the vet*

Processor at Step 5: actions taken:
Discourse focus stack <-- CDF
CDF <-- member of ALFL of *the vet*
Actor stack <-- CAF
CAF <-- member of PAF of *my dog*

---

The expected focus of D15 is $5.98, and the actor is Harriet. By the focussing algorithm, $5.98 is established as focus since *it* in D15-2 may co-specify the focus. By decision 4 of the agent pronoun rules, the actor focus of D15-2 is Harriet and the first use of *she* in D15-2 co-specifies with Harriet; however, since there is a single potential actor from D15-1 (*she* is not able to co-specify with the pig), an actor ambiguity condition holds. This condition accounts for the ambiguity of the second *she* in D15-2. This particular use cannot be easily resolved on the basis of the first two sentences of D15. However, when D15-3 is seen, the actor focus, if it is Harriet, is odd (Harriet does not have the money after the transaction in D15-1). However, since the same pronoun is used for the actor focus as in the previous sentence, this suggests an unreliable pronoun use. The resolution of such cases is not considered to be part of this report; the manner in which people proceed from failure conditions remains a matter of further study. It is significant that the focussing rules predict when such cases occur.

**figure 4.6. Determining the Anaphora Co-specifications in D14-1 and 2b**

CDF: *my dog* --> Dog201
owner: speaker310
specification: Woofie
syntax: np102

CAF: *I* --> speaker310

PDF: *the vet, yesterday, I,* verb phrase of D14-1

PAF: *my dog, the vet*

Sentence: He injected him with a new medicine.

Pronoun focussing rules: Agent *he*: --> *the vet*
Non-agent *him* --> CF of *my dog*

---

## 4.6 Rules for Possessives

The pronoun rules discussed in the last section do hold for first and second person personal possessive pronouns, but not for third person pronouns such as *his, her, its* and *their*. Because the semantic class of agent cannot be recognized for possessives, a separate set of rules is needed which incorporates the use of the discourse and actor foci. As with non-possessives, when the discourse focus is more longstanding than the actor focus, the discourse focus is preferred. Furthermore, recency does not apply to possessive personal pronouns. The rules are given in the flow diagram in figure 7. The conventions for abbreviations in the figure follow those of figures 1, 2, 3.

The focus set and forward co-specification state in the possessive rules are exemplified by the set of examples below. The first shows how focus sets are used, the second two are cases of non-backwards antecedent pronoun conditions, the first being a case of forward co-specification and the second a non-antecedent pronoun. Non-antecedent pronouns are generally unacceptable to most readers; this behavior will be discussed in a later section.

D16-1 Jill went to Spain this year,
2 while Harry went to the Caribbean.
3 Their vacations were cheaper separately than together.

D17-1 If you take away their lollipops,
2 the children will cry.

D18-1 Football is a tough game to play.
2 Their injuries went up drastically this year.

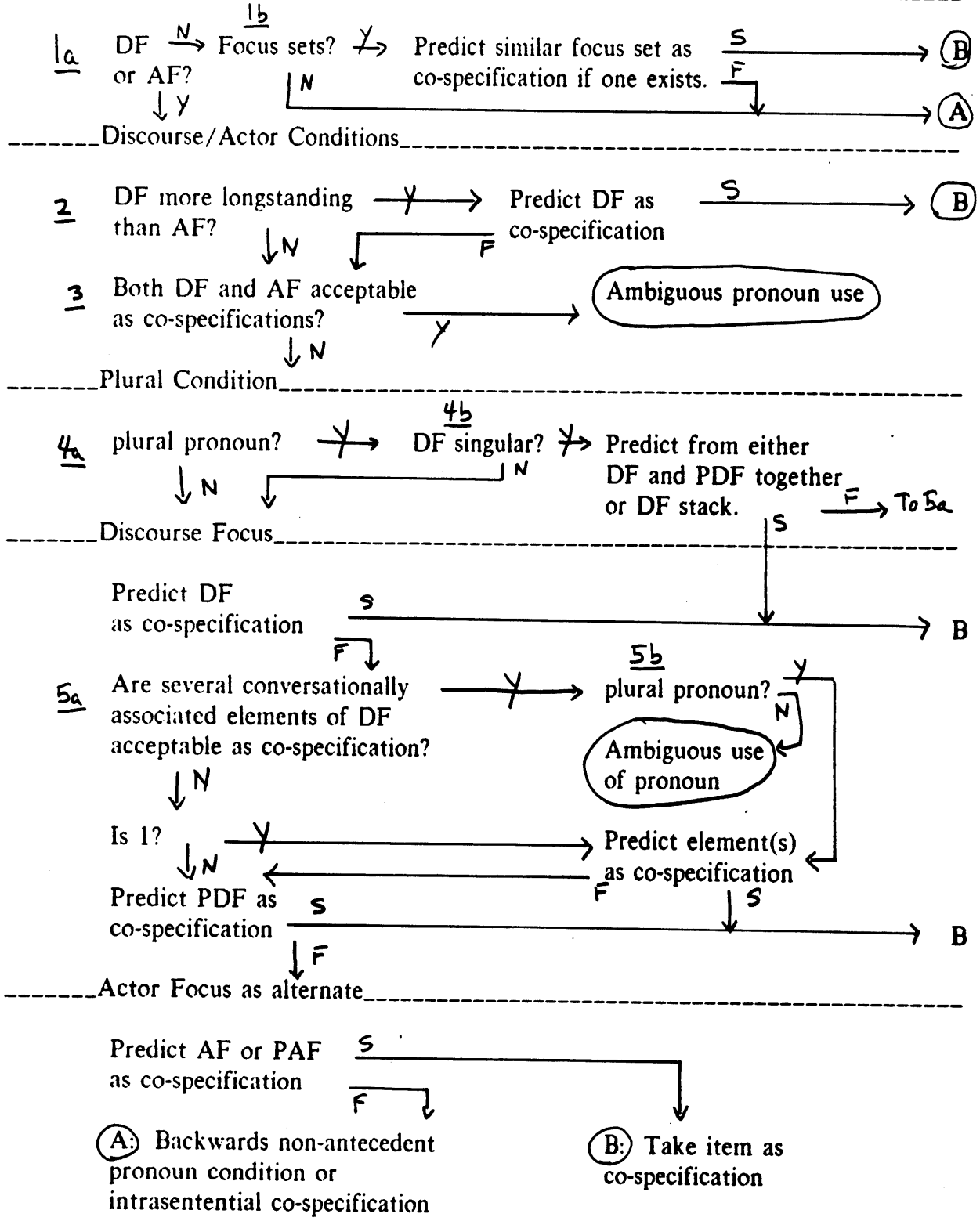
Decisions 5a and 4a are exemplified by the pronoun co-specifications which are illustrated in several examples below. In the first, the conversational associations of the focus set up three classes of possible co-specifiers for *his*, John, Wilbur and the speaker. Of these, the speaker may be ruled out as invalid due to syntactic person disagreement, but because the other two are acceptable, there is no method for determining which is the intended co-specification of *his*.

D19-1 We are going to have a big dinner.
2 John will be the cook, and Wilbur the baker.
3 I will be the wine master.
4 We will eat at his house.

In the example below, the discourse focus is toys while the actor focus is Bonnie and

**figure 4.7. Rule for Third Person Personal Possessive Pronouns**

GIVEN: DF -- discourse focus AF -- actor focus
PDF -- potential discourse foci PAF -- potential actor foci



Dickie. The use of the possessive rules given in figure 7 cause decision 4a to be considered where the discourse focus, as well as the potential foci, are inferentially unacceptable. The processor must then pass through decision 6 to the prediction of the actor focus as co-specification. Since the actor focus of Bonnie and Dickie is acceptable as the co-specification of *their*, it is chosen.

D20-1 Bonnie and Dickie got the most amazing collection of toys this Christmas.
2 Unfortunately, now their toys are scattered all over the house,
3 and it's dangerous to walk around.

The final example shows the similarity between the rules for possessives and non-possessives for pronouns which may co-specify sententially as well as intersententially. Compare D21 with D12 discussed previously.

D21-1 Bill has a real problem.
2 Shem loves his sister,
3 but the woman is in love with someone else.
4 Bill is going to have to tell Shem, who will be heart broken.

Suppose that D21-2 occurred in isolation or as the first sentence of a discourse. The use of *his* would co-specify with Shem. However, in the context of the discourse D21, *his* co-specifies with the actor focus. The actor focus for the discourse is reset after each sentence is interpreted, so at the time pronoun rules are applied to *his*, the actor focus is Bill, just as native speakers would use the language.

D21 is a useful example because it illustrates how speakers might confuse pronoun co-specifications. Suppose the speaker wanted *his* to co-specify with Shem. That is, the speaker, intending for Shem to be the actor focus, could have meant that Shem loved his own sister. To be perfectly clear, the word *own* would need to be present. Of course, often speakers do not make such careful distinctions, and their discourses are correspondingly confusing, with the hearer's actor focus differening from that of the speaker. It is an advantage of the focussing rules that these cases can be explained.

The pronoun rules presented so far assumed a simple representation of the phrases in focus. The focus is a piece of the hierarchical network of associations, where the associations actually mentioned in the discourse are distinguished from those which have not been mentioned. The next section will indicate what further representations

are needed to use focussing pronoun rules effectively.

## 4.7 More About the Representation of Focus

In chapters 1 and 2 the problem of scoping of anaphora was explored briefly. Chapter 1 demonstrates that scoping must be considered in order to interpret certain anaphoric expressions while chapter 2 demonstrates that the representation of focus must include the syntax and semantics of the anaphoric expression if examples such as D22, repeated below, are to be interpreted. This section will substantiate the claim made in chapter 1 that focus together with a representation such as Webber's [1978] provides an interpretation of pragmatic anaphora. However, a brief digression is in order first.

Van Lehn [1978] presents extensive evidence for the view that people do not disambiguate the scope of quantification ambiguities during normal sentence understanding for sentences such as D22-1.[1]

D22-1 Wendy gave each girl Bruce knows a crayon.
2 She used it to draw a Christmas card for her mother.

He reviews the major theories of scoping phenomena and shows how each accounts for the data; he also presents examples which each theory fails to explain. He concludes that disambiguation of quantifier scope is not a linguistic process, and that the correlations of quantifier scope with surface structure are the result of other linguistic processing. In the discussion which follows, focussing does not rely on scope of quantification judgments; instead it seems to require only that some means be given to represent the ambiguities. For cases of bound anaphora, a slightly stronger assumption is made, namely that syntactic rules govern whether a pronoun must be disjoint in reference from a quantified noun phrase.

To explain the comprehension of D22 above, some underlying semantic form representing scope is needed at some point in the processing of the discourse. The interesting questions are why it is needed and when. Webber [1978] argues carefully for why a representation of scope is needed. Furthermore, focussing rules require such a

---

1. Some informants find this example bad English because they cannot decide whether *she* is Wendy or someone else.

representation, for example, to choose the proper co-specifications of *she* and *it* in D22-2. Focussing rules indicate that such a representation is needed not in the initial determination of expected focus, but in the process of confirming the actor and discourse foci. This use of scope information is compatible with Van Lehn's findings because ambiguities are not considered until additional discourse material beyond the single sentence is presented.

*A crayon* in D22 has two possible interpretations, given by Webber in the following formal representations:

(R1) iz: **maxset** (lambda (u:C) [(Ex $e_3$) . Gave $e_2$,x,u & evoke $S_1$,u])z
where C is the crayon predicate
i is the iota operator
$e_2$ is Wendy
$e_3$ is the girls Bruce knows
$S_1$ is the sentence

(R1) is interpreted as the set of crayons, each of which is associated with D22-1 such that Wendy gave it to one of those girls Bruce knows.

(R2) ($e_4$) iy: Cy & Gave $e_2$,$e_3$',y & evoke $S_1$,y
where $e_3$' is a prototype of the form x:G (G is the girl predicate)

(R2) is interpreted as the crayon mentioned in D22-1 which Wendy gave to the prototypic girl Bruce knows.

For some readers there is another interpretation for *a crayon* which says the crayon was given by Wendy to the set of girls Bruce knows:

(R3) iy: Cy & Gave $e_2$,$e_3$,y & evoke $S_1$,y

Suppose now that D22-1 is vague,[1] and no processing of it chooses between the

---

1. A question, often raised by Martin [in preparation], is whether a sentence which is ambiguous between several readings must have several different structure representations. He offers a semantic representation which preserves ambiguity until some processor demands a refinement. Whether this approach or an alternative representation containing both readings is best is an open question in need of further research.

two readings (R1) and (R2) for *a crayon*. When focussing rules settle the co-specification of *it* in D22-2 with the discourse focus, both readings must be available. The set reading may be eliminated immediately because of syntactic constraints on co-specification, so (R2) is left, forcing a reading for *each girl* as the prototypic girl. If however, a speaker accepts (R3) as an alternative reading, then D22 is ambiguous since both one crayon (R3) and a prototypic crayon (R2) are available to focussing.

To determine the actor focus specification for *she* in D22-2, the ambiguity between Wendy and *each girl* must be settled since this is a case of potential actor ambiguity. Actually the interpretation is three ways ambiguous; there is Wendy, *each girl* interpreted as a set as in (R1), or as a prototype as in (R2). The set reading may be eliminated immediately because *she* is singular. To choose between the remaining two, an inference mechanism must decide which is more sensible as in the case of Darby and Monty. The decisions are subtle, so to choose one of these, the inference machine may need to know the co-specification of *it*, which implies that the ambiguity on the readings for *a crayon* must be resolved as well. The interactive behavior for D22 between the rules governing anaphora in agent position and the rules for other positions is similiar to those of D14 given in a previous section.

A case of semantic ambiguity similar to the one in D22 is illustrated in D23 below.

D23-1 Sally wanted to buy a vegomatic.
2 She had seen it advertised on TV.

*A vegomatic* may be interpreted specifically (some particular vegomatic which Sally bought) or non-specifically (one of the many vegomatics). The focus does not distinguish between the two after D23-1 because, like D22-1, D23-1 is ambiguous, and neither interpretation can be chosen with certainty. When *it* is resolved for co-specification in D23-2, the inference mechanism must decide that Sally does not want the very one advertised on TV, but one like it, i.e., *a vegomatic* is a non-specific use. The value of focus in this case is that focus immediately defines the co-specification problem as one of choosing between interpretations of the focussed phrase.

The discussion of representation above does not account for certain uses which remain problematic. Consider the case shown in D24.

D24-1 Sally bought a vegomatic which had a broken cutting blade.
2 She had seen it advertised on TV.

*A vegomatic which had a broken cutting blade* is usually interpreted specifically as some particular vegomatic which Sally bought. However, *it* is ambiguous between the vegomatic Sally bought and one like it, that is, between the extensional and intensional readings shown in figure 8. Note that a characteristic instance is not the same as a prototypic instance since prototypes never exist as real objects while a characteristic instance is a real entity used in a prototypic way.

Focussing does not distinguish between the two readings since it focusses on the extensional one. When *it* is resolved for co-specification in D24-2, the inference mechanism must discover that it is slightly odd for Sally to have bought the very vegomatic she saw advertised on TV. Then if no other choices for co-specification are available, an intensional reading must be chosen. This example is problematic because it places much weight on the inference machine to decide that the reading is odd. However, this is likely to be just where the weight of the decision ought to be; many native speakers find D24 slightly bizarre because their first reading is that Sally had seen the vegomatic with the broken blade advertised on TV. In fact, it appears that when a specific indefinite noun phrase such as *a vegomatic* is introduced, and the speaker wants to turn attention to the intensional reading, a plural pronoun is used as shown below:

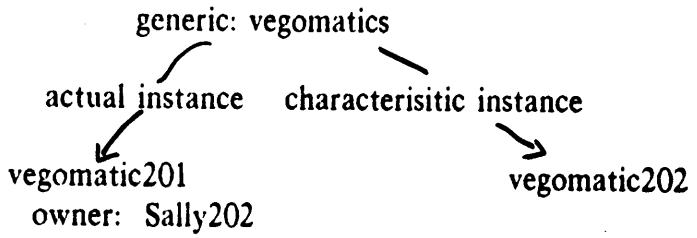
(10) She had seen them advertised on TV.

The use of the intensional plural is not covered in the pronoun rules.

Examples such as D24 are perplexing for another reason: they are examples of

---

**figure 4.8. Intensional and Extensional Specifications of *it* in D24-2**




EXTENSIONAL reading of *it* INTENSIONAL reading of *it*

what I will call, following Fahlman [1978], the copy phenomena. The ambiguity centers around the fact that there are many copies of an abstract prototype. Automobiles, computer programs, airplane flights and money are other common cases of entities which result in copy phenomena. In D25, the interpretation of *it* depends on whether the speaker is referring to a particular flight or the normal Sunday flight, a copy of which occurred on *this Sunday.*

D25-1 TWA 384 was so bumpy this Sunday I almost got sick.
2 It usually is a very smooth flight.

Note that the *it* cannot co-specify with the particular flight[1] on *this Sunday.* However, it is possible that the speaker intended TWA384 to refer to a particular flight, and so there is a mixing of interpretations for the co-specification of *it.* This should not be displeasing to the reader since in general speakers seem confused about whether they are talking about the prototype or a particular copy of it.

Other characteristic uses of anaphora are the bound variable cases given by Partee [1972, 1978]. In D26 below, *him* co-specifies with Archibald.

D26-1 Archibald sat down on the floor.
2 Every man put a screen in front of him.

If bound variable cases provide syntactic and semantic disjoint reference conditions, then focussing accounts for D26 since *him* must be disjoint in reference from *every man,* and by the focussing algorithm, *him* may co-specify Archibald, the expected focus. However, for bound anaphora, syntactic and semantic rules must also indicate sentence anaphora which must be bound by the quantifier. In D27-3 below, the *he* is bound to *every child* so that the reading of the sentence is, roughly, "every child is such that he will insist about himself that he is not sleepy." Without some rules that indicate that *he* in D27-3 is bound, the focussing rules will predict that the co-specification of *he* is Billie by step 3 of the agent rules.

---

1. Intuitively it is odd to speak of a particular flight as always having some property when it only in fact occurs once, so it is natural to expect the co-specification not to mix generic and specific readings of TWA 384.

D27-1 Billie is a typical eight year old.
2 He likes to stay up past his bedtime.
3 Since every child will insist he's not sleepy,
4 Billie will try to stay up late
5 even though he's ready to fall over with exhaustion.

Then the inference mechanism would have to reject this choice. To do so, it must conclude something like "it is odd for every child to insist anything to be true of a particular child." Then the bound variable reading must be confirmed with evidence like "it is okay for every child to say of himself that some characteristic is true." While such behavior ought to be possible in principle for the inference mechanism, it should not be necessary because of syntactic structures in English. Therefore, rules for bound anaphora and disjoint reference must be part of the syntactic anaphoric filters.

One other possible explanation could be given for D27. If D27-3 could be interpreted on the basis of tense, or some other indicators, as being generic, a bound variable reading would be forced. Since the interpretation of generics (see Lawler [1972] and C. Smith [1969]) is not yet well understood, this alternative explanation seems doubtful.

## 4.8 Restrictions on Co-specification Using Focus

There are other restrictions on co-specification besides those due to sentence syntax and semantics. In chapter 2 the stacked focus constraint was outlined: since anaphors co-specify with the focus or a potential focus, an anaphor, which is intended to co-specify with a stacked focus, must not be acceptable as a co-specification with either the focus or potential focus. An example from Chaim Potok's [1975] *In the Beginning* of co-specification with a stacked focus is presented below.[1]

> Was that old lady evil, the one Saul and I had seen sitting on the porch? I had dreamed about her. When the trolley car took me and Saul past her house again this morning, she was gone. Evil, it had a queer sound to it in English.
>
> {Here the narrative moves on to the speaker and an incident in a school classroom.

---

1. "In the Beginning" by Chaim Potok, page 212, chapter 4, Fawcett Publications, Inc. Conn, 1975.

A discussion between the speaker and a male teacher ensues for five paragraphs. The following paragraph begins:}

She had worn an old brown coat and a green scarf over her head.

In this example, *she* co-specifies with the old lady discussed previously. The stacked focus constraint is needed to reduce interference between a pronoun and other material before the antecedent. If Potok had told of a discussion between the speaker and a female teacher, it would no longer be possible to tell that *she* was co-specifying with the old woman. The reading of *she* as teacher might be a bit surprising because what the teacher is wearing was not relevant to the previous conversation, but it certainly is not the case that an inference mechanism would decide that teachers do not wear old brown coats and so forth. The stacked focus constraint restricts co-specification interpretation in extended discourse.

The stacked focus constraint is not stated directly within the focussing rules. Instead it is implicit in the function of them. When a current focus is rejected in favor of a stacked focus, if the anaphor causing rejection was acceptable as a co-specification of the current focus, the rejection could not have occurred. The stacked focus constraint is a consequence of the movement of focus in focussing.

The stacked focus constraint, however, may be overridden. An astonishing set of examples was identified by Grosz (see Deutsch [1974], Deutsch [1975]) in protocols of two person dialogues. Two typical cases are given below. Each contains a pronoun co-specifying with the stacked focus when intervening material contains possible co-specifications. These examples are taken from a dialogue between two people, one of whom is assembling an air compressor while the other one gives advice.

Case 1:
A: Bolt the pump to the base plate. There are 4 bolts, 4 nuts and 4 washers. <here follows an explanation of where to put the bolts and what tools to use.>

B: I would like to know if I can take off the back plate.

A: You shouldn't have to. Are you having trouble with the bolts?

B: Yes

A: <Now follows a long discussion of the use of the ratchet wrench, the extension and the socket for the wrench. The discussion ends with:> You will use the 2" extension and a 1/2" socket.

B: It is bolted. Now what should I do?

Case 2:
A: One of the bolts is stuck and I'm trying to use both the pliers and the wrench to get it unstuck.

E: Don't use the pliers. Show me what you are doing. Show me the 1/2" combination wrench.

A: Ok

E: Show me the 1/2" box wrench

A: I already got it loosened.

In most of these cases, the intervening possible co-specifications could be ruled out on the basis of some complicated set of inferences about what can and cannot be bolted to what, or loosened. However, it does not appear that enough delay in understanding occurs even for people who read this excerpt (the original speaker and hearer were not being tested for delay) to suggest that they are ruling out multiple possibilities.[1] Hence one may conclude that another process is helping drive to the understanding of what is meant. In particular it appears that understanding depends on knowing that something being bolted completes the task which A originally specifies in case 1. The focus of A's first command indicates exactly what is bolted. Here it appears that focus provides the co-specification for the object under discussion, but some other mechanism, which interprets completion of task goals, indicates where to pop back in the focus stack.

How many such discourse interpretation mechanisms exist? While this report does not address this question directly, some speculation is possible on the basis of research which will be reported on in chapter 6. In general, it appears that discourses which permit violation of the stacked focus constraint must contain an implicit structure

---

1. This informal evidence needs to be tested out in a psychological laboratory. The author has not done so but the results of such experimentation would be revealing.

of task completion or contain some other well specified structure which guides the hearer in understanding. Without this structure the hearer has no means for deciding that the speaker intended for the pronoun to co-specify with something other than the object currently under discussion.

## 4.9 Pronouns Which Have No Co-specifications

The focussing algorithm as stated in chapter 2 contains a condition called the non-antecedent pronoun condition. This section will describe how that condition occurs and what difficulties it poses in understanding definite anaphora.

There are many uses of pronouns where the pronoun has no co-specification in the preceding discourse, where it is not used to refer forward, and where it is not used in conjunction with some action such as pointing. Several cases are given below; the first three are from Postal [1969], the fourth from Chafe [1975], the fifth from dialogues collected for PAL, a system discussed in chapter 6, and the last was spoken by a lecturer at a presentation this author attended.

D28-1 I saw Mr. Smith the other day; you know, she died last year.
2 John is an orphan. He misses them very much.
3 Pro-Castro people don't believe he is a monster.
4 I went to a concert last night. They played Beethoven's ninth.
5 I want to meet with Bruce next week. Please arrange it for us.
6 I used to be quite a tennis player. Now when I get together with the young guys to play, I can hardly get it over the net.

With the exception of D28-1, most speaker-hearers are able to say which is the intended co-specification of the pronoun in the cases above. D28-1 can be understood if the hearer is informed that Mr. Smith had a wife. However, some of these, especially 1 and 2, are so odd for most hearers that they first indicate a lack of comprehension of the pronoun use. Hearers are divided on the acceptability of 3, and most hearers find 4 and 5 acceptable. Furthermore, such examples, as far as I can tell, do not occur at all in written samples.

This report will not give an account of just how such cases are understood. However, focussing seems to provide an indication of how to handle such situations. In all the multi-sentence cases, the pronoun co-specifies something which is closely associated with the focus. What is problematic is the fact that some speakers do not

permit co-specification with elements which no longer exist, such as John's parents in light of John's orphanhood. Whatever the manner in which hearers recover co-specifications for such pronouns, it ought to be that if focus is used, it ought to account for why some are acceptable and others not.

These examples do suggest that the focus mechanism must distinguish between entities associated with the focus which are mentioned during a discourse (and therefore are available for co-specification) and those which are associated with the focus in a long term way as part of the database, but not mentioned in the current conversation. This constraint may seem odd at first glance, but it can be motivated. Human communication (and perhaps one day, machine communication) proceeds in the context of a very rich collection of memories and general knowledge. There is no reason to assume that the memories and knowledge in two people are similar enough for the speaker to know that the hearer will identify the same memory or knowledge. Furthermore because the memories and knowledge are rich in detail, the speaker cannot be certain that the hearer will attend to the right detail when there are several to choose from. Instead the speaker must indicate which detail is of interest. Speakers do this by means of language, and though not explained here, by gestures, glances and other non-linguistic sounds. Accordingly, focus must indicate which part of the structure has been attended to during the conversation, and which part is potentially of interest but which has not been pointed out by the speaker.

In computational terms the above requirement stipulates that data structures or programs which represent concepts must distinguish between information previously established and information resulting from the conversation itself. One possible implementation, suggested by FRL (see Roberts and Goldstein [1977]), is the use of a comment on a frame value; in semantic networks the use of an extra link also comes to mind. Yet neither of these is quite correct. Both allow for the focus mechanism to have access to this necessary information at some later point in processing. Since the motivation for focus has been to reduce the search for co-specification, the representation must be sectioned off into those parts resulting from conversation and those resulting from long term knowledge. Furthermore, the non-conversational section must be hidden from view until the section representing the conversation fails to provide some expected co-specification. For example, in D28-4, the specification for *they* is the

orchestra which performed the concert in focus. However, the focus mechanism must not even notice the orchestra as a possible specification for *they* until the conversational elements are rejected as possible co-specifications. Further evidence for this behavior is shown by D29 below.

D29-1 I went to a concert with some friends last night.
2 They played Beethoven's ninth.

Until the hearer decides[1] that it is a bit odd for the speaker's friends to have been the players, the specification of *they* to orchestra must not be considered.

What is needed to indicate sectioning is partitioning such as that used by Grosz in semantic network implementations of her focus mechanism. She used the partitioning to create a space of objects currently in focus. The above examples show that at least one of those focus spaces (several may be present concurrently) must indicate all the entities in conversation as distinct from the rest of memory. Knowledge representation systems lacking this capability will fail to make proper use of focus mechanisms and may result in incorrect co-specification predictions.

---

1. Native speakers differ on whether *friends* is odd as the antecedent. Differences of opinion seem to indicate when an inference is not knowledge shared by all speakers and hearers.

## 4.10 The Problem of Parallelism

Focussing gives incorrect predictions for certain cases of co-specification which are difficult to define. Intuitively, they may be characterized as instances of parallel structure between sentences of a discourse. In many of these cases focussing predicts the wrong co-specification. To understand what is meant by parallel structure, two simple cases, one in which focus does predict correctly, and another in which it fails, will be discussed. In D30, the pronoun co-specifies with the mud pack, as focus would predict. The parallelism of these sentences is reflected in the semantics of *put on* and *pull off* as well as in the similarity of the syntactic structure of the two sentences, each containing an imperative voice main clause.

D30-1 Put the mud pack on your face.
2 After 5 minutes, pull it off.

Focussing rules predict the proper co-specification in D30 because the thematic relations of the verb follow the similarity of structure. In D31, the pronoun *it* co-specifies with rose and not with the green Whitierleaf. The expected focus after the first sentence is Whitierleaf, but the parallel structure of the sentences seems to govern the choice of co-specification. To summarize, between similarity of structure and focus, similarity is preferred as a means of choosing co-specification, so when each gives a different prediction, the similarity of structure must be used.

D31-1 The green Whitierleaf is most commonly found near the wild rose.
2 The wild violet is found near it too.

On first glance it appears that the focussing rules could be "fixed" by simply remarking that the expected focus is wrong and that a potential focus should be chosen. No such option is available, for such a fix requires that the inference machine reject the expected focus. To do so, the inference mechanism needs some knowledge about the world that indicates the unacceptability. For D31 no such knowledge could possibly be forthcoming since all the flora involved are found near one another. There is no knowledge which states that violets are found near wild roses and not near

Whitierleafs.[1]

Another example of parallel structures is shown in D32. The parallel structures again are reflected in the similarity of the syntactic forms as well as the semantics of *most* and *mine*. After D32-1, the expected focus is the radiator on most cars (that is the prototypic radiator). Using the focussing rules, *it* will be taken to co-specify with that radiator. But this prediction is incorrect; *it* co-specifies with the radiator of the speaker's car.[2]

D32-1 On most cars the radiator has a free bolt hook.
2 But on mine, it has a floating bolt hook.

*It* is similar to intensional use of *it* in D24. What makes it different is that D32-2 has an underlying semantic form which is similar to D32-1. D32-1 specifies a univeral set of cars and says something about one of the parts for those cars; D32-2 specifies a set of one thing, the speaker's car, and says something about a part of it, where the part is related to the universal car by intension. Thus *it* in D32-2 is not pointing to some intension of radiator; it co-specifies with the radiator of the speaker's car, but *it* is related intensionally to the radiator in D32-1. The similarity in the underlying semantics of D32-1 and cars-2 must be used in interpreting the pronoun uses.

One might wish to construct some special purpose mechanism that looks for similarities in structure between two sentences. This method is doomed for two reasons. Parallelism exists in many aspects of language, and it happens at arbitrary levels of structure. Furthermore, at any given level, the problem of recognition of parallelism has plagued computational models of language since such models were first suggested. For example, parsing of English sentences containing conjunction is as yet an unsolved problem. Methods tried, such as those of Woods [1973] in LUNAR, fail because of overgeneralization.

---

1. In certain cases a special audience may have different responses to the parallelism above. For example, botanists who know what flowers are near others might behave differently. But even special audiences must sometimes use general techniques. Such is the case in the D31 example, because Whitierleafs exist only in the imagination of this author.
2. The author thanks Bob Moore for suggesting this example.

The fact that interpretation of parallelism has failed for other aspects of computational models of language should only indicate that the problem is a deep one. Just how deep it affects co-specification is indicated in the example below. Since the example is long, and readers who are not familiar with procedural specifications may not understand the algorithm, analysis of the focus mechanism and of the procedure itself will clarify the example.[1] The reader will note that a "sol" is a schedule, a collection of timing information which contains several different kinds of timing entries.

D33-1 The sol is searched for an entry for the subscriber.
2 If one is found, the subscriber's relative transmission time is computed according to formula-1.
3 The subscriber's clock transmission time is computed according to formula-2.
4 When the transmission time has been computed, it is inserted as the primary entry in a transmission schedule.
5 For each rats entry, the rats' relative transmission time is computed according to formula-1,
6 and the rats' clock transmission time is computed according to formula-2.
7 The rats' transmission times are entered into the schedule.

This example accomplishes the simple task of taking an entry in a schedule with the special name of sol (line 1), performing two computations on the entry (lines 2 and 3), and putting the results in another schedule (line 4). Then the same process is repeated for a set of entries in the sol called rats entries (lines 5 through 7).

Using the focussing algorithm, the focussing mechanism chooses the expected focus as the sol. Since the focussing algorithm says nothing about *one* anaphora, the sol is confirmed as focus.[2] Both the relative transmission time and the clock transmission time are associated with the focus, the sol. The use of *it* in D33-4, moves the focus to *the transmission time.* At line 5 the focus switches to the rats entry since both *the rats relative transmission time* and *the rats clock transmission time* are associated with it. A problem occurs in line 7. What is the co-specification of *the schedule*? There are two

---

1. This example comes from Balzer [1977].
2. If *one* anaphora function like definite anaphora, then the focus of D33-2 would be the entry.

schedules to choose from: the sol, and the transmission schedule mentioned in line 4. By focussing, the sol will be chosen as the co-specification.

However, people who read procedural specifications without foreknowledge of what this particular procedure should do, know that *the schedule* co-specifies with the transmission schedule. They are not using knowledge which tells them that procedures ought to be written in a certain way which eliminates the sol, since there is no such knowledge. Instead the reader is relying on a structure which he or she computes for the example. The structure was stated explicitly two paragraphs back:

> Then the same process is repeated for a set of entries in the sol called rats entries (lines 5 through 7).

Like many procedures, D33 specifies a process and repeats it on several sets of data. *The schedule* is disambiguated by use of the structure of the procedure, that is, take some data out of the sol, perform a set of computations on the data, and put the results in the schedule used for results (the one called a transmission schedule). At a very different level of description, a similar structure, one parallel to the first part of the description, indicates the co-specification of the noun phrase *the schedule.*

An interesting question to ask here is whether most readers know upon reading D33 that it has the structure given above. Because some readers (computer scientists no less) find this example confusing, it may well be that most do not; that they compute it by re-reading the text and looking for a structure which "fits" the example. Whether or not the structure is obvious, focussing does not give an account of the similarity of the two parts of D33.

Why is focussing unable to encode this similarity? Focussing encodes what is being talked about in the discourse, but it gives no account of what sub-parts are. Focussing relies on a sentence by sentence set of connections either to maintain what is being talked about or to indicate the movement to a new entity for discussion. In examples such as D33, what is being talked about is the sol, certain of its entries and the properties of those entries. There is also mention of another schedule, the transmission schedule. The "talking about" involves focus recognition, focus movement and potential focus introduction. The difficulty with focus centers on the notion of

"sub-part." Rather than being a sentence by sentence account of relations, a sub-part represents, intuitively speaking, a relation resulting from the two sets of several sentences, each set organized into some kind of unit. Focussing is meant to capture the antecedent relations among the sentences of a unit; it does not appear to capture the relations between units in the discourse. "Sub-part" is probably only one of several relations among sentence units. Whatever those relations are, they are understood by some means besides focussing.

One possible consequence of this observation could be that the focus mechanism should be abandoned in favor of some as yet unspecified mechanism that is able to determine parallel relations among sets of sentences in a discourse. However, parallelism is not relevant to most of the examples presented in this report. Many cases of co-specification occur where there is no similarity of structure other than the common subject-verb-object pattern typical of English sentences. Since what is being talked about appears in many constituent positions in sentences of a discourse, the s-v-o pattern seems too gross a level to specify similarity. Hence while parallelism is needed to deal with a certain set of cases for which focus gives incorrect predictions, focus is effective for many other cases of co-specification where parallelism would not be helpful. Those cases where parallelism occurs are intuitively well defined although not specifiable within the focussing rules. One may conclude that focus mechanisms account for one aspect of pragmatic anaphora, and that some different mechanism is needed to encode similarities in structure which are used in human languages. This report will not give an account of such a mechanism. Rather, the examples above provide some additional observations about the nature of parallelism in natural languages.

## 4.11 Comparison with Other Recent Computational Algorithms

Hobbs [1976] discusses a syntactic algorithm for anaphoric expressions. For sentential antecedence Hobbs' algorithm uses the intuitions of Lasnik's precede and kommand rules.[1] It has the disadvantage that sentential antecedents are considered first

---

1. Hobbs' algorithm is stated in terms of establishing antecedence rather than disjoint reference. This difference is significant since Hobbs' claim is a stronger one, and as will be shown, less reliable.

so that it chooses Oscar for the co-specification of *he* in D9-3. For intersentential anaphora, the syntactic algorithm works on previous sentences in left to right order. Often this gives the same effect as the focus but there is no theoretical explanation for why this process performs as it does. In fact, the Hobbs' algorithm is accurate in a large number of cases for anaphors occuring in text. In a case study of three different texts, Hobbs reports 88.2% accuracy for the basic algorithm and 91.7% accuracy for a version augmented with selectional restrictions. In the remaining cases, the algorithm predicts the wrong phrase as antecedent (see Hobbs [1977]).

Hobbs wants a more effective algorithm. He chooses one which he calls a "semantic algorithm," consisting of two parts. First each sentence is reduced to its semantic primitives. Then a search is done to determine how those primitives match with the previous sentence expressed in semantic primitive form. To reduce the search size, a bidirectional search is used. For (11), Winograd's example, slightly modified to use only pronouns, the following facts and general inferences are needed.

(11) They prohibited them from demonstrating because they feared violence.

Facts:

(w demonstrate) cause violence

((x cause y) and (z diswant[1] y)) cause (z diswant x)

General inferences:

x prohibits y --> x diswants y

x fears z --> x diswants z

By forward inferencing and unification, it can be proven that the *they* who want fear is the first *they* in (11). A similar technique, but more general and without the need for primitives, is discussed in Isner [1975].

The use of a focus mechanism does not deny the need for the kind of inferencing detailed by Isner and Hobbs. However, focussing makes the instantiations simpler: choose the focus (except for recency cases when potential focus is tried first) as the co-specification of the pronoun in question and look for a proof or a contradiction. Contradictions tell us to drop the focus and test the potential focus in the same way, while correct proofs mean the co-specification is the intended one. Hence focussing

---

1. "Diswant" is used to avoid the negation problems of "not want."

eliminates combinatorial search for antecedents by reducing the matter to one of confirming the co-specification predicted on the basis of focussing. In effect, focussing offers an additional control constraint on Hobbs' semantic algorithm.

Another approach to antecedence can be found in Lockman [1978]. Lockman's purpose is more general since he wants an algorithm to establish "contextual reference relations" between noun phrases which are not anaphoric ones. For example, he wants to determine the relation between *a harp and wings* and *heaven* in:

D34-1 After dying, Algernon went directly to heaven.
2 A harp and wings had already been prepared when he arrived.

Lockman states a premise which is not held by this author, i.e. "it appears that any item appearing in a text can be a referent for some possible following sentence of the text."[1] Lockman suggests that all noun phrases must be guessed at and tested. Thus his algorithm is somewhat like focus in that a choice is made and confirmed or denied, but unlike focus, all noun phrases may be choices. Since the number of choices is potentially large, to reduce the number of choices, Lockman develops the notion of structural context to capture the coherency between sentences in a text. He says:

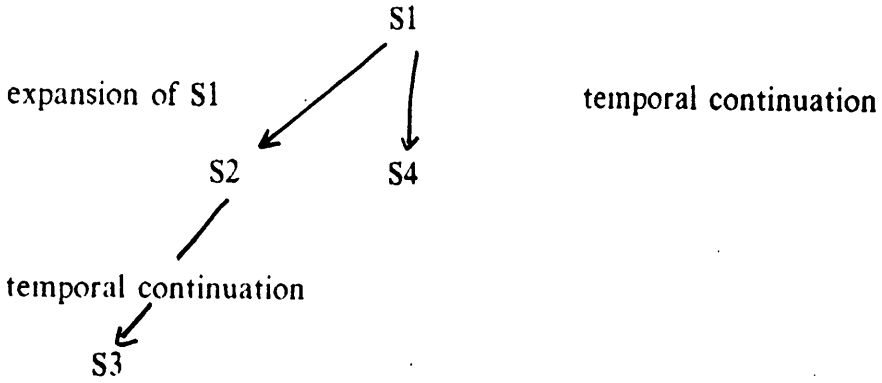
> The default search path for finding which sentence or sentences to fit a new sentence $S_{n+1}$ under (to feasibly connect it to) lies back up the path of expansions and temporal continuations that resulted in $S_{n+1}$'s immediate predecessor $S_n$.[2]

An expansion is defined as a "sentence $S_j$ which gives us further information (i.e. further detail, clarification, manner, cause, etc.) about either the entire conceptualization expressed by $S_i$ [a previous sentence] or about some concept within the meaning of $S_i$."[3] A temporal continuation is a sentence that describes an action that occurs later in time than the action of a previous sentence. Lockman uses these notions to construct a structural tree of sentences in text, as depicted in figure 9. The tree expresses the expansions and continuations of all sentences.

---

1. Lockman, *op. cit.*, page. 30. Note that Lockman mistakenly uses *referent* for something like antecedent or co-specification. It appears that he means antecedent.
2. Lockman, *op. cit.*, page 106.
3. Lockman, *op. cit.*, page 100.

**figure 4.9. Structural tree of Lockman expansions and continuations**

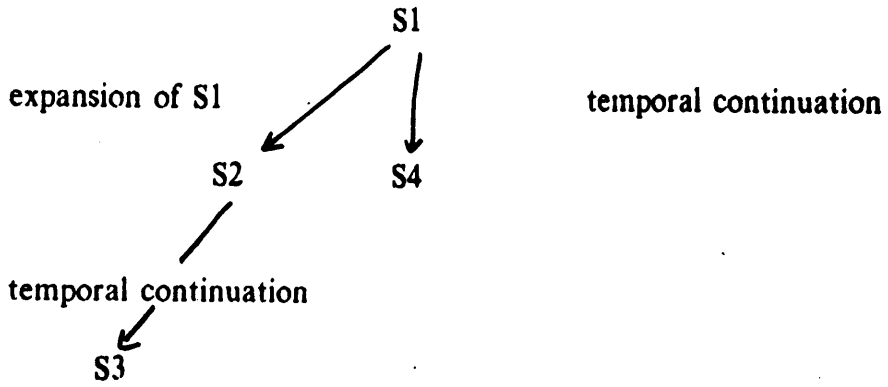



---

Certainly many texts may be described as having sentences which are related by expansion or temporal continuation. I have not studied in detail whether there is any difficulty in determining temporal continuations. On first reflection, it seems that tense information helps make this possible. Using temporal and expansion continuations is one method for determining the structure of discourse. The resulting tree form which Lockman's work produces models a structure which seems implicit in many discourses. Structure must be found in order to handle such problems as parallelism. Recent work by Hobbs [1979] uses coherency patterns to describe several kinds of structures inherent in discourse.

What is implausible about Lockman's approach is how one decides that an expansion had occurred (or failed to occur) unless one has first determined how the anaphoric expressions co-specify with noun phrases preceding in the context. For example, D9-3, given previously, is coherent in the text because *he* is predicted as co-specifying with Jeff and thereby establishing the sentences as saying something more about what is already under discussion, namely that Jeff is sick in Oscar's opinion. A reasonable guess would be Carl, but the resulting interpretation of the sentence is not coherent because of the focus behavior. Why would Oscar say of Carl that he is sick when Jeff is under discussion? The expansion relation must occur after one has decided that the co-specification is reasonable. Deciding on an expansion relation before the co-specification in D9-3 is possible only by guessing, based on all possible people mentioned in the discourse. Focus provides a means for predicting the co-specification

**figure 4.9. Structural tree of Lockman expansions and continuations**




---

Certainly many texts may be described as having sentences which are related by expansion or temporal continuation. I have not studied in detail whether there is any difficulty in determining temporal continuations. On first reflection, it seems that tense information helps make this possible. Using temporal and expansion continuations is one method for determining the structure of discourse. The resulting tree form which Lockman's work produces models a structure which seems implicit in many discourses. Structure must be found in order to handle such problems as parallelism. Recent work by Hobbs [1979] uses coherency patterns to describe several kinds of structures inherent in discourse.

What is implausible about Lockman's approach is how one decides that an expansion had occurred (or failed to occur) unless one has first determined how the anaphoric expressions co-specify with noun phrases preceding in the context. For example, D9-3, given previously, is coherent in the text because *he* is predicted as co-specifying with Jeff and thereby establishing the sentences as saying something more about what is already under discussion, namely that Jeff is sick in Oscar's opinion. A reasonable guess would be Carl, but the resulting interpretation of the sentence is not coherent because of the focus behavior. Why would Oscar say of Carl that he is sick when Jeff is under discussion? The expansion relation must occur after one has decided that the co-specification is reasonable. Deciding on an expansion relation before the co-specification in D9-3 is possible only by guessing, based on all possible people mentioned in the discourse. Focus provides a means for predicting the co-specification

without a general guess and test method on all the noun phrases, and without the additional burden of an even more general guess and search technique concerning continuations of previous sentences.

Lockman's continuations and temporal expansions for contextual structure might constitute a possible method noticing similarity of structure between sub-parts of a discourse. Used in this way, such structures could provide a control on recognizing parallel sub-parts of a discourse. Linked to focussing, these structures could suggest that rather than the focus, some other object is meant to be co-specified by a defnp in cases such as D33. This speculation suggests a point of further research.

## 4.12 Conclusions

Two claims have been substantiated in this chapter. One claim concerns the use of focussing with linguistic rules for sentential disjoint reference and syntactic scope, and with representations delimiting sentential scope information. Both the rules and representations have been shown to be compatible with the process of focussing, and necessary to focusssing in providing information relevant to pronoun co-specification. The item which is focussed on must include an abstract structure which includes scope and disjoint reference information.

The second claim concerns the control of inference used in interpretation of pronouns. Because the focussing algorithm predicts a co-specification and then asks for confirmation of it, with the stipulation that contradictions indicate incorrect predictions, inferencing is controlled by focussing. In previous AI natural language systems co-specification results from binding of free variables during inferencing. The inference process may be characterized as one of proving a consequent from a set of premises. However, the consequent is not entirely well specified since the pronouns in it are made free variables. As a result, many inferences will be made with incorrect choices for the free variables. Focussing eliminates this kind of inference. Focussing asks that a discourse sentence, with the definite anaphora replaced by co-specifications to the discourse and actor foci, be judged for consistency with the preceeding sentences of the discourse.

A part of this chapter has specified and illustrated how the focussing rules

predict co-specification of pronouns with the actor and the discourse focus. Both actor and discourse foci are necessary. In English, definite anaphora signal two classes of phrases which are previously mentioned in the discourse, the element of chief interest, and the agent acting in relation to that element. Focussing models the function of these two signalling behaviors. The rules show how many constraints, syntactic, semantic and phonological, affect the choice of antecedences for pronouns; they also show how many kinds of information about the world of the speaker and hearer play a part in distinguishing the co-specifications of pronouns. The focussing rules, by means of the discourse and actor foci, differentially apply the constraints; the predictions which result are tested in the database representation of the speaker and hearer's world. Focussing captures the effects of foregrounding (cf. Chafe [1977]) since focussing accounts for the co-specification of pronouns by means of the foci, and focus movement indicates how new entities may be foregrounded and pronominalized.

The focus popping cases described by Grosz, and the need for parallelism indicate the role of higher structures in focus interpretation. The Grosz examples violate the stacked focus constraint, by the use of goal structures in the discourse. The parallelism examples show that another kind of structure is also used in discourses. While focus popping makes use of the focus mechanism, the parallel structure cases seem to rely on a mechanism which is different in kind from focussing.

This chapter further specifies the nature of focussing as it applies to pronoun interpretation. Focussing captures one of the relations that exist between sentences of the discourse. It is a sentence by sentence relation which specifies what it being talked about in each sentence, and how elements of the discourse which are talked about are carried forward into the discourse. Pronoun interpretation relies fundamentally on the focus since the pronoun is a signal that what was talked about before is still under discussion.

# 5. Co-present Foci in Anaphora Disambiguation

## 5.1 The Occurrence of Co-present Foci in Discourse

In chapter 2 discourse was defined as being about one central concept. It was pointed out there that sometimes speakers discuss several concepts at once without indicating that they are doing so. Generally this results in a discourse which really is not a discourse, that is, the discourse is confusing enough to keep the hearer from understanding. Sometimes however people discuss more than one thing without confusion. How is it possible to discuss several things at once? One such case has already been presented; in chapter 4, an actor focus was shown to be present in many discourses in addition to the discourse focus. It is also possible to have co-present discourse foci within the discourse focus. Such foci are the subject of this chapter.

This chapter investigates three topics. The first introduces the phenomena which constitute co-present foci. A second topic concerns algorithmic rules which account for these phenomena. The third topic considers the function of the focussing mechanism of chapter 2 in connection with the anaphor rules for co-present foci. The completion of the three topics in this chapter will permit me to support two claims about co-present foci:

> 1. The use of *the one...the other* and *this* and *that* definite anaphora signal the speaker's focus in the discourse.
>
> 2. No changes in the focussing mechanism are needed to incorporate the rules for co-present foci.

In the sections which follow I will investigate the three topics for each of the co-present foci phenomena. Thus the form of the discussion will be to present some phenomena, then some rules and an explanation of the focussing mechanism. At the conclusion, I will present the additional arguments for the claims I wish to support.

Just what is meant by co-present foci? When more than one element is introduced in a discourse and each is discussed relative to the other or relative to a class in which both occur, the discourse is said to be maintaining co-present foci. An

example will be helpful for understanding how this behavior occurs.[1]

D1-1 Two men held up the Park Street bank at 3:05 pm today.
2 (The) one, a 6 foot 210 pound dark haired male, wore a motorcycle helmet with a darkened visor.
3 The other, a 5 foot 2 inch 100 pound blonde male, wore a stocking cap.
4 The one held a gun on the patrons,
5 while the other scooped cash and securities from five tellers.

D1-1 introduces two elements, while 2 and 3 mention each using *the one...the other* construction. At this point, both are in focus. Both continue to be in focus in 4 and 5. Why is this example not simply a matter of a discourse that talks about one thing and then another thing?

The simultaneous discussion of two or more things must somehow indicate which of the two the speaker is talking about right now. So one might expect that maintaining co-present foci would require some clear means of marking that the noun phrases are being used co-presently. The variety of definite anaphora in English is limited to personal pronouns and definite noun phrases. Personal pronouns maintain co-present foci by distinguishing on semantic criteria such as animacy. From the cases of defnps discussed in chapter 3, it appears that no definite noun phrases maintain co-present foci. But chapter 3 does not contain an analysis of *the one...the other* anaphora or *this-that* anaphora. Both are a kind of definite noun phrase; they are treated separately here because they permit and mark the use of co-present foci which other defnps do not.

D1 contains a use of co-present foci because the specification links of its noun phrases depend on their relation to the initial focus of two men. The specification of *the one* and *the other* must be made relative to the initial focus. Each time one of these two phrases is used, the initial focus of two men (and what has been learned about each) must be used to decide on the specification links. The specification of either phrase cannot take place without the initial focus, since the phrases alone incompletely

---

1. Readers disagree about the second sentence of the example. Some prefer *one* to *the one*. In the discussion which follows, the disagreement will be indicated by the use of parentheses around *the*.

specify an element. By comparison with defnps in chapter 3, *the one...the other* phrases are like the implicit specifications discussed there, because both require the focus to completely specify the phrase. More details about focus relationships for *the one...the other* cases will be presented in the next section.

Each of the kinds of co-present foci markings work in a slightly different fashion. D1 presents an example of *the one...the other* as co-present foci. Cases of *this-that* as co-present foci will be considered in an upcoming section. *This-that* phenomena are complex because *this* and *that* play an additional role in other settings. In the presence of another focus, *this* indicates a focus movement while *that* indicates a secondary dual focus. Rules for interpreting *this-that* as co-present foci as well as rules for *this* and *that* to mark focus movement will be given in the sections that follow.

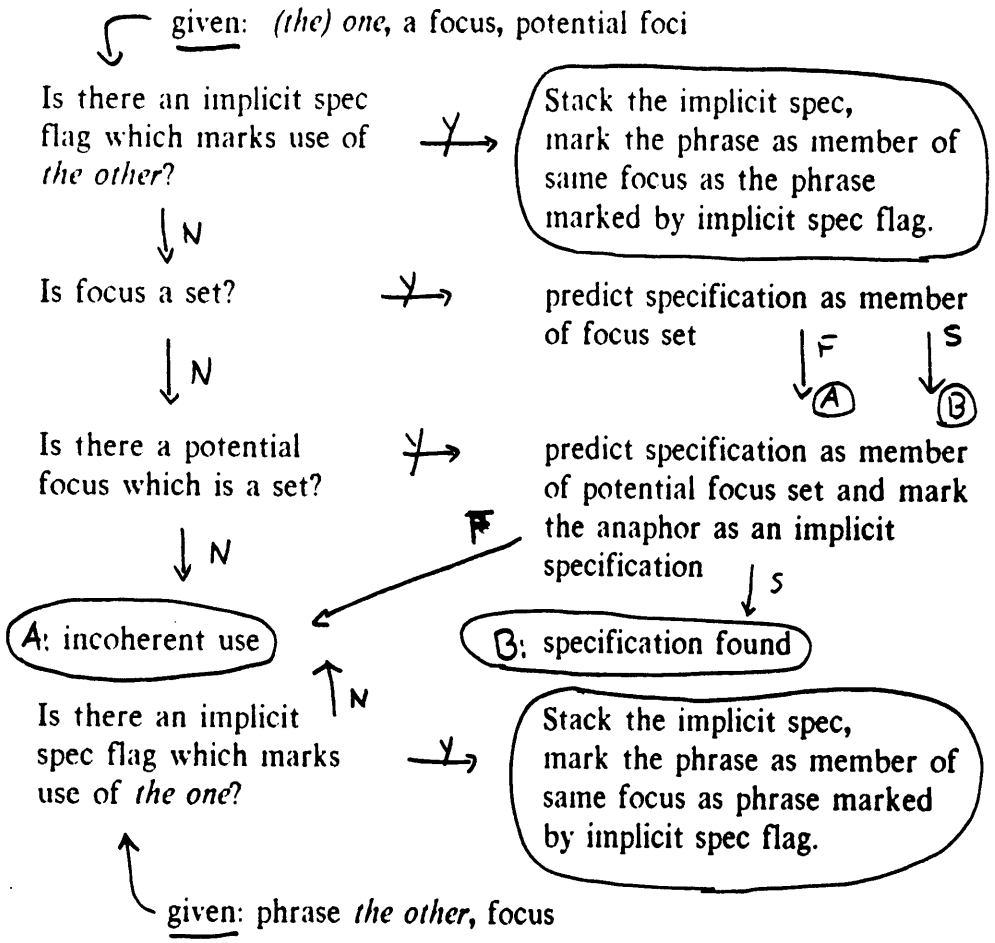
## 5.2 Rules for Co-present *the one...the other*

The rules for interpretation of *the one...the other* may now be added to the other available defnp rules. They are presented in a flow diagram in figure 1 below and are part of the implicit specification rules for defnps.

The use of these rules for D1 results in the following behavior. The initial discourse focus is *the Park Street bank* while the actor focus is *two men*; potential discourse foci include the two men, the time of the hold up, and the action. By decision 3 of the rules, a potential focus from D1-1 provides the set for the specification of *the one* in D1-2. As with the uses of co-specification rules for pronouns and the specification rules for defnps, the focus and the potential focus are sources of specification of the defnp. But the full set of rules is not quite so simple. To interpret *the other* in robbers-3, the implicit specification flag in the focussing mechanism must be used. It was included in focussing for defnps implicitly related to the focus; defnps, for example, such as *the flavor* when ice cream cones are the focus. For *the other*, the rules indicate that the implicit spec flag must mark the use of *the one*. When it does, the implicit specification is stacked and *the other* takes its place. All these changes take place as part of the rules of interpretation. Meanwhile, what is the focussing mechanism doing?

The answer is, basically, what it always does. When *the one* is encountered in

**figure 5.1. Rules for Interpretation of *the one...the other***




---

D1-2, its specification is found relative to a potential focus of *two men*. Now suppose after all sentence interpretation, the focussing mechanism runs. Step 7 of the algorithm indicates that the focus must move to the potential focus because of the implicit specification, and that *the one* becomes flagged by the implicit spec flag. Of course, *the Park Street bank* is stacked as are all default initial foci which turn out not to be of interest. When D1-3 is encountered, *the other* is interpreted as specifying through the focus. When the focussing mechanism runs, step 7 indicates that the focus should be kept where it is when an implicit specification specifies by focus association. Since this is the case for *the other*, the focus does not move from *two men*.

The focussing mechanism makes heavy use of the implicit spec flag in *the*

*one...the other* constructions. However, this is not a disadvantage because such constructions work so much like other implicit defnps and because the focus must remain on the element set of which they are a member. Here, then, is one instance where the focussing mechanism needs no changes for a co-present focus.

The rules for *the one...the other* result in two states, one where a specification is found and the other which is labelled "incoherent use." This second state is really a comment about the comprehension process. When the focussing mechanism reaches this state, it is assumed to halt with no other results than the message and the contents of the focus, spec flags, stack and the like. This state is intended to reflect the sort of situation which people reach, the "HUH?" reaction. Many incomprehensible uses of definite anaphora seem to cause this behavior; what people do when they reach it is unclear. As I have indicated before, they seem to have some strategies for trying to decide what the speaker might have meant, as well as some strategies for deciding if they should even care about whether they should comprehend. These strategies are not part of the theory presented here. The rules delimit a part of language behavior. The strategies for recovery from incoherency remain to be explored.

## 5.3 Movement with Co-present Foci *The one...the other*

The focus of the discourse may also move to center on one of the co-present foci. An example of this behavior will show how co-present foci appear and disappear as the focus changes.

D2-1 I have two dogs.
2 (The) one is a poodle;
3 the other is a cocker spaniel.
4 The poodle has some weird habits.
5 He eats plastic flowers and likes to sleep in a paper bag.
6 It's a real problem keeping him away from plastic flowers.
7 The cocker is pretty normal,
8 and he's a good watch dog.
9 I like having them as pets.

D2-1 introduces two elements, while 2 and 3 specify each using *the one...the other* construction. At this point, both have been in focus together with the initial focus of two dogs. D2-4 moves the focus solely to the poodle, and it moves on to the cocker at D2-7. In 9, the initial focus of two dogs is re-established. Why is this example not

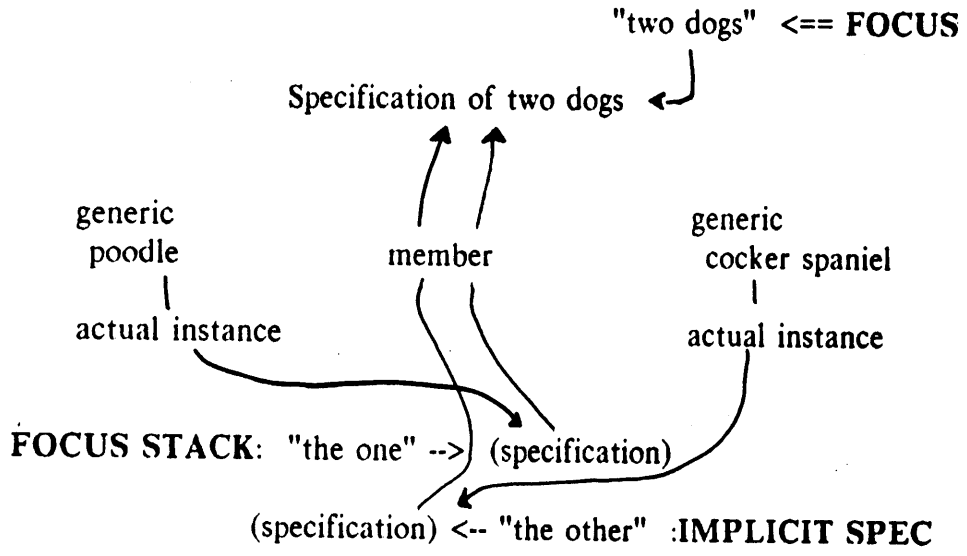
simply a matter of a discourse that talks about one thing and then another thing? What happens to the co-present foci when focus moves to a single element?

D2 contains a use of co-present foci because the specification links of its noun phrases depend on their relation to the initial focus of two dogs. The specification of *the one* and *the other* must be made relative to the initial focus. The co-specification of *the poodle* in D2-4 and of *the cocker* in 7 depend on *the one* and *the other* since they have been identified by 2 and 3 as a poodle and cocker respectively. A detailed look at how this occurs will clarify the specification and co-specification relations.

Figure 2 shows the status of the focus and of the specification links at the time that D2-4 is processed. Comprehension of *the poodle* occurs by explicit backwards co-specification to a stacked focus (*the one*) by step 3 of the defnp rules in chapter 3. When the focussing algorithm is run, the focus of the two dogs and the phrase *the other* are stacked. Using pronoun rules from chapter 4, the uses of *he* and *him* in D2-5 and 6 are interpreted as co-specifying *the poodle*. When D2-7 is processed, *the cocker* does not co-specify with either the actor or discourse focus, but it does co-specify with the use of *the other* which is in the focus stack. When the focussing algorithm runs, *the cocker* moves into focus.

---

**figure 5.2. Co-present focus links for D2**



Co-present foci reflect a special kind of structure that occurs in discourse. Several elements are introduced. When continuing discussion of one of the elements extends the discourse, the focus moves to that element. When that discussion is complete, the focus cannot simply move onto any other thing the speaker wants to mention. The discussion should return to the other elements, and those elements discussed. However, the discussion of one element for an extended part of the discourse may involve introduction and consideration of other elements. The real constraint in the foregoing analysis is that discussion should eventually return to the other elements introduced via co-present foci. When it does not, the hearer is left to wonder why co-presence was used in the first place. As with other constraints, the focussing algorithm does not force this behavior, but the presence of co-present foci in the focus stack indicates where a speaker may have failed to meet the constraint. For example, in D2, no intervening potential foci are sensible co-specifications for *the cocker*. However, in the following addition to D2, *that cocker* does not co-specify with the one introduced in D3-3:

D3-1 I have two dogs.
2 The one is a poodle;
3 the other is a cocker spaniel.
4 The poodle likes to play with another cocker from down the block.
5 That cocker is very feisty,
6 but that doesn't seem to interfere in the games they play.
7 My cocker is pretty normal,
8 and he's a good watch dog.
9 I like having them as pets.

This discourse would be incomplete only if discussion never returned to the speaker's two dogs; for then, the hearer might wonder why it was that the speaker did not simply begin the discussion by telling the hearer about the poodle. To summarize, co-present foci affect what should eventually be said to make any discourse complete. The eventualities, however, do not affect the interpretation of definite anaphors in between.

Membership relations are a reliable indication of elements which are multiply focussed upon in the same way as *the one...the other*. Two typical examples are the set-element identification, discussed in chapter 3, and conjoined noun phrases such as *a boy and a girl*. These cases may be comprehended in the same way as *the one...the other* examples.

## 5.4 Exophoric Demonstrative Uses of *This* and *That*

Many times in discourse *this* and *that* are used exophorically, i.e. to point to some object in the view of the speaker and hearer. Thus if a speaker says: "I want that picture on the wall," while pointing, *that* is being used to point (i.e. demonstratively), and to point to something external to the conversation (i.e. exophorically). Most linguists regard these uses as explained by a notion of what is near and not near the speaker (see Halliday and Hasan [1976] and Fillmore [1971]). However the matter turns out, exophoric demonstratives are not a topic of this chapter simply because an adequate explanation of them requires a theory of perception to account for what is meant by the notion of pointing and just how the hearer comes to view the object referred to. Instead, this chapter will consider cases of *this* and *that* where a co-specifying noun phrase appears explicitly in the discourse, or in the case of *this*, where the antecedent follows in the discourse.

## 5.5 Use of *This* and *That* for Co-present Foci

*This* and *that* are a marked means for carrying two foci. However, before stating the arguments in favor of this view, I will review previous work on *this* and *that*,

Halliday and Hasan [1976] claim that in general *this* and *these* imply proximity to the speaker; *that* and *those* imply distance from the speaker which may or may not involve proximity to the addressee. The meaning of *that* is "near you or not near either of us, but at any rate not near me."[1] They also claim that *this* is used to convey a sense of immediacy and of solidarity with the hearer, a kind of shared interest and attention. Even the non-anaphoric use of *this* in "There was this man..." is used to emphasize common experience and common interest. Halliday and Hasan point out that the most common use of *this* and *that* is as extended reference, a term they define to mean use of any identifiable portion of a text. For example, in D4-2, *this* is understood as referring to the whole previous sentence.

D4-1 No one will take it seriously.
2 This is the frightening thing.

---

1. Halliday and Hasan [1976], *op. cit.*, page 58-59.

Halliday and Hasan note that in many instances of *this* and *that* there is an ambiguity between reference to a specific object and to extended reference. How such ambiguities arise will be discussed in the following section, as well as an alternate proposal for use of *this* and *that*.

Lakoff [1974] does not offer a theory for *this* and *that*; she observes several catagories of use. The first is spatio-temporal deixis which includes the same observations of proximity as Halliday and Hasan. The second is discourse deixis, i.e., those uses where the anaphor is used to refer back or forward in the text. An observation of the behavior of *this* suggests that *this* must refer back to some object specifically in the text; thus D5 is acceptable but D6 is not.

D5-1 I saw Fred in his new sombrero.
2 This hat is really something.

D6-1 I saw Fred in his new outfit.
2 This hat is really something.

The third category of use is emotional deixis. One type uses *this* to allude to something already mentioned, but outside the discourse proper as in:

D7-1 I see there's going to be peace in the mideast.
2 This Henry Kissinger is really something!

Another type uses *this* to replace indefinites in contexts where one wants to express vividness as in:

(1) He kissed her with this unbelieveable passion.

A third type of emotional deixis is the use of *this* to provide additional information about some person who has been referred to previously. The next section will provide an alternative account of some of these cases.

Webber [1978] points out a particular use of *that* and *those* as a representative of *one* anaphora: cases of *that* or *those* followed by a post-modifier. A typical example is:

> It was known that if one took certain mutant bacteria and mixed them with an extract of normal bacteria, some cells became like those that provided the extract and gave rise to transformed progeny cells.[1]

---

1. Luria [1975], *Thirty Six Lectures in Biology*. MIT Press, Cambridge, MA., page 43.

Although *one* anaphora will not be discussed in this report, this particular use of *that* and *those* is well marked in the language. It also adheres to the general claim about *this* and *that* being put forth here; *that* marks a secondary focus of interest. The next section presents arguments supporting this claim.

Focus provides a different method for accounting for the behavior of *this* and *that*. The choice of which to call *this* and which *that* depends upon which the speaker wishes to be of chief concern or interest in the discourse. Consider the example below:

D8-1 I'm having a party tomorrow night;
2 it will be like the one I had last week.
3 That party was a big success
4 because everyone danced.
5 This one will have better food.
6 I've asked everyone to bring something special.
7 Want to come?

Two different parties are talked about; that is, both of them are in focus. To indicate that the speaker wants to discuss both, *that* is introduced to co-specify with the one mentioned second. The second party is used as a means for comparison to the first; hence *this* indicates the main concern of the speaker while *that* a secondary concern.

The use of *this* and *that* differs from *the one...the other* because the initial foci in the two cases are different. As has been shown, the initial focus for *the one...the other* is a phrase which represents a set to which specifications for *the one...the other* are linked. For *this* and *that*, no initial focus representing a set is needed. Instead, whichever of the two noun phrases occurs first acts as a source for specification of the other noun phrase. By a "source" for specification, I mean a phrase which has similar properties but a possibly different specification. As D8 shows, the initial focus receives a *this* use and the second focus a *that* use.

To consider *this* and *that*, D8 will be analyzed in detail for a summation of the focus movement rules. D8-1 and 2 establish a focus according to the focussing algorithm. Among the potential foci of D8-2 is *the one I had last week*. D8-3 indicates that the speaker wants to say more about the potential focus while maintaining the first focus; this is accomplished by means of using *that* instead of *the* to co-specify with the party last week. If *the* had been used, it would cause the hearer to suppose initially in processing that the speaker was talking about the upcoming party; then the hearer

would need to reject the choice because of the tense of the verb. *That* is a much clearer means of telling the hearer which one is under discussion.

How is the first focus maintained for D8? One possibility is to use the focus/co-present focus markers of *the one...the other*. However, since that mechanism assumes set relationships for focussing which do not exist in *this* and *that* cases, this possibility cannot be realized in this discourse. Instead just stacking the first focus is sufficent. When a noun phrase with *this* as determiner is encountered, the co-present focus from the stack must be chosen.

The basic kernel of the rule for *this* and *that* is: *this* is a determiner used for main focus, i.e. *this* + <noun phrase> determines main focus, while *that* + <noun phrase> co-specifies with an old focus. However, if the focus has been mentioned using *that*, then a *this* defnp must co-specify with an old focus. D8 is a case where the second item discussed (which is focus by focus movement) is given *that* as determiner, so that when the first item is discussed again, *this* must be used. An example of the normal rule instantiation is given below in D9.

B: What are the plans for the banana raid?
A: According to Hilda's plan, you and I stay here until everyone else is in position. I don't much like it because I think we'll miss all the action. Wait. I've got a better plan: we'll be the guide party, and Eloise and Hilda the search party. Then we'll be in on the action. Well, what do you think, isn't this a better plan than that one?

The rule for *this* and *that* reflects the locus of the speaker's concern. In D8 the first thing introduced is the chief concern, while in D9, A indicates concern with her own plan rather than Hilda's. Hence the second part of the *this* and *that* rule reflects how the speaker may stipulate his/her interest: when the speaker uses *that* as determiner for a defnp which co-specifies with the focus, the speaker is indicating that chief concern lies with another element. This observation helps explain some of Halliday and Hasan's conclusions about empathy using *this* and *that*. Furthermore, it offers an alternative explanation for the speaker-hearer proximity account of *this* and *that*. This observation may be stated explicitly as the concern constraint: *this* as a determiner of a noun phrase indicates that the element co-specified by the defnp is the speaker's chief concern, while *that* indicates secondary concern.

The above explanation is incomplete. The concern constraint functions only when speaker and hearer are focussed on the same elements. In D8 and D9, speaker and hearer are focussed on what the speaker makes the focus of the discourse. However, in some dialogues, the speaker and hearer do not always share focussed items. Consider D10 below.

D10-1 A: Let's flip a coin and see who calls it.
2 B: Heads.
3 A: That's what it is. (* This is what it is.)

The focus of this dialogue is the coin which is being flipped. B has a second focus which is the result of the toss. When A speaks of B's focus, A uses *that* to refer to it; *this* cannot be so used. When several examples are considered, the proper formulation of the rule becomes clear: when speaker and hearer have different focus, use *that* as the determiner of a defnp which co-identifies the non-shared item, and use *this* for shared items.

Halliday and Hasan put forth a different explanation. They claim that the rule governing *this* and *that* between speakers is the following. Use *this* to refer to what you are talking about and *that* to refer to what the other person is talking about. This rule is inadequate although it will provide an account of D10. Using the notion of focus and shared focus, however, accounts for cases such as the one below[1] on which the Halliday and Hasan rule fails.

D11-1 (The expert asks to be shown a small screw which the apprentice has found.)
2 Apprentice: Ok, I'm going to show this little screw to the consultant. (when saying this sentence, the apprentice is talking to himself)
3 Apprentice: (Speaking to the TV camera person) Can you focus on that little screw?

In D11, when the apprentice is speaking to a hearer (that is, himself) who shares his focus on the screw, *this* is used. When speaking to a hearer who does not share focus, the apprentice must use *that*. In summary, the full statement of the concern constraint is this:

---

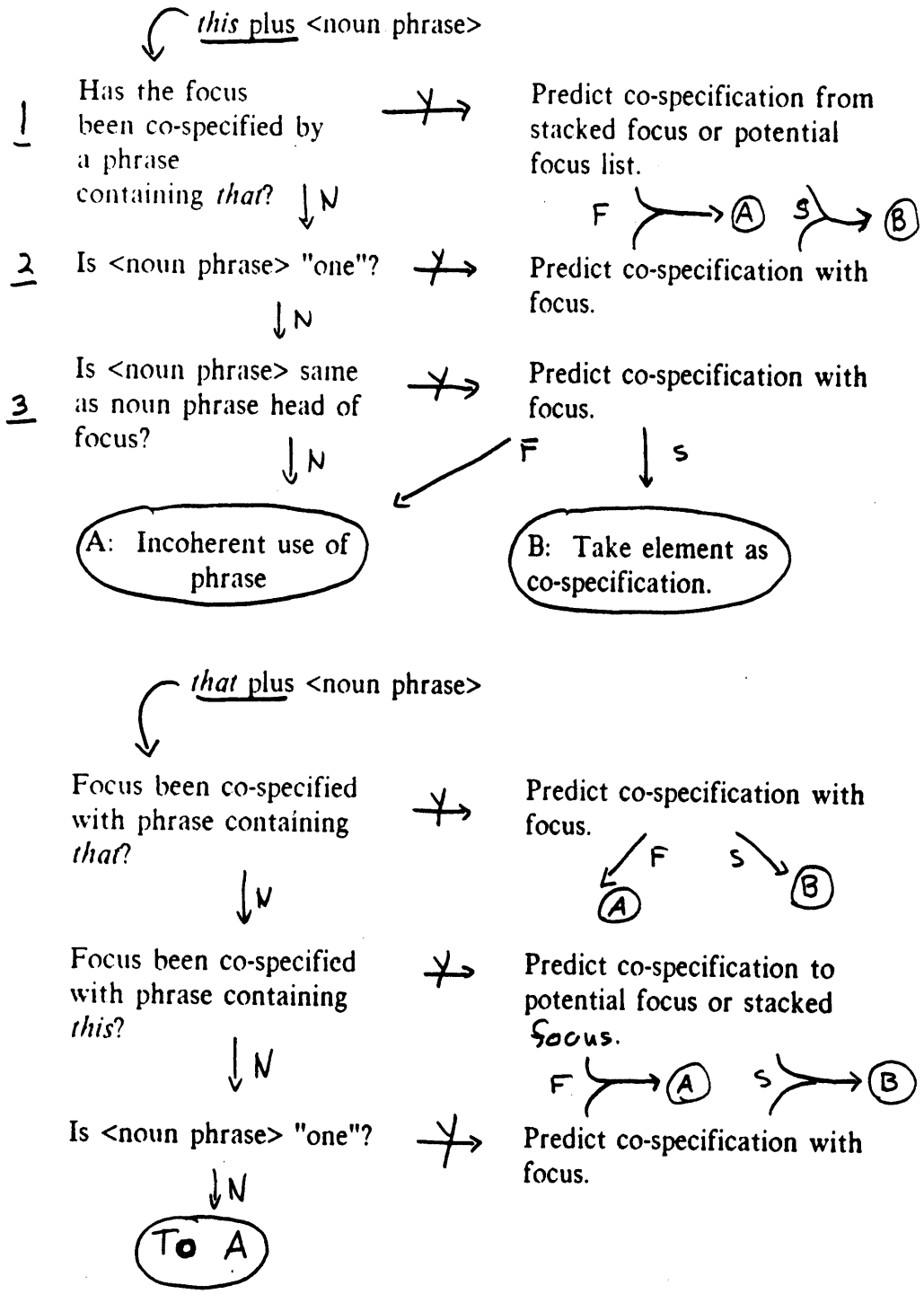
1. This example comes from Dialogue 1 of Grosz [Deutsch, 1974].

> Concern constraint: When the speaker and hearer share a focus, *this* as a determiner of a noun phrase indicates that the element co-identified by the defnp is the speaker's chief concern while *that* indicates secondary concern. When speaker and hearer do not share a focus, use *that* as the determiner of a defnp which co-identifies the non-shared item, and use *this* for shared items.

Given the previous discussion, the rules for use of co-present *this* and *that* are given in figure 3. The rules will cover only shared focus conditions since these are assumed in this report. An algorithmic account of the non-shared conditions involves a model of the speaker's beliefs of the hearer, which, as indicated previously, is beyond the scope of this report.

In summary, *this* and *that* move focus when the two determiners occur together in a discourse. They are essentially co-present foci since the speaker uses them to mark discussion on more than one thing in the discourse, but unlike *the one...the other*, they do not require the specification flag which is needed for *the one...the other*; they require only rules specialized to the presence of determiners on the focussed elements. Just as in the case of *the one...the other*, use of co-present *this* and *that* requires no adjustment to the focussing algorithm. The focussing algorithm makes its moves based on the co-specification and specification relations found in the sentence by the anaphor rules for co-present foci. Hence co-present foci anaphora are analogous to explicit and implicit defnps and pronoun cases.

**figure 5.3. Co-specification for co-present *this* and *that***



## 5.6 *This* and *That* in Focus Movement

*This* and *that* play two major and different roles as definite anaphora in English; they mark co-present foci and when only one of the two occurs, they mark focus movement. Because of these two roles, it has been difficult to analyze *this* and *that* and specify rules for their behavior. The last section presented examples of co-present foci and rules which govern their behavior. This section will show how these words signal focus movement.

When considering the behavior of *this* in discourse, one may observe that a *this* defnp moves the focus to whatever the noun phrase of the defnp specifies. As the rules in the previous section stipulate, usually the focus moves to the leading potential focus in the potential focus list. Yet sometimes the focus moves to the entire description given by the previous sentences; sometimes, surprisingly, the focus does not really move in the sense that a new element is co-specified; the same element is specified but from a different perspective. These cases are differentiated by the lack of *this* and *that* use together. In non-co-presence, *this* and *that* are not used to contrast between two items. Instead, *this* moves the focus, while a *that* defnp does not necessarily mark focus movement. If this claim is true, then the important question to address is: how are co-present foci distinguished from non-co-present cases? Before this question is considered, a detailed analysis of *this* movement cases must be undertaken. Each of the different cases will be reviewed and explained as focussing phenomena. This discussion will start with examples from texts in which each of the different uses of *this* occur.

The most common use of full *this* defnps[1] is to mark focus on an element in the potential focus list. However, since most of this report has concerned movement to a potential focus, that claim is hardly surprising. What is surprising is that the potential focus and focus have the same noun phrase head. Normally in such a circumstance, the focus would be preferred, but with *this* defnps the opposite behavior results. One might wonder why this would be a use of *this* defnps. The reason is clear: when both the

---

1. I distinguish here between the use of *this* plus a noun phrase (e.g. *this book*), called a full *this* defnp, and *this* and a null noun phrase (e.g. *this*), called an empty *this* defnp. The latter case will be treated last.

focus and a potential focus are the same kind of elements, if the speaker wants to talk about the potential focus, *the* as a determiner will signal focus, while a pronoun would at best be ambiguous. *This* provides a method for marking that the potential focus is going to be discussed. An example[1] of this behavior is given in D12. In this example, both the focus and the potential focus are axons. The *this* defnp co-specifies with the potential focus of *man's longest axon.*

D12-1 The axon may run for a long distance....
2 Man's longest axon runs for several feet, from the spinal column to muscles that control movements of the toes.
3 In spite of its great length, this axon, like all nerve fibers, is a part of a single cell.
4 It is living matter.

A second use of *this* co-specifies with the focus from a different perspective. While the example above shows that potential foci should be considered rather than the focus, the example below shows that when the potential foci are unacceptable as co-specifications, then the focus must be considered. An example with two such uses is this:

D13-1 Consider the roomful of electronic equipment that makes up a modern, high-speed digital computer.
2 Rack after rack of transistors, diodes, magnetic core memories, magnetic film memories--
3 all laced together by an intricate system of wiring many miles in length.
4 Imagine the room, and everything in it, shrunk to about the size of a cigarette package.
5 Now suppose we give this marvelous box to a clever electrical engineer, a man working, however, not in our own midcentury, but about the year 1900.
6 We present our gift
7 and demonstrate a few of the remarkable feats it can perform: several hundred thousand additions in one second...
8 We leave this tantalizing device with the suggestion that he try to find out what's inside the cigarette package...

---

1. This example and the next are from Denes and Pinson [1973], *The Speech Chain: The Physics and Biology of Spoken Language.*, Anchor Press, Garden City, New York, pages 124, and 122, respectively.

At D13-4 focus centers on *the room*. In the next sentence, the room is referred to as *this marvelous box*, a very different way of seeing a room, but not a surprising one if it has been shrunk to the size of a cigarette package. D13-6 uses a defnp (*our gift*) to place focus on the act of giving in the previous sentence.[1] Focus remains there in D13-7. D13-8 contains *this tantalizing device*. This defnp co-specifies with the focus, but it describes the gift from a very different perspective, not one of a more generalized concept, but one which reflects the fact that the gift can perform a lot of fast computation and other properties (not repeated in the text) ascribed to it in D13-8.

Empty *this* defnps are the most difficult to disambiguate in the theory of definite anaphora given here. However, this is not necessarily a flaw in the theory since it appears that native hearers also have difficulty with empty *this*; rhetoric books often warn writers against use of it because of its ambiguity. In spite of this warning, empty *this* usage occurs in text (including in this paragraph!). It is generally used to co-specify the predication of the entire preceding sentence. For example[2] :

D14-1 A basic fact of gravitation is that two masses exert forces on one another.
2 We can think of this as a direct interaction between the two mass particles, if we wish.
3 This point of view is called *action-at-a-distance*, the particles interacting even though they are not in contact.

D15-1 Since however, the interpretation has been put forward as a hypothesis, some weight will be added to it
2 if it can be shown to have an antecedent probability.
3 This is what I shall endeavor to do in the remaining pages.
4 {additional text of several sentences}
5 The circumstances of his life were such as to encourage, if not to cause, both naturalistic and transcendental elements in his philosophy.
6 To support this contention, it will be necessary to review those aspects of his environment and personal career which have a bearing on the tendencies in question.

---

1. One might want to claim that *our gift* co-specifies with box without focussing on the act of presentation. But the only reason the box is a gift is that it is stated as having been given; boxes are not inherently gifts, they just happen to be in this case.
2. Resnick, Robert and David Halliday [1966]. *Physics: Part I.* John Wiley and Sons, Inc. New York, page 404, and Goudge, [1969], page 326.

Both D14-2 and D15-3 (without the *if*), use empty *this* phrase to co-specify the entire predication in the previous sentence.

Like full *this* defnps, empty *this* co-specifies with a member of the list of potential foci; empty *this* preferring the verb phrase and full *this* the lead noun phrase in the list. A decision must be made concerning whether that preference holds for a particular use of *this* or not; in D15-6, the full *this* defnp, *this contention*, co-specifies the predication of D15-5, and so the preference does not hold. Deciding whether the preferred potential focus is the co-specification of a full *this* defnp therefore requires a judgment by the inference mechanism that the co-specification is acceptable. In D15-6 neither circumstances nor elements are judged as contentions. Given such a judgment, the other member of the potential focus list is accepted as co-specification of the focus of *this contention*.

*this* is also used to single out a potential focus when the focus and potential focus are different kinds of elements. Such behavior is similar to the use of *the* in defnps, but it most commonly occurs when a *this* defnp has just been used to move the focus. In doing so, the speaker is marking several very rapid movements of focus. In the example below, there are three uses of *this* in five sentences.[1]

D16-1 The number of such elements need not, I think, be finite,
2 just as long as the system assumes that it can determine all the properties of each element.
3 Clearly if a system does not make this assumption,
4 (akin to the "closed world" assumption for databases discussed in [Reiter, 1977])
5 then it must be able to derive and manipulate IDs.
6 For example, this assumption was not made in the Travel Budget Manager's Assistant system developed at BBN as part of the Speech Understanding Project.
7 If this system were told....

One other use of full *this* defnps which occurs in discourses must be noted here. Inside of a quantified phrase a full *this* defnp does not move the focus to itself. Furthermore, the *this* defnp takes its co-specification from the quantified variable; such

---

1. From Webber [1978], page 173.

cases are similar to the bound variable pronouns discussed in chapter 4. Quantified phrase patterns also use *that* in a similar way. In the example[1] below, the quantified phrase and *this* anaphora are underlined.

D17-1 We can, therefore, associate with each point near the earth a vector **g** which is the acceleration that a body would experience if it were released at this point.
2 We call **g** the *gravitational field strength* at the point in question.

The above example introduces a vector of a certain type and defines it. Part of the definition requires another item, a point, to be mentioned more than once. Even though another item is in focus, namely the body, which is co-specified by *it*, the author wants to keep attention on the point, so *this* is used. However, the point discussed in D17 is not the focus, because in D17-2, it must be spoken of using the definite noun phrase *the point in question*; if *it* were used, the reader would understand *it* to mean the body experiencing acceleration.

Now consider the cases of *that* used non-co-presently. *That* anaphora are used to single out an element in the text. The kinds of elements singled out are surprisingly different. One use of *that*, called new mention *that*, describes an element which has not been mentioned previously in the text. Unlike use of *this*, new mention *that* does not force focus movement. Examples of new mention *that* with focus movement as well as without focus movement are given below.[2]

D18-1 This is a course in biology.
2 Biology studies those entities that are called organisms: men, worms, yeast cells, bacterial cells are organisms.
3 Some organisms are unicellular,
4 some are multi-cellular.

D19-1 In Marigold's garden, roses grow everywhere.
2 She likes roses of the Eastern gorge variety more than those of the Western shore,
3 so she has a lot of them in her collection.
4 They grow to prize winning shapes and sizes.

D18 contains a *this* used exophorically; the focussed element is the one described by *that*, i.e. *those entities that are called organisms*. By using a *that* determiner, the author

---

1. From Halliday and Resnik, *op. cit.*, page 405.
2. The first example comes from Luria, *op. cit.*, page 3.

is permitted to bring into discussion some element which has not been mentioned before. This behavior is reminiscent of the concern constraint because the speaker is talking about something which the hearer does not share as a focus. *That* defnps move into focus in a manner similar to *the* defnps, i.e. by the potential focus list. *That* defnps do not automatically become the focus. When an anaphor following the *that* defnp co-specifies with it, then the *that* defnp is the focus; otherwise the focus never moves to the element specified by the *that* defnp. In D19 the behavior of *that* is similar; the speaker introduces *those of the Western shore* without turning discussion to them.

The use of *that* in D18 may be contrasted with its use in D19. In D18, the *that* use specifies a new element of the discourse without help from the focus, that is, there is no implicit specification with the focus as in the cases discussed in chapter 3. However, in D19, the *that* phrase has a deleted noun phrase head, so that *those of the Western shore* specifies a kind of rose. This specification depends on the use of the focus. While new mention *that* phrases describe an element not previously mentioned in the text, the phrase may be implicitly specified by means of the focus.

The other use of *that* will be called previous mention *that*. A previous mention *that* phrase takes as antecedent some phrase and its interpretation, the phrase having been mentioned previously in the discourse. An example[1] is given below.

D20-1 If MNMSD is referred to by D either as "the mayor of San Diego" or "D's neighbor,"
2 then node 'MNMSD' represents the individual referred to.
3 The problem is that only looking at that node provides no reflection of the differences in the two references to MNMSD,
4 even though the surface DEFNPs do express this difference.
5 Focus spaces provide a means of representing this difference.

In D20, *that node* co-specifies with the node of D20-2. If the rest of the discourse is ignored, D20-3 would have been equally acceptable using *this*. However, the author does not want to keep focussing on *that node*; in the next sentence use of *this difference* forces the focus onto the differences. Note that while the presence of *this* appears to be significant in previous mention *that*, it is not necessary; it merely clarifies what is in focus. An example of previous mention *that* where no *this* is given may be found in

---

1. From Grosz [1977], page 82.

Goudge [1969], page 69:

D21-1 But the aim of counting is to assist reasoning.
2 In order to do that, it must carry a *form* akin to that of reasoning.
3 Now the inseparable form of reasoning is that of proceeding from a starting point through something else, to a result.

Like empty *this*, empty *that* seems to prefer the predication as its co-specification.

The important question about previous mention *that* is why it exists at all in the language. It is clear from D20 why *this* cannot be used, but what about *the* or *it*? In the examples I have found, *it* is generally ambiguous, and hence is an ineffective means for indicating focus. However, *the* in place of *that* is possible. Suppose D20-3 were:

(2) The problem is that only looking at the node provides no reflection of the differences in the two references to MNMSD...

(2) could be used in the D20 discourse. However, the use of *the* forces a movement of focus from the person to the node, when what the author actually wants to turn her attention to is the differences in the two references to the person. In other words, an intervening, and in this case unnecessary, focus movement occurs.[1] Hence *that* serves a useful function in the language, that is, to introduce terms without the term becoming the focus of the speaker's (and therefore the hearer's) attention.

Now that the various uses of *this* and *that* used non-co-presently have been shown, rules for their behavior may be stated in process terms. The rules given in figures 4 and 5 include the co-present cases as well. They are included because it is necessary to be able to distinguish co-present from non-co-present uses, and this algorithm provides that distinction.

As with the other defnps considered in this chapter, the *this* and *that* rules in figures 4 and 5 do not change the focussing mechanism. The mechanism proceeds using

---

1. There is another reason for using *that*. The context which precedes the text of D20 makes reference to a figure in the text. *That node* is a pointing behavior. This example suggests that there is an important relation between focussing and pointing behavior. As has been stated earlier, it will not be discussed in this report, but remains a matter for further research.

the co-specification information as before. Only the *that* rules suggest any new behavior in the rules. The B state of the rules indicates that co-specification information must be masked from the focussing algorithm. The mask prevents the focussing algorithm from noticing a co-specification, which is normally how it moves the focus. Some readers may object that this is really a change in the focussing mechanism, but in fact, the mask must happen in the anaphora rules because it is there that the situation must be noticed.

A more important effect of *that* defnps must be noted here. Normally potential foci are collected on the basis of sentence order. Only syntactic effects such as clefting, discussed in chapter 2, are permitted to change the order of what is added to the potential foci list. *That* defnps also change the order since they are more likely to become a focus than other potential objects. Hence the potential focus list algorithm must include a step which puts phrases with *that* definite articles first in the list.

**figure 5.4. Interpretation of *This* Noun Phrases**

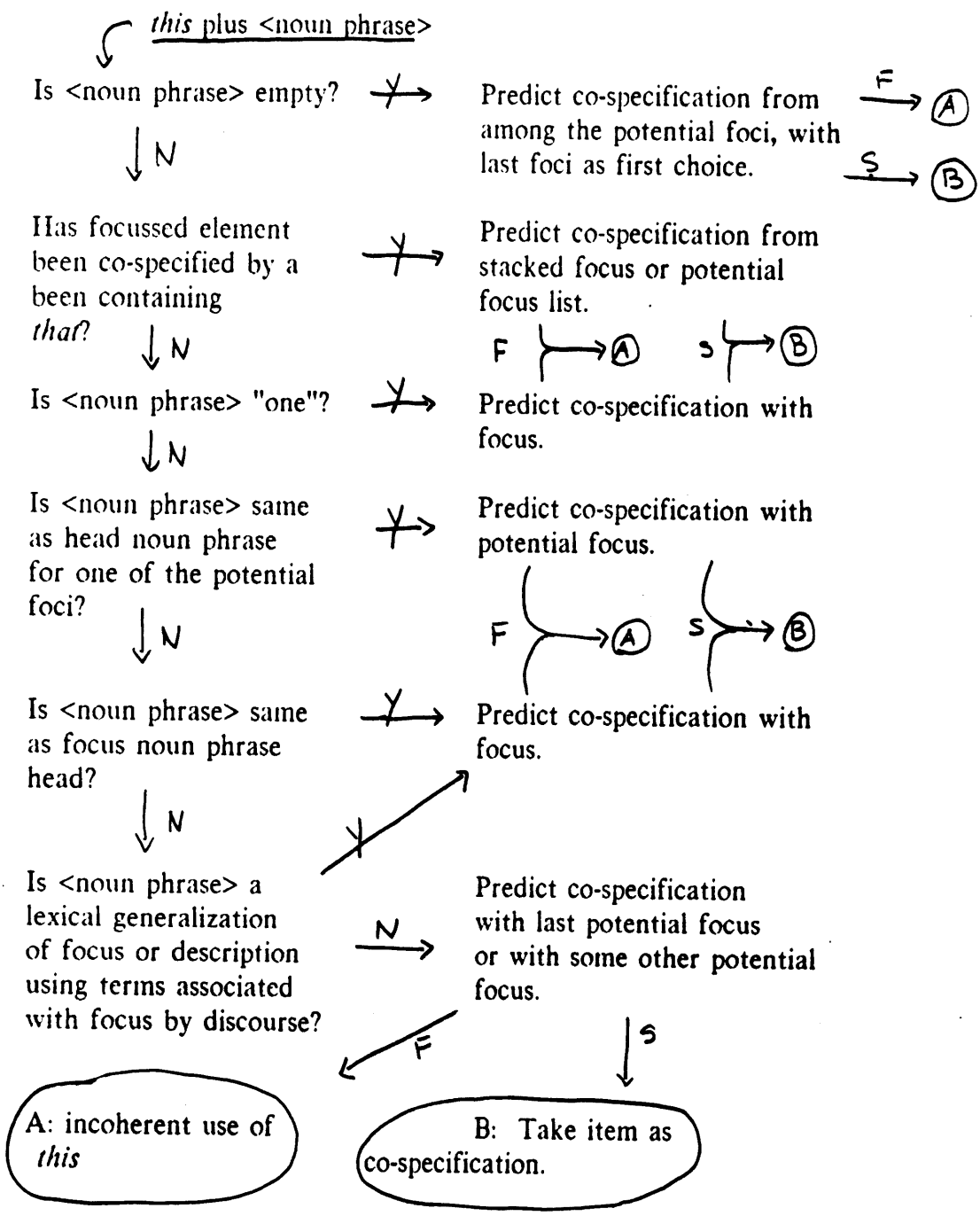
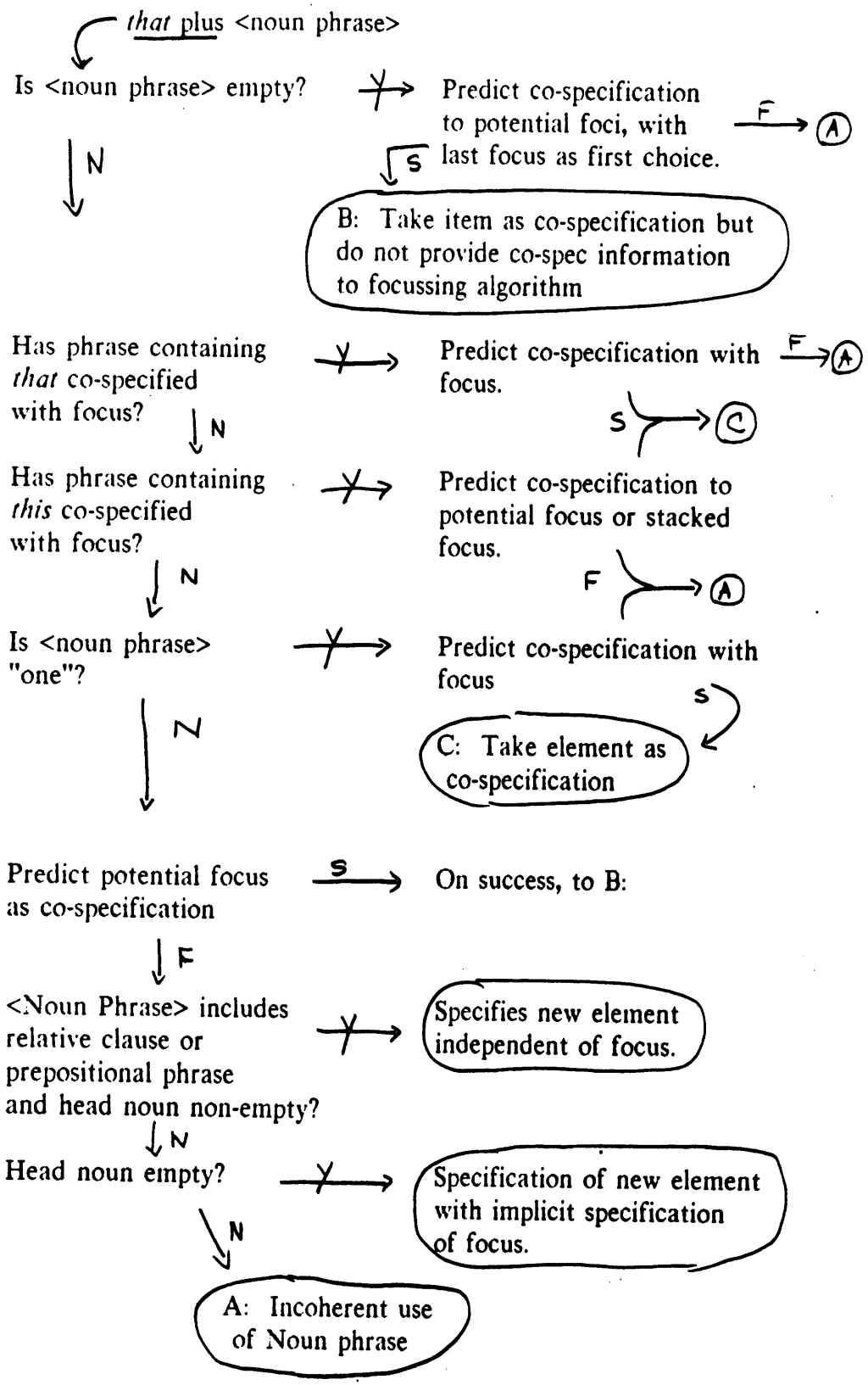


**figure 5.5. Interpretation of *that* Noun Phrases**



## 5.7 Using the Focus Movement Algorithm

Because the co-present foci rules have some complexity, the use of them with focussing is difficult to grasp. So that the phenomena may be more fully understood, two examples will be analyzed in detail. The use of the rules and the focussing algorithm will be described fully. The first example was shown previously and discussed informally:

D22-1 I have two dogs.
2 (The) one is a poodle;
3 the other is a cocker spaniel.
4 The poodle likes to play with another cocker from down the block.
5 That cocker is very feisty,
6 but that doesn't seem to interfere in the games they play.
7 My cocker is pretty normal,
8 and he's a good watch dog.
9 I like having them as pets.

The expected focus of D22-1 is *two dogs*. D22-2 contains a use of *the one*, which is predicted to specify as a member of the focus set according to the rules given in figure 1. This specification causes the expected focus to be confirmed as focus according to step 7 of the focussing algorithm. So far, the actor focus is the speaker. D22-3 contains *the other* which causes the implicit specification flag to be switched and *the one* stacked. The focus remains on the speaker's two dogs. In D22-4, *the poodle* co-specifies with the stacked focus of *the one* by lexical generalization. Although *the poodle* is in agent position, by step 5 of the focussing algorithm, since no other anaphora are present in dogs2-4, poodle becomes both actor and discourse focus. The potential discourse focus list after D22-4 contains *another cocker, the block* and the entire verb phrase. D22-5 keeps the discourse focus on the poodle since the rules in figure 5 state that the focussing algorithm is not provided with the co-specification of *that cocker.*

D22-6 contains two interesting anaphora, *that* and *they*. Since *that* is an empty *that* use, by figure 5, it is predicted to co-specify with the last member of the potential focus list (that is, the verb predication). This co-specification prediction is acceptable, but focus does not move to it. The interpretation of *they* requires the use of the agent position pronoun rules of chapter 4. At D22-6, the actor focus and the discourse focus center on the poodle; the potential actor list includes the cocker. Decision 6b applies,

and a prediction of the actor and potential actor as co-specification is acceptable, thereby making the co-specification of *they* the poodle and the cocker from down the block. Actor and discourse focus move to *they*, since it is the only anaphor in the sentence available to the focussing mechanism for focus movement.

*My cocker* cannot co-specify with the discourse focus because the discourse focus points at two dogs, one of which is not known to be owned by the speaker, the other of which is owned by the speaker but is a poodle. Instead the focus stack is searched for a co-specification. The stack includes (from bottom to top): *two dogs, the one* (a cocker spaniel). *The one* is taken as the co-specification. This choice also pops the focus. The co-specification of *he* with the focus for D22-8 keeps the focus on the speaker's cocker spaniel. Finally, the co-specification of *them* is resolved to *two dogs* by decision 3b of the non-agent pronoun rules. The first prediction of the discourse focus and potential focus is rejected (the only potential focus is the predication of D22-8); the second prediction is to choose a co-specification from the focus stack. Since *two dogs* is syntactically and inferentially acceptable, it is chosen as co-specification of *them*. Note that D22 meets the return constraint for uses of terms like *the one...the other*.

Another example, which may be analyzed using *this* and *that* rules, uses *this* and *that* non-co-presently.

D23-1 One day Bill's father bought Bill a new softball.
2 Bill and his friends played with it daily.
3 Then Harry got a hardball.
4 This ball had more speed and accuracy than Bill's,
5 so Harry and Bill and company played with it.
6 That bothered Bill's father
7 because he didn't like to see Bill neglect his toys.

The expected focus of D23 is *a softball*. It is confirmed by the use of *it* in D23-2. D23-3 introduces a hardball, which is a potential focus for the discourse. Decision 4 of figure 4 requires that *this* defnps use the potential focus list as a source for co-specifications, when no *that* is present and when the source passes syntactic and inference criteria. It causes *this ball* in D23-4 to co-specify with the potential focus of *a hardball*. The focus moves to *this ball*. D23-5 contains a pronoun which is taken as co-specifying the focus of hardball. In the next sentence an empty *that* occurs. The potential focus list of the previous sentence contains Harry, Bill and company and the predication expressed by the verb phrase. The rules for *that* predict that the last

member of the potential focus list is the co-specification of *that*, a predication, which is what, intuitively, *that* co-specifies with.

## 5.8 *This* Noun Phrases Used Non-Anaphorically

One use of *this* noun phrases has not been discussed so far. It is best presented by Prince [1978], and is a phenomenon which occurs only in spoken discourse. *This* noun phrases may be used as specific indefinite noun phrases. Prince cites among many examples:

D24-1 I work in electronic and auto shows.
2 Companies hire me to stay in their booth and talk about products.
3 I have this speech to tell.

In D24-4 *this* could be replaced by *a*, and the noun phrase is still a specific indefinite. Prince also cites indefinite *this* as one of the categories distinguished by Lakoff, which were discussed in a previous section (see (1)).

Why does colloquial English allow for two different ways of marking specification indefinites? The answer lies behind some statistics cited by Prince. She indicates that in 209 of 243 cases (86%), the element introduced by an indefinite *this* use is mentioned again within a few clauses. Hence indefinite *this* is a strong, though not an entirely obligatory, means for marking the focus of attention. In contrast to the other uses of *this*, indefinite *this* does not co-specify with another phrase in the discourse. In initial discourse, such cases are straightforward to distinguish because no focus exists; in fact, the focussing algorithm may be modified without difficulty to expect the use of initial *this* to be a preferred initial focus. In mid-discourse, non-co-specifying (i.e. indefinite) *this* could be recognized as part of the decision rules in figure 4. These modifications have not been included here since further research is needed to see whether such changes account for indefinite *this* use in numerous examples. Since this behavior seems to occur only in spoken discourse, it requires further consideration to determine why it only occurs there.

## 5.9 Conclusions about *This* and *That*

This chapter has shown examples of a variety of ways in which four kinds of noun phrases are used. *The one...the other* require a few simple rules and use of the implcit specification flag in the focussing algorithm. *The one...the other* are used to indicate two secondary foci relative to the main focus. This relationship is a more general form of direct association specification with focus which was discussed in chapter 3. The behavior of *the one...the other* is easily recognized and therefore requires simple rules.

By contrast, the rules for *this* and *that* show some complexity, which reflects the variety of ways the determiners are used in English discourses. Part of this chapter has provided an explanation of why these different behaviors must be present. Briefly summarized, *this* serves to indicate the focus of the speaker's attention when the previous focus is the same type of noun phrase or when the speaker wishes to view the focus from a different perspective. *That* is used to indicate a secondary concern which is not to be focussed upon. The role of *that* is the same whether a *this* noun phrase is present or not. Both *this* and *that* have an additional use. They provide a means for indicating a newly specified element. In the case of *this*, however, the element is likely to become focus, while for *that*, the focus is likely to be kept elsewhere. It may be concluded that *this* and *that* used co-presently allow the maintainence of two foci, one of main concern and the other of secondary concern to the speaker. *This* and *that* used non-co-presently, that is, when only one of the two types of noun phrases is found, also indicate main concern (*this*) and secondary concern (*that*) relative to some other focus.

Co-presence is a means for talking about two or more discourse elements that are related to each other. Because language is linear, and perhaps because people have trouble paying close attention to two things at once, it is not really possible to focus on both elements simultaneously. Instead, two elements are set up for discussion and considered in turn using the normal focussing process. Co-presence cases are well signalled in language behavior, perhaps to prevent confusion for hearer. Since hearers may be confused by single foci, it is not surprising that co-present foci should be signalled clearly enough so that some of the potential confusion is reduced. It may well be that the signalling is necessary for the speaker as well, to help keep track of what he or she is trying to say. This is mere speculation until focussing is applied to the

generation of language, and a theory of its behavior is given.

In contrast to co-present foci use of *this* and *that* noun phrases, non-co-present uses of them allow the speaker to indicate which of all the things she or he has mentioned is most important to the discussion. *This* and *that* used in non-co-presence allow the speaker to point at the relevant material with the least confusion. Hence the real difference in these uses is the difference of expectation in what will be talked about.

This chapter has supported the explanation of co-presence by stating and explaining a number of rules for the phenomena of *the one...the other* and *this-that* phrases. In addition, the use of the rules has been shown. These rules require no change in the focussing mechanism which was given in chapter 2. Not only is it desirable to retain the simplicity of the focussing mechanism as originally stated, but also the lack of change in the mechanism argues for it as a computational framework for explaining the comprehension of definite anaphora.

Finally, it must be said that the phenomena covered here are a tip of the iceberg. The relation between *this* and *that* as a reflection of the speaker's focus and pointing behavior remains to be explored. The kinds of questions to be asked include whether *this* marks an object in the speaker's and hearer's view, whether *that* marks something not in the hearer's view, whether *that* marks something in the speaker's view but of less concern, or whether some other behavior is happening as well. Furthermore, the rules presented in this chapter do not make it easy to understand if and why they give closure on the *this* and *that* phenomena. Yet the ball is rolling on *this* and *that*, for focussing behavior shows a surprising insight into the nature of these definite anaphora.

# 6. Some Experimental Systems Using Focus

## 6.1 Use of Focus in Discourse Understanding

Chapters 2, 3, 4 and 5 of this report have presented algorithms and rules for the use of focussing in discourse; no work would be complete without a discussion of the function of some of these rules in discourse understanding programs. In this chapter, two experimental natural language systems in which focussing has been used for definite anaphora resolution will be reported on. Two aspects of these systems demand particular attention. First, specific rules and parts of the algorithms of chapters 2, 3 and 4 have been implemented. Their behavior in a computational model provides further understanding of focussing as a process. Second, several different kinds of dialogues have been tested in the experimental systems. The effectiveness of focussing for these dialogues also provides a better understanding of the nature of discourse. This chapter will also report on a supplementary issue, which has been noted in chapter 2 and more extensively described in chapter 4, that is, the need for higher level control in interpreting these dialogues. This issue must be considered to explain how focussing fits into the overall problem of discourse understanding. Briefly, while it has been shown that focus is a necessary part, it is only one of several processes used in discourse understanding. This chapter will attempt to show what other processes are needed to guide the process of focussing.

The two systems which will be discussed are the Personal Assistant Language Understanding Program, called PAL, built as part of the Personal Assistant project at the Artificial Intelligence Laboratory at the Massachusetts Institute of Technology; and the Task Dialogue Understanding System (TDUS) built at the Artificial Intelligence Center at SRI International. Focussing mechanisms have been implemented for each of these systems. The systems are designed to allow interaction about a specific set of problems between a computer and a person; the interaction takes place in a discourse of several sentences. In each case, the computer is being requested to perform some set of actions related to the problem domain. While the problems are different, some similarities in the manner in which the actions are discussed will be shown. One fundamental difference between these systems is the role of the computer in the task. In PAL, the computer is an assistant, that is, it takes orders from a person about what to do and is generally thought of as simply assisting the system user. In the TDUS, the

computer is the expert and the user an apprentice. The user takes orders from the system, asks questions and provides information about how s/he is succeeding with the task. In other words, PAL assists a user and accepts the demands of that individual, while TDUS demands both certain actions to be performed and information about how the individual is performing a task. This difference does not bear in any direct way upon the use of anaphora: but the reader will see that TDUS discourses have other features which do affect the use of anaphora.

## 6.2 Focussing in the PAL System

PAL, the Personal Assistant Language Understanding Program, is designed to understand the English (type written) form of requests for arranging various events. The user is expected to tell PAL all the details s/he wants concerning some event, using one or more sentences of freely chosen English. PAL then attempts to inform a PA scheduler what to schedule, what people are supposed to participate, and so on, so that the scheduler takes into account other activities on the participants' calendars that may conflict. To understand a discourse, PAL must have several natural language skills:

a. parsing for the syntactic structure.

b. interpretation of predicate-argument relations.

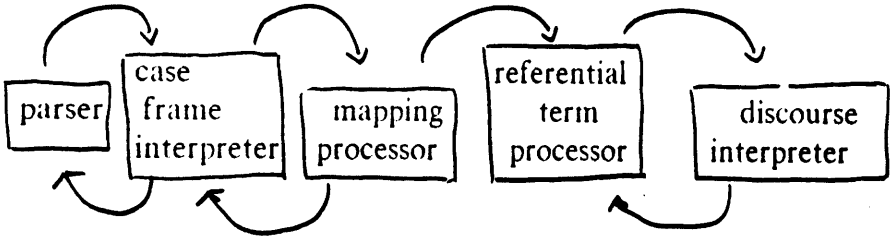
c. mapping of the words of each sentence to a representation used by the underlying database and programs.

d. disambiguation of the referential terms.

e. interpretation of each sentence for its discourse purpose.

Each of these skills is defined by a module of PAL as shown in figure 1. This figure also indicates the flow of control and data between the modules. Using the FRL language [Roberts and Goldstein, 1977], the mapping processor creates data structures called frames. The processor of PAL called the PA scheduler creates events and schedules them using these frame representations of the sentences. It accepts commands to schedule events at particular times and places for particular individuals from the

**figure 6.1. Modules of PAL**




---

discourse interpreter.[1]

An example of the PAL system in operation may be given using the previously given discourse:

D1-1 I want to schedule a meeting with Ira.
2 It should be at 3 pm on Thursday.
3 We can meet in his office.
4 Invite Bruce.

The first three modules of PAL turn D1-1 into several frames, one a type of **schedule,** one a type of **meeting,** one a type of **name** with a pointer to the string *Ira,* one a type of referential expression with a pointer to the string *I.* The referential component of PAL resolves the referring expressions of *I, a meeting* and *Ira.* Since no focus exists at the start of the dialogue, it uses various heuristics to determine the specification of names and indefinites and the co-specification of first person pronouns.

The discourse interpreter of PAL has two tasks. First it activates the focussing mechanism which creates an expected focus, and a list of default expected foci. PAL uses a very simple algorithm for choosing expected focus. Using information from

---

1. The implementation details are discussed in Sidner [1978]. The reader who wishes to know about how frames are created and to see some sample frames is directed to that paper.

the case frame interpreter, it chooses the first filled case of the verb from a list of ordered cases. This method is somewhat like the algorithm of chapter 2 for finding expected focus although the semantics underlying PAL were not rich enough to insure specifying a theme for every verb. The method also does not take into account marked syntactic forms. It proved adequate for the kinds of sentences in PAL discourses although it was clear that a more general treatment, such as the one given in chapter 2, was needed.

The discourse interpreter also has the task of determining the purpose of a given sentence in the discourse. The discourse purpose of a sentence is the request or other purposive behavior which the speaker is trying to elicit from the hearer with the sentence. The discourse purpose may be nothing more than "add this fact to the others I've told you," especially in the case of story telling. Or it may be a more subtle demand, such as the demand to close the window by saying "It's cold in here." Discourse purposes in PAL are a formulation of direct speech acts (see Searle [1969]). PAL's discourse interpreter represents an attempt to specify computationally the discourse purposes of sentences of the speaker. Another system which analyzes discourse purposes is described in the work of Bobrow et al [1977]. However, that system is designed so that the computer controls the discourse. The computer questions the user, so that most of the user's contributions to the discourse have the simple discourse purpose of providing answers to questions.

For D1-1, the discourse interpreter must determine its discourse purpose. While this may seem obvious to the reader, for PAL, the purpose is not obvious because the sentence is actually a request and not just an inform; the indirect request must be computed. PAL has several general schemas for the kinds of general requests which the scheduler performs; one of these is SCHEDULE. The schema is recognized as a sentence that is either an imperative, or declarative indicating desired action. Since D1 is a case of a declarative with a modal indicating desired action, the discourse interpreter decides that what is being requested is an act of scheduling and in particular, scheduling a meeting among certain people. This is what it indicates to the PA scheduling program.

The succeeding sentences of D1 create additional frames which PAL interprets as asserting the time and place of the meeting. The focus of the discourse is confirmed

in D1-2 as the meeting and re-confirmed in D1-3 because of the assertion of *meet*, a nominalization of the focus. PAL does not distinguish nominalizations in the straightforward way discussed in chapter 2 because it was unclear what the behavior of nominalizations was at the time that PAL was built; instead, it uses a combination of checks on valid discourse purposes and associations to focus.

The last sentence of D1 is a request which is taken to be about the meeting because people are invited to events, and because the event is left unspecified in the sentence. The request of D1-4 is interpreted as being a part of the scheduling operation although invitations in other contexts may be requested without scheduling of events. Certain requests such as INVITE may be interpreted in several ways because PAL uses general schemas to model discourse requests. PAL uses such schemas to interpret declaratives such as "The time should be 3 pm" as assertions about the time of the meeting it has been asked to schedule. Once a scheduling request has been made, as is done by the first sentence of D1-1, assertions and commands which follow are taken as providing additional constraints on the request for scheduling. In D1-4 inviting is taken as an additional constraint on the membership of the meeting. By contrast, when no other schema is in operation, INVITE, which is also a schema, is taken as a simple action of its own.

Focussing helps PAL understand the referential expressions of D1. While simple heuristics might have been used for all these cases of anaphora, focussing is a unified technique for approaching anaphora in PAL. All the pronoun anaphora resolution uses the focus either as co-specification or as source of a co-specification. The focus mechanism also provides the event left unmentioned in D1-4 in a manner like step 8 of the focussing algorithm.

In general, PAL dialogues use the explicit backwards co-specification rules and the associated co-specification rules of chapter 3 for anaphoric defnps. PAL also computes specifications for defnps and names which refer outside the discourse using the context growing technique discussed in chapter 3. For pronouns, PAL uses the focussing strategies to determine co-specification with focus (including the use of the recency rule discussed in chapter 4). The most complex rules used are for pronouns which co-specify a term associated with the focus. These associations are restricted to those which have been mentioned in the discourse. For example, while meetings have participants

associated with them, the participants cannot be pronominalized until some specific one is mentioned. In D1, after D1-1, Ira may be referred to as *he*, since he is a participant in the meeting and has been mentioned. The association rule also predicts ambiguity; if D1 continued as shown below, the use of *his* in D1-6 is ambiguous. The conversationally associated elements rule for possessive pronouns predicts this ambiguity.[1]

D1-5 We have a lot of topics to cover;
6 we also need to read his report.

The pronoun rules in PAL differ from those of chapter 4 because the focussing mechanism in PAL does not use an actor focus.

The referential term processor also includes a computation of Lasnik's precede and kommand rules for disjoint anaphora. Using this computation, all classes of sentential referential terms which must be disjoint from each other are removed from the classes of possibly co-specifying expressions. For those pronouns which cannot be resolved to focus or its associations, the computation of precede and kommand and sentential co-specifying classes is used. This computation determines that *his* in D2 co-specifies Josh but not in D3. Co-specifying classes function in a manner similar to the use of actor foci and the pronoun co-specification rules of chapter 4; the computation of co-specifying classes lacks the explanatory foundation given by the actor focus as well as the integrated treatment of actor and discourse focus of chapter 4.

D2-1 I am having a meeting with Sandra today.
2 Josh says we can use his office.

D3-1 I am having a meeting with Sam today.
2 Josh says we can use his office.

PAL is designed so that the focus mechanism is controlled by a higher level

---

1. Some readers may be surprised at this ambiguity since they will not find it ambiguous themselves. The reason for this behavior is that there are two interpretations, when spoken, for D1-6. One is unstressed, and the other stressed. The unstressed reading is captured by the rule above. The stressed reading is not, and must be accounted for in some as yet unspecified way. Here is another case where stress and prosodics are important indicators of anaphora resolution patterns which need to be accounted for.

process, the discourse interpreter. Why is this a necessary feature of a discourse interpreter? Consider the following sample of a PAL type dialogue with two alternate second sentences.

D4-1 I want to have a meeting with the head of our department.
2 (a) You can choose the time with Roger's help because he will be going too.
(b) You can pass the time with Roger's help because he will be going too.

D4-2b is somehow odd. However, strictly on the basis of focus, the oddity of the sentence cannot be recognized. The focussing mechanism only decides that *the time* designates the time of the meeting of D4 since it is possible to pass the time of an event with another person's help. Even more bizarre is:

D5-1 I want to have a meeting with my piano teacher.
2 a) Choose the place for me.
b) Eat at the place for me.

In these examples, the oddity should not be predicted by the inference mechanism which operates as part of the anaphora rules. When the sentence D5-2b is considered in isolation, the inference mechanism would decide that it is perfectly possible to eat at the place of a meeting with any person, including piano teachers. What is odd is the request in the context at hand. Focussing predictions are sensible only when the sentence has a discourse purpose that is acceptable for the task specified.

The need for higher level control in PAL is provided by the discourse interpreter, which guides the activation of the focussing mechanism. It restricts the kinds of sentences which are generally applicable in the context of PAL schedules. However, guiding the activation of the focussing mechanism is often not enough. PAL is designed so that the referential term processor activates before the discourse interpreter. This introduces a bit of bottom-up behavior into what is otherwise a top-down process. However, that design may be flawed. Using the focus, the referential term processor applies rules for backwards co-specification; it fails to notice the oddity of dialogues such as D5-2b because it will predict that *the place* means the place of the meeting. People may actually settle on this as the interpretation of D5-2b. What is questionable is whether people choose the interpretation of *the place* in the same manner as PAL. PAL interprets the noun phrase and then proceeds to analyze the discourse purpose of the sentence. There is another alternative: begin analysis of the discourse purpose during which the noun phrase is interpreted. If noun phrase interpretation

results in a discourse purpose which is bizarre, the noun phrase interpretation should be rejected. The difference here is a subtle one and which behavior actually occurs in human language is not clear. However, the discussion in the next section will indicate that a top-down control may be necessary.

To summarize, the focus mechanism in PAL uses only the discourse focus and its related anaphora rules, including the recency rule. It chooses expected focus with a simplified version of the expected focus algorithm, and it handles focus movement in the manner of the focussing algorithm. Because the focus mechanism of PAL has no actor focus, it uses some sentential rules in conjunction with precede and kommand to determine the co-specification of pronouns which fail to follow the focus rules. The focus mechanism is activated by the discourse interpreter which indicates whether the sentence and its co-specifications are valid discourse requests.

## 6.3 Focussing in the Task Dialogue Understanding System

The Task Dialogue Understanding System (see Robinson [1978]) simulates the behavior of a person who is an expert at assembling a small appliance. The user of the Task Dialogue Understanding System, TDUS hereafter, is assumed to be an apprentice who is learning how to assemble some or all of the appliance with help from the expert. The TDUS tells the user what parts to assemble in what order and answers questions about what tools are needed for the job and how to perform the steps of the task. As indicated in the discussion of Grosz' work in chapter 1, the TDUS uses a task model to guide its commands to the user and to understand what the user asks about the task, as well as to understand what the user says about completing various parts of the task. TDUS relies on a global focussing mechanism, built by Grosz. It determines the co-specifications for anaphoric defnps and the specifications of other referring defnps. Since natural user-machine communication includes uses of pronouns, TDUS needed a mechanism for pronoun interpretation, which could provide some unusual capabilities. This section will explain how focussing as described in this report was used to provide that capability. "Focussing" will be meant in a more immediate sense than the global mechanism created by Grosz. The two are integrated in TDUS because the local focussing mechanism inherits information from the global focus during the anaphoric interpretation process.

Two different kinds of pronoun anaphora occur in TDUS dialogues. These anaphora are found in the dialogues of native speakers. Sample behaviors have been observed in transcripts collected by Grosz [Deutsch, 1974] of dialogues between two people, one an expert and the other an apprentice who is learning to assemble an air compressor. Pronouns are used to co-specify with the current focus of discussion as in D6 below:

D6-1 A: I haven't encountered any problems yet.
2 A: Removing the pump seems very straightforward.
3 E: Have you disconnected **the air line** first?
4 A: I loosened **it**; I didn't disconnect **it**.

Pronouns are also used to co-specify with some focus much further back in the conversation:

D7-1 A: **One of the bolts** is stuck and I'm trying to use both the pliers and the wrench to get it unstuck.
2 E: Don't use the pliers. Show me what you are doing. Show me the 1/2" combination wrench.
3 A: Ok
4 E: Show me the 1/2" box wrench.
5 A: I already got **it** loosened.

The *it* in D7 co-specifies the bolt mentioned in the first sentence of the dialogue. Intervening between *one of the bolts* and *it* are other objects which could be co-specified by *it*. However, the native speaker interprets *it* as co-specifying the bolt without ambiguity. Surprisingly, once this behavior is isolated, it is not limited to task dialogues. The same behavior may be produced within a story context such as the one below:

D8-1 The other day I had a problem with my radio
2 because the speaker made a buzzing noise.
3 I decided to take it to be fixed.
4 Then Jane called to tell me that Ronald was in town.
5 I made plans to have her bring him to dinner.
6 Later Lily came by to look at my new car.
7 I never did get a chance to take it to be fixed.

The kind of focussing which occurs for D7 and D8 is not simple focus movement. It is more readily described as a focus shift. This shifting is a more abrupt means of changing focus back to some element previously discussed. It is marked by certain kinds of discourse behavior, other than the simple introduction of a term which is focussed upon in the next sentence or the use of an anaphor which cannot co-specify any object currently in focus. The difficulty of shifting of focus may be stated simply:

how is focus shifting to be distinguished from focus movement? How is focussing used to interpret these cases? If focussing is used, is any additional machinery needed?

Since many examples similar to D7 occur in the transcripts which Grosz collected, these have been used to discover the nature of the behavior. Task dialogues are structured according to the goals of the task which the apprentice is trying to complete. For example, when the apprentice notes that a bolt is stuck, it is a comment that takes place as part of disassembling the compressor, the sub-task the apprentice has been requested to do. The apprentice's comments are also indications of new sub-tasks. When the apprentice says that a bolt is stuck and then says that she is trying to loosen the bolt with some tools, she indicates a sub-task she has set up for herself. This sub-task is not a usual part of the disassembly, thought it certainly might happen. Furthermore, the apprentice indicates the sub-task is completed when she states that the bolt is loosened.

These observations permit a conclusion about task dialogue behavior; the speaker introduces goals through conversation, and the speaker announces the completion of a goal. Some of these goals are those typically associated with performing the task, while others represent unusual or new situations which occur as a result of the particular skills of the apprentice. The introduction and completion of goals through linguistic announcements must be recognized if a computational model is to interpret dialogues such as D7.

There are two ways in which goals are recognized as undertaken and as completed. One is to use a task model such as the one associated with TDUS. The task model in TDUS represents the typical steps in the assembly of a compressor, including which steps of the task must be done before others, and which steps may be done in any order. TDUS is also designed so that if the apprentice indicates that s/he has completed a step B, which had to be preceded by step A, then TDUS assumes step A was done as well. Using a task model, when the expert commands the apprentice to perform some step, and the apprentice announces that s/he is doing that step, then goal introduction could be recognized.

Recognition of both the introduction and completion of goals is also accomplished by a second means, by recognizing how language structure and form are

used to signal the goals of a discourse. In D7, a declarative sentence describing a state announces the apprentice's goal while another declarative of the same form indicates completion. Several other forms commonly appear as well. D9 below exemplifies the kinds of discourse structure reflected in language structure and form.

D9-1 E: Bolt the pump to the base plate.
2 A: What do I use?
3 E: There are 4 bolts, 4 nuts and 4 washers.
4 A: Can I take off the back plate?
5 E: No. Are you having trouble with the bolts?
6 E: Use the ratchet wrench, on the top to hold the bolt and hold the nut stationary on the bottom with a box wrench.
7 A: What is a ratchet wrench?
8 E: Show me the table. The ratchet wrench is the object between the wheel puller and the box wrenches on the table. Show it to me.
9 A: It is bolted. What do I do now?

In task dialogues, commands in the imperative mood indicate goals which the expert wants completed (D9-1) while goals initiated by the appentice use a state description. Both the expert and the apprentice use questions to specify sub-tasks which are unique to the situation (D9-2) or to indicate a problem with the sub-task which may require another sub-task to solve (D9-5). Most answers to questions in the declarative usually provide some information for completing a sub-task (D9-8) while denials such as "no" or "don't do that" deny the sub-task suggested. The completion of goals is indicated by declaratives which are passives and pseudo-passives (D7-5 and D9-9). Also the use of "nevermind" and "ok" indicate completion of a goal.

Using these observations, a goal interpretation component was designed for TDUS. It contained a few rules for the recognition of goal introduction and completion based on language signals. These are summarized in the list below. The list is by no means complete, but it offered a starting point for a description of discourse goals which could be used to determine focus shifting.

1. If a command is used, the goal is the command given.

2. Indirect commands of the form below indicate a goal, but modals are not incorporated in the system:
   a. I want you to <verb phrase> --> goal of <verb phrase>
   b. You should <verb phrase> --> goal is <verb phrase>

3. An assertion of the forms below has goal of <verb phrase>:
   a. I try to <verb phrase>
   b. I can {not, only, still} <verb phrase>
   c. I want to <verb phrase>
   d. The x is <adjective>. --> change-state x

4. Question of the form "how do I <verb phrase>?" has goal of <verb phrase>.

5. Question of the form "where is <noun phrase>?" has goal FIND <noun phrase>.

6. There-insertions introduce a new focus.

7. Goals are recognized in the following five forms:
   a. I have finished, am done.
   b. I have <verb> the <noun phrase>.
   c. {The, a, <null>} <noun phrase> is <verb>.
   d. okay
   e. nevermind

The goal interpretation module of TDUS uses the above rules to recognize whether a sentence contains a new goal or completes a goal already commanded by the speaker or indicated by the hearer. Focus processing proceeds in a top-down manner based on whether a new goal was seen, an old one completed or the current goal continued. For each of these three cases, a focus is found.

The simplest of the three alternatives is the continuation of the current goal. In such a case, focussing precedes in a fashion similar to the focussing algorithm. The process is not identical to the focussing algorithm because there are actually two algorithms, one to confirm the expected focus and one to move the focus. The similarity in confirmation and movement was not observed by the author until after the process was implemented. When a new goal is seen, the current goal and the focus are stacked away for later use; it is expected that eventually the new goal will be met and the old goal reinstantiated. The most complex behavior results from noticing the completion of a goal, particularly one on the stack. In this case, the old goal environment is re-instated, and the anaphora are interpreted relative to the focus of that environment. With this reinstatement, anaphora are never compared first to an inappropriate environment and then to the previous goal environment. Instead just the

previous goal environment is used.

The recognition of an old goal and the reinstatement of that goal environment is exemplified in D10, a simplified version of D9. The initial goal is given by D10-1. It is stacked in favor of the goal suggested by the apprentice's question D10-2 which is completed by the expert's answer. Another sub-goal is then introduced in D10-4, completed by the expert's answer, and this behavior is repeated with goal completion given in D10-8. *Okay* indicates the completion of some task, but it is not a clear enough statement to indicate which goal. D10-9b provides the environment; the goal interpreter reinstates the goal of bolting the pump from D10-1 and decides that it is completed because of the form of the sentence. When that goal is reinstated, the focus of D10-1 is re-established as well. It is the focus of pump which is used as the co-specification of *it* in D10-9b.

D10-1 E: Bolt the pump to the platform.
2 A: Where are the bolts?
3 E: They are in the tool box.
4 A: What tools should I use?
5 E: The ratchet wrench.
6 A: Where is it?
7 E: It's on the table.
8 A: I found it.
9 (a) A: Okay.
(b) A: It is bolted.
(c) A: What should I do now?

There are difficulties in trying to determine the proper environment to instantiate as the result of goal completion. Sometimes the verb alone may not be sufficient to re-instate a goal. Consider the example of D11.

D11-1 E: Tighten the pump.
2 E: To do so, tighten the bolts.
3 A: The pump is tightened.

D11-1 sets up tightening as a goal with the focus of *the pump*. D11-2 resets the focus (because of the *do-so* anaphor), so pump is stacked. D11-2 also introduces a second goal of tighten, with its own focus on bolts. D11-3, the apprentice's response to commands, indicates the goal of D11-1 is completed, and because tightening the bolts is a sub-task, that it is completed also. However, in order for TDUS to choose among the two goals, both the verb and the noun phrase are needed. Focus shifting back to a

previous goal is accomplished by use of both the command designated by the verb and the noun phrase associated with that verb. The goal interpreter in TDUS matches first on verb centered commands, but it also uses the noun phrase to choose the proper command in cases such as D11.

TDUS might be even more powerful if the goal interpretation component took advantage of what the task model provided about the tasks which needed to be performed. For example suppose that the task model for bolting a pump to a platform included the subtask of locating bolts and locating tools. Then in D10 when A asks where the bolts are, TDUS could use the task model to be sure that the question is part of the task of locating the bolts. Taking advantage of the task model in goal interpretation would avoid the confusion TDUS might reach in answering the question below:

D12-1 E: Bolt the pump to the platform.
2 A: Where is the hose?

While a hose is part of a pump, in the bolting operation, the hose is irrelevant. TDUS should respond not with its location but with information indicating that finding it is inappropriate, and with deletion of the goal associated with the question from the goal stack. At present, TDUS does not use sub-tasks in the model in this way.

Use of the task model of TDUS also reduces the system's dependency on the rules predicting goal introduction and completion. The seven rules presented previously are clearly general heuristics. Their use alone may indicate the apprentice's preceived goals, but not in a totally reliable way. By combining the use of the rules with the task model, TDUS could avoid potential errors resulting from the heuristics. For example, heuristic rule 3d states that a sentence such as "the X is <adjective>" indicates a goal of changing the state described by <adjective> for X. However, in some cases this heuristic is not appropriate:

D13-1 E: Take off the bolts.
2 A: I am loosening them with the pliers.
3 A: These pliers are really expensive looking.
4 A: I guess I should be careful with them.

Here the apprentice is not suggesting that s/he should make the pliers more or less expensive looking; the comment is kind of chit-chat. Only certain adjectives such as "broken, loose, too big" or the like actually indicate that a tool is not useful and must

be changed in some way. By comparing the predicted goal of the heuristics to possible goals for items which are unusable, TDUS could decide that the goal interpretation D13-3 is outside the normal range of goals and may be due to the mislabelling of the sentence by heuristic rules.

TDUS and PAL discourse bear some resemblance. The PAL discourses discussed previously are a version of TDUS discourse which is only one goal deep versions, because PAL interprets one goal and instantiates all parts of it. As a result, in PAL focus movement is most common, but a focus shift caused by a statement of goal introduction or completion does not often occur. A similar behavior is possible in a PAL-like discourse where the second *it* co-specifies with meeting and not with the time:

D14-1 U: Schedule a meeting with John.
2 U: Set the time for 3 pm.
3 U: If it's a bad time for him, try Tuesday.
4 PAL: I've scheduled it for Tuesday at 9.

Like TDUS, if PAL is to understand U and produce a shift such as that in D14-4, PAL must contain a discourse controller which recognizes embedded discourse purposes just as TDUS recognizes embedded discourse goals.

## 6.4 Conclusions

Focussing provides a significant advantage for language systems which seek to understand extended discourse. It provides for each goal environment, or discourse purpose environment, a focus of the discourse. When speakers then shift their attention to a previous goal or purpose in the discourse, the focussing mechanism provides the connecting concept which was under discourse there. Thus focussing provides the stack of previously discussed foci. When the discourse interpreter decides to which goal in the conversation the speaker is returning, the focus of that part of the conversation offers the co-specification for the anaphora which are often used in the sentences which indicate the return. Previous natural language understanding systems offer no machinery for interpreting such cases since they have no means for indicating the focus of discussion.

The direction for further work in discourse understanding systems is clear. Higher level controls need to be understood through research and experimentation.

Scheduling and task dialogues offer instances of a variety of behaviors which require control of a higher level than just the focussing mechanisms. The example of D15 in chapter 4 provides another important and intriguing behavior. Focussing is an important step because it delineates what the speaker is concerned about within a discourse environment. Focussing makes a clear set of predictions about which elements of the discourse the speaker can mention, but its purpose in discourse is limited to capturing the focus of the discourse discussion. The use of higher discourse control in PAL and TDUS shows how discourse control guides the use of the focussing mechanism. This guidance is necessary to determine whether the relationships between the anaphoric phrase and the rest of the sentence are sensible ones, not from the perspective of what is semantically understandable in a given sentence, but from the view of what is a sensible comment within the context of the ongoing discourse. Focussing predicts what the comments are about, but cannot determine whether those comments are reasonable ones. Only a higher level of control may decide how to use the focus predictions.

# 7. Summary and Conclusions

## 7.1 The Relation of Focussing to Anaphora Comprehension

At the beginning of this report, five claims were stated concerning the relation of focussing to definite anaphora comprehension. These claims support the hypothesis of this work, namely, that definite anaphora are signals which the speaker uses to tell the hearer what element in the discourse is the current discourse focus; at the same time, the element in focus constrains which anaphoric expressions are used to signal the focus. The five claims may now be summarized.

First, focussing provides a means for distinguishing definite noun phrases used anaphorically from those used non-anaphorically. In particular the focus and its associations sort definite noun phrases into three categories, those which have co-specifications in the discourse, those which specify some element associated with the discourse, and those which specify some element outside the discourse.

Second, focussing distinguishes pragmatic anaphora from bound variable and inter-sentential anaphora, using in addition sentence syntactic and semantic information on disjoint reference. Focussing does not indicate what the syntactic or semantic information from a sentence must be, but it does differentially use this information in determining the co-specification of an anaphor.

Third, the focussing mechanism provides a structure for writing and a machine for running rules which determine the co-specifications of pragmatic anaphora. The rules for definite pragmatic anaphora have been exhibited and discussed in the previous chapters, and the focussing machine's behavior has been demonstrated on numerous examples. The simplicity of the rules as well as their success in modelling many different anaphoric uses argues for the focussing theory of anaphora comprehension.

Fourth, focussing reduces the search for inferences which support the predicted co-specification chosen for an anaphor. Previous computational theories have proposed extensive inferencing for finding the co-specifications of anaphora. With focussing, inferencing is reduced, partly because syntactic and semantic criteria eliminate some predictions, but largely because inferencing is controlled by the focussing predictions.

The focussing mechanism reduces inferences to those that directly support a focus prediction, rather than producing the co-specification of an anaphor as a by-product of large scale general inferencing. Control is important not only for reducing search, but also because it may make it possible for researchers to use inference strategies, such as supposition, which heretofore appeared combinatorially infeasible.

Fifth, the data structure representing the element in focus indicates what items may be associated with the focus and what phrases are used to mention those items. By choice of an element in focus, certain noun phrases are used to mention items which are intuitively linked with focus while other noun phrases cannot be so used. That only certain noun phrases are acceptable indicates what the data structure representing the element in focus must include. Previous chapters show that generics, instances, and prototypes of generics, all linked with inheritance, must be part of that structure. In addition, certain common, limited associations among elements must be represented in the data structure. Furthermore, the data structure must permit three kinds of computations: relative sequence computation, forward inference and supposition. The data structure must also allow for attribution, that is, a means of indicating that a speaker thinks a definite noun phrase has a specification without stipulating which one it is. Finally, since others have shown the need for representing scope of quantification and bound variables, previous chapters have shown how to taken advantage of this information when available to the focussing mechanism.

The evidence for these claims not only supports the hypothesis of the signalling relation between an element in focus and an anaphor; it also enriches what is known about the nature of definite anaphora in English discourse by providing a process model of discourse comprehension. With this computational description, focussing via the focus mechanism becomes a tool for furthering research in comprehension.

What focussing does not include must be stated clearly. Three general topics mentioned during this report are not part of focussing: exophoric reference, that is, reference to objects which the speaker and/or hearer perceive by seeing or hearing; task models or other knowledge structures which represent a complex sequence of events or goals; and discourse purposes, that is, the speaker's intent in uttering a sentence in a discourse. These topics have been mentioned in this report because they link focussing to other aspects of discourse comprehension.

## 7.2 Limitations of Focussing in Anaphora Comprehension

In previous chapters the focussing mechanism is shown to give some counter-intuitive results on certain examples, those which have been characterized as involving parallelism. As was pointed out, there may be some limits on the kinds of discourses which may be understood in normal comprehension when parallelism is involved. Yet clearly hearers easily understand discourses with a limited kind of parallelism. In particular, they take advantage of the parallel structure between two successive sentences of a discourse in comprehending definite anaphora. Focussing cannot account for the detection of parallel structure, not only because the computation of such structure is poorly understood, but also because focussing chooses different defaults for co-specification than those required for parallelism.

That focussing cannot account for definite anaphora which occur in parallel structure does not argue against focussing as a means of explaining anaphora comprehension. On the contrary, focussing explains a characteristic use of anaphora while parallelism characterizes an anaphoric behavior which has different properties from the anaphora which depend on focus. The focussing and parallel uses are fundamentally different because parallel structure plays no role, as far as I can tell, in the uses governed by focussing. The comprehension of definite anaphora which relies on parallelism falls outside of focussing, and some mechanism governing their behavior remains to be discovered.

Focussing may be an incomplete explanation for another anaphora phenomenon, pronouns without co-specifications. These pronoun uses are puzzling because some, such as D1 below, are easily understood by hearers, while others, such as D2, are not.

D1-1 I went to a concert last night.
2 They played Beethoven's Ninth.

D2-1 John is an orphan.
2 He misses them very much.

From focussing behavior one might conclude that for such cases, the hearer searches among elements associated with the expected focus for an element which acts like a co-specification. What is unexplained is why John's parents, the specification of *them*, is more difficult to find than the orchestra, the specification of *they*. Focussing behavior

at least suggests that the associations of any given element must encode some associations differently than others.

## 7.3 Further research

There are five basic directions of research that are suggested by this report. First, verb phrase and indefinite anaphora are treated only peripherally in this work. The do-*so* and do-*it* cases are only briefly considered, and other verb phrase anaphora (see Webber [1978] for a complete list) left unexplored. In addition, *one* and *such* anaphora have not been considered at all in this work. There is reason to believe that focussing may be useful for these types of anaphora as well; the immediate questions to raise are just how the use of focussing and the focus mechanism may be applied, what rules are produced, and what constraints arise from syntax and semantics.

A second basic direction of research is the relationship of stress to anaphora and focussing. Throughout this report examples of stress differences have been mentioned. Some of these suggest that stress is a means of indicating what will be the focus in the initial part of a discourse; other examples seem to show that stress, particularly contrastive stress, indicates a shift of focus. These indications need to be clarified by careful research. Furthermore, stress as a computation needs to be better understood if it is to be used by the focussing mechanism. The focussing mechanism is constructed so that were stress information computable, it could be used just as syntactic and semantic information are.

Further understanding of generic anaphora constitutes a third direction of research. Chapter 3 develops some heuristic rules, but a full treatment of generics is needed. Chapter 4 shows that some generic use of pronouns is distinguished from non-generic use by a plural pronoun where a singular pronoun would co-specify with the focus. The examples below show that *this* and *that* also require a set of rules which determine whether the phrase is generic.[1] D3 suggests that generic interpretation of *that* noun phrases influences focussing since *that* co-specifies with the generic

---

1. Examples are due to Chafe [1975] and Stenning [1978].

Beethoven's Sixth while *it* co-specifies with the version played when the speaker was present. Stenning claims of the second example that the first use of ***this engine*** is specific while the second is generic.

D3-1 I went to a concert last night.
2 The orchestra played Beethoven's Sixth.
3 You don't hear that too often.
4 I enjoyed it very much.

D4-1 On my left is a reciprocating engine.
2 In this engine *a crankshaft* transmits the gas pressure to the thrumjit.
3 This engine powered **every car on the road**
4 until Herr Wankel suggested that *the crankshaft* was inelegant and invented his rotary engine.

A fourth direction for further research centers on the differences in use of referring expressions. Stenning [1978b] suggest that there are differences in the use of referring expressions in the different kinds of discourse, namely expository writing, stories and arguments. Most of the examples in this report are of expository text or of examples people say; a few examples are from stories. These examples corroborate Stenning's insight. They lead us to ask the following questions: In what way do anaphora have preferred or different functions in particular kinds of texts? What different assumptions does a speaker make about a hearer's knowledge in different discourse circumstances? Do the inferences which a hearer must make in understanding how an anaphor co-specifies with a focus indicate the difficulty of a text for the hearer? The answer to these questions will tell us more about the function of anaphora in language, and more about speaker-hearer knowledge and its role in language.

Fifth among the research directions is the use of focussing for language generation. The focussing mechanism discussed in this report interprets anaphora from the point of view of a hearer, that is, the mechanism seeks to understand how the speaker uses an anaphor to co-specify. By symmetry the focussing mechanism ought to be able to produce anaphora in the way that a speaker does. Generating discourses using the focussing mechanism and the associated constraints seems feasible; it is yet to be undertaken fully although McDonald [1978] reports on some preliminary use.

Finally some speculation on theories of pragmatics is warranted. In his well known William James lectures, Grice [1975] defined several maxims of conversation, one

of which was the maxim of relevance. Grice[1] says about this maxim:

> Under the category of *Relation* I place a single maxim, namely, "Be relevant." Though the maxim itself is terse, its formulation conceals a number of problems which exercise me a good deal; questions about what different kinds and foci of relevance there may be, how these shift in the course of a talk exchange, how to allow for the fact that subjects of conversation are legitimately changed, and so on.

As long as relevance is a part of a theory of pragmatics, focussing must be included in that theory, whether it is the theory which Grice has begun to unfold or some other one. Focussing and focus as they have been used in this report bear directly on Grice's concerns; for they suggest a means for carrying out the maxim of relevance. Namely, a speaker is speaking relevantly in a discourse if he or she introduces a focus, and proceeds to another one by mentioning it and re-mentioning it with definite anaphora. Old foci are re-invoked by a definite noun phrase which points out which old focus is co-specified or in one of the other, less direct, ways discussed in the previous chapters. Nothing less than the use of focus will suffice for relevance; for the moment the speaker fails to provide a focus for the hearer and to point back to it in successive utterances, the hearer has no means of knowing what is relevant in the discourse at hand. In some sense the discourse ceases to be a discourse.

Perhaps it is surprising that focussing should play such a central role in a theory of pragmatics. In particular, it is surprising that focussing one's attention on something and signalling one's focus is part of the criteria for speaking relevantly. One expects relevance to be a matter of what is said about some thing, rather than that the thing is mentioned consistently. But if we remember that focussing allows for the speaker to tell the hearer that the same thing is still under discussion and without needing to say explicitly what that thing is, then the role of focussing is not so surprising. Focussing must then be the first criterion for speaking relevantly, since it explains how a hearer decides what the speaker is talking about.

---

1. Grice, *op. cit.*, page 67.

## 8. List of Cited Works

Akmajian, A. [1973], *The Role of Focus in the Interpretation of Anaphoric Expressions.* in Anderson and Kiparsky (eds.), Festschriff for Morris Halle, Holt, Rinehart and Winston, New York, New Yqrk.

Allen, James [1978] *Recognizing Intention in Dialogue* Ph.D. thesis, Department of Computer Science, University of Toronto.

Austin, J.L. [1962] *How to Do Things With Words.* Oxford University Press.

Balzer, R., N. Goldman and D. Wilie [1977] *Information in Program Specification.* Proceedings of the Fifth International Joint Conference in Artificial Intelligence, August, pages 389-397.

Baranovsky, Sharon [1973] *Some heuristics for Automatic Detection and Resolution of Anaphora in Discourse.* M.S. thesis, Dept. of Computer Sciences, University of Texas, Austin.

Bobrow, D., R. Kaplan, M. Kay, D. Norman, H. Thompson, and T. Winograd, [1977] *GUS, A Frame-Driven Dialogue System.* Artificial Intelligence, Volume 8, Number 2, April.

Bobrow, D.G. and T. Winograd. [1977] *An Overview of KRL, A Knowledge Representation Language.* Cognitive Science 1.

Bullwinkle, C. [1975] *Picnics, Kittens and Wigs: Computer Performance of the Sentence Completion Task.* Proceedings of the Fourth International Joint Conference in Artificial Intelligence, August 1975.

Bullwinkle, C. [1977] *Levels of Compexity in Discourse for Anaphora Disambiguation and Speech Act Interpretation.* Proceedings of the Fifth International Joint Conference in Artificial Intelligence, August 1977; also M.I.T. A.I. Lab Memo 413.

Carpenter, Patricia A. and Marcel A. Just. [1977] *Integrative Processes in Comprehension.* in LaBerge and Samuels (eds.) Perception and Comprehension. Hillsdale, N.J., Lawrence Erlbaum Associates.

Carpenter, Patricia A. and Marcel A. Just. [1978] *Reading Comprehension as Eyes See It.* in Just and Carpenter (eds.) Cognitive Processes in Comprehension. Hillsdale, N.J., Lawrence Erlbaum Associates.

Chafe, W. [1976] *Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View.* in Li (ed.) Subject and Topic, Academic Press, New York.

Chafe, W. [1975] *Structures and Human Knowledge.* in J.B. Carrol and R.O. Freedle (eds.) Language Comprehension and the Acquisition of Knowledge, V.H. Winston and Sons.

Charniak, E. [1972] *Toward a Model Of Children's Story Comprehension.* M.I.T. A.I. Lab TR-266.

Chomsky, N. [1965] *Aspects of a Theory of Syntax.* M.I.T. Press, Cambridge, MA.

Chomsky, N. [1971]. *Deep Structure, Surface Structure and Semantic Interpretation.* in Steinberg and Jakobovitz (eds.) Semantics: An Interdisciplinary Reader in Philosophy, Linguistics and Psychology. Cambridge University Press, London.

Chomsky, N. [1976] *Conditions on Rules of Grammar.* Linguistic Analysis Volume 2, p. 303-351.

Clark, H.H. and Susan Haviland. [1977] *Comprehension and the Given-New Contract.* in Freedle (ed.) Discourse Production and Comprehension, Ablex Publishing Corporation, Norwood, N.J.

Clark, H.H. and Catherine Marshall. [1978] *Reference Diaries.* in Tinlap-2, ACM and ACL, New York.

Cohen, Philip R. [1978] *On Knowing What to Say: Planning Speech Acts.* University of Toronto Department of Computer Science, T.R. No. 118.

Collins, A., J. S. Brown, and K. Larkin. [1977] *Inference in Text Understanding.* BBN Tech report no. 40.

Donnellan, Keith S. [1977] *Reference and Definite Descriptions.* in Naming, Necessity, and Natural Kinds, S.P. Schwartz (ed.), Cornell University Press, Ithaca.

Deutsch, B. [1974] *Typescripts of Task Oriented Dialogs,* unpublished manuscript, SRI-International, Menlo Park, Ca.

Deutsch, B. [1975] *Establishing Context In Task-Oriented Dialogues.* Proceedings of the 13 Annual Meeting of ACL, AJCL Microfiche 35.

Doyle, Jon [1978] *Truth Maintainence Systems for Problem Solving.* M.I.T. A.I. Laboratory AI-TR-419, January.

Evans, Gareth [1977] *Pronouns, Quantifiers and Relative Caluses (1),* Canadian Journal of Philosophy, Volume VII, Number 3, September, pages 467-536.

Fahlman, Scott E. [1977] *A System for Representing and Using Real-World Knowledge.* M.I.T. A.I. Laboratory AI-TR-450, December.

Fillmore, C.F. [1971] *Lectures on Deixis.* Unpublished manuscript, Linguistics Department, University of California, Berkeley.

Firbas [1971] *On the Concept of Communicative Dynamism in the Theory of Functional Sentence Perspective.* Unpublished manuscript, Faculty of applied philosophy, Brnenske University, Czechloslovakia.

Fodor, J. [1978] *Methodologial Sollipcism.* Unpublished manuscript, Cambridge, MA.

Frege, Gottlob [1892] *Ueber Sinn und Bedeutung.* Zeitscrift fuer Philosophie und philosophsche Kritik, Vol. 100, pp. 25-50. English translation (*On Sense and Reference*) in Frege, Philosophical Writings.

Geach, Peter T. [1962] *Reference and Generality: An Examination of Some Medieval and Modern Theories.* Cornell University Press, Ithaca, New York.

Goldstein, I. P. and R. B. Roberts.[1977] *NUDGE, A Knowledge-based Scheduling Program.* M.I.T. A.I. Lab memo 405.

Goudge, Thomas A. [1969] *The Thought of C.S. Pierce.* Dover Publications, Inc., New York.

Grice, H.P. [1975] *Logic and Conversation.* in P. Cole and J. Morgan (eds.), Syntax and Semantics Vol. 3 Speech Acts. Academic Press, New York.

Grimes, Joseph [1975] *The Thread of Discourse.* Janua Linguarum Series, Mouton Press, The Hague.

Gruber, Jeffrey [1976] *Lexical Structure in Syntax and Semantics.* North Holland Publishers, New York.

Grosz, Barbara [1977] *The Representation and Use of Focus in Dialogue Understanding.* Stanford Research Institute Technical Note 151, Menlo Park, California.

Grosz, Barbara [1978] *Focussing in Dialog.* Tinlap-2: Theoretical Issues in Natural Language Processing, ACM and ACL, New York.

Halliday, M.A.K. [1967] *Notes on Transitivity and Theme in English.* Journal of Linguistics 3, pages 117-274.

Halliday, M.A.K. and Ruqaiya Hasan [1976] *Cohesion in English.* Longman Press,

London.

Hawkinson, L. [1975] *The Representation of Concepts in OWL.* Proceedings of the Fourth International Joint Conference in Artificial Intelligence, August 1975.

Hendrix, Gary L. [1975] *Partitioned Networks for the Mathematical Modeling of Natural Language Semantics.* Technical report NL-28, Department of Computer Science, University of Texas, Austin, Texas.

Hobbs, Jerry R. [1976] *Pronoun Resolution.* Research Report #76-1, City College, City University of New York, New York.

Hobbs, Jerry R. [1977] *38 Examples of Elusive Antecedents from Published Texts.* Research Report #76-2, City College, City University of New York, New York.

Hobbs, Jerry R. [1979] *Coherence and Co-reference.* Cognitive Science. Vol. 3, No. 1, pages 67-90.

Isner, Dale W. [1975] *Understanding "Understanding" Through Representation and Reasoning.* Unpublished dissertation, University of Pittsburgh, Department of Computer Science.

Jackendoff, Ray S. [1972] *Semantic Interpretation in Generative Grammar.* MIT Press, Cambridge, MA.

Karttunen, Lauri [1968] *What Makes Definite Noun Phrases Definite?* Paper p-3871, The RAND corporation, Santa Moncia, California.

Kripke, Saul A. [1972] *Naming and Necessity.* in Semantics of Natural Language, Davidson and Harman (eds.) Reidel Publishing Co., Boston.

Katz, Jerrold J. and Fodor, Jerry A. [1963] *The Structure of a Semantic Theory.* Language, Volume 39, Number 2.

Kuno, Susumo [1975] *Three Perspectives in the Functional Approach to Syntax.* in Grossman et al (eds.) Papers from the Parasession on Functionalism, Chicago Linguistics Society, Chicago, Ill.

Lakoff, Robin [1974] *Remarks on This and That.* Proceedings of the Tenth Regional Meeting of the Chicago Linguistic Society. Chicago, Ill.

Langacker, P. [1969] *On Pronouns and the Chain of Command.* in Reibel and Schane (eds.) Modern Studies in English, Prentice-Hall, Englewood Cliffs, N.J.

Lasnik, Howard [1976] *Remarks on Co-reference.* Linguistic Analysis, Volume 2, Number 1.

Lawler, J.M. [1972] *Studies in English Generics,* unpublished dissertation, Department of Linguistics, University of Michigan, Ann Arbor, MI.

Lockman, Abe D. [1978] *Contextual Reference Resolution in Natural Language Processing.* Dept. of Computer Science TR-70, Rutgers University, New Brunswick, N.J.

Lyons, J. [1968] *Introduction to Theoretical Linguistics.* Cambridge University Press, London.

Marcus, Mitchell [1977] *A Theory of Syntactic Recognition for Natural Language,* unpublished dissertation, Massachusettts Institute of Technology, Cambridge, MA.

Martin, William [in preparation] *Philosophical Foundations for a Linguistically Oriented Semantic Network,* Laboratory for Computer Science, M.I.T., Cambridge, MA.

McDonald, David [1978] *Subsequent Reference: Syntactic and Rhetorical Constraints.* Tinlap-2: Theoretical Issues in Natural Language Processing, ACM and ACL, New York.

Minsky, Marvin [1975] *A framework for Representing Knowledge.* in Winston (ed.) The Psychology of Compuer Vision, McGraw-Hill Book Co., New York.

Moore, Robert C. [forthcoming] *Reasoning about Knowledge and Action.* SRI-technical report. Menlo Park, California.

Norman, Donald A. and David E. Rumelhart [1975] *Explorations in Cognition.* W.H. Freeman, San Francisco.

Partee, Barbara H. [1972] *Opacity, Co-reference and Pronouns.* in Davidson and Harman (eds.) Semantics of Natural Language, D. Reidel Publishing Co., Boston.

Partee, Barbara H. [1978] *Bound Variables and Other Anaphors.* Tinlap-2: Theoretical Issues in Natural Language Processing, ACM and ACL, New York.

Prince, Ellen F. [1978] *On the Inferencing of Indefinite This Defnps,* unpublished manuscript, University of Pennsylvania.

Pople, Harry E. and Decision Systems Laboratory [1977] *The Formation of Composite Hypotheses in Diagnostic Problem Solving.* Proceedings of the Fifth International Joint Conference in Artificial Intelligence, August 1977.

Postal, P. [1969] *On the So-called Pronouns in English.* in Reibel and Schane (eds.) Modern Studies in English, Prentice-Hall, Englewood Cliffs, N.J.

Postal, P. [1969] *Anaphoric Islands,* Proceedings of the Fifth Meeting of the Chicago Linguistic Society. Chicago.

Putnam, Hilary. [1978] *Reference and Understanding.* in Meaning and the Moral Sciences. Routledge and Kegan Paul, Boston.

Quillian, M.R. [1968] *Semantic Memory.* in Minsky (ed.) Semantic Information Processing, M.I.T. Press, Cambridge, MA.

Quirk, Randolph and Sidney Greenbaum [1973]. *A Concise Grammar of Contempory English.* Harcourt, Brace and Jovanovich, Inc., New York.

Reichman, R. [1978] *Conversational Coherency.* CSR Technical Report No. 95, Bolt, Beranek and Newman, Inc., Cambridge, MA.

Reinhart, Tanya [1976] *The Syntactic Domain of Anaphora.* unpublished Ph.D. dissertation, Department of Foreign Literature and Linguistics, M.I.T.

Rieger, Charles J. [1974] *Conceptual Memory: A Theory and Computer Program for Processing the Meaning Content of Natural Language Utterances.* Stanford Artificial Intelligence Lab Memo AIM-233.

Roberts, R.B. and I.P. Goldstein [1977] *The FRL Manual.* AIM-409, Artificial Intelligence Lab, M.I.T., Cambridge, Ma.

Robinson, Ann [1978] *Investigating the Process of Natural Language Communication: A Status Report.* SRI Technical Note 165. SRI International, Menlo Park, Calif. July.

Ross, J.R. [1967] *Constraints on Variables in Syntax.* unpublished MIT doctoral dissertation, Cambridge, Massachusetts.

Rumelhart, D. [1975]. *Notes on a Schema for Stories.* in Bobrow and Collins (eds.) Representation and Understanding: Studies in Cognitive Science, Academic Press, Ne York.

Russell, Bertrand [1905] *On Denoting.* reprinted in Logic and Knowledge, R.W. March (ed.), Allen and Unwin, London, 1956.

Schank, R.C. et al. [1975] *SAM - A Story Understander.* Research Report 43, Yale University Department of Computer Science, New Haven, Connecticut.

Searle, J.R. [1969] *Speech Acts: An Essay in the Philoshpy of Language.* Cambridge University Press.

Sgall, Petr, Eva Hajicova, and Eva Benesova. [1973]. *Topic, Focus and Generative Semantics.* Scriptor Verlag, Democratic Republic of Germany.

Simmons, R.F. [1973] *Semantic Networks: Their Computation and Use for Understanding English Sentences.* in Schank and Colby (eds.) Computer Models of Thought and Language, W.H. Freeman and Company San Francisco.

Smith, Brian C. [1978] *Levels, Layers and Plans: The Framework of a Theory of Knowledge Representation,* unpublished Master's thesis, Dept. of Electrical Engineering and computer Science, M.I.T., Cambridge, MA.

Smith, Carlotta S. [1969] *Determiners and Relative Clauses in a Generative Grammar.* in Reibel and Schane (eds.) Modern Studies in English, Prentice-Hall, Englewood Cliffs, N.J.

Springston, F.J. [1975] *Some Cognitive Aspects of Presupposed Co-reference Anaphora.* unpublished dissertation, Stanford University, Stanford, CA.

Sidner, Candace L. [1978] *A Progess Report on the Discourse and Reference Components of PAL,* Proceedings of the Second National Conference of CSCSI, Toronto, July, 1978. Also MIT AI Lab Memo 468, April, 1978.

Stenning, Keith [1978a] *Anaphora as an Approach to Pragmatics.* in Halle, Bresnan, and Miller (eds.) Linguistic Theory and Psychological Reality, MIT Press, Cambridge, MA.

Stenning Keith [1978b] *Why Making Reference out of Sense Makes It So Hard to Make Sense of Reference.* unpublished manuscript presented for Sloan Foundation Workshop on Indefinite Reference, Amherst, MA, December.

Van Lehn, Kurt A. [1978] *Determining the Scope of English Quantifiers.* AI-TR-483, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, MA.

Vendler, Zeno [1967] *Linguistics in Philosophy.* Cornell University Press, Ithaca, N.Y.

Walker, Donald E. (ed.) [1976] *Speech Understanding Research.* Final Technical Report, Project 4762. Artificial Intelligence Center, Stanford Research Institute, Menlo Park, CA.

Wasow, Thomas [1972]. *Anaphoric Relations in English.* unpublished dissertation,

Massachusetts Institute of Technology, Cambridge, MA.

Webber, Bonnie Lynn [1978] *A Formal Approach to Discourse Anaphora.* Technical Report 3761, Bolt, Beranek and Newman, Cambridge MA.

Williams, Edwin S. [1977] *Discourse and Logical Form.* Linguistic Inquiry, Vol 8, (Winter, 1977), p. 101-139.

Winograd, Terry [1971] *Procedures as a Representation for Data in a Computer Program for Understanding Natural Language.* M.I.T. dissertation.

Woolfe, Virginia [1957] *A Room of One's Own.* Harcourt, Brace and Jovanovich, Inc., New York.

Woods, William [1973] *An Experimental Parsing System for Transition Network Grammars.* in Rustin (ed.) Natural Language Processing. Algorithmics Press, N.Y.

Woods, William et al [1976] *Speech Understanding Systems: Final Technical Progress Report - 30 October 1974 through 29 October 1976.* Technical, Report 3438, Bolt, Beranek and Newman Inc., Cambridge, MA.

## Appendix I - Algorithms

### The Expected Focus Algorithm:

Choose an expected focus as:

1. The subject of a sentence if the sentence is an *is-a* or a there-insertion sentence.
   This step presumes information from a parse tree about what the subject, and verb are and about whether the sentence is there-insertion.
2. The first element of the default expected focus list (DEF list), computed from the thematic relations of the verb, as follows:
   Order the set of phrases in the sentence using the following preference schema:
   theme unless the theme is a verb complement in which case the theme from the complement is used.
   all other thematic positions with the agent last
   the verb phrase
   This step requires a list of the surface order of the noun phrases, and a data structure which indicates which noun phrases fill which thematic slots in the verb. Such a data structure must be computed by a case frame mechanism like the one reported in Marcus [1977].

### The Focussing Algorithm

A. Steps 1-9 use an ALFL (alternative focus list). It is initiated to be either the DEF or PFL depending on whether the discourse is initial or in progress. A stack called the focus stack is globally available to this algorithm. On first use of this algorithm, the stack is empty.

B. Set the current focus (CF) to either the expected focus found from the expected focus algorithm or the focus of the discourse when discourse is in progress.

To confirm the current focus as focus or to reject the current focus for another focus in the next sentence of the discourse:

1. **DO-ANAPHORA**: If the sentence contains do-anaphora, take the last member of the ALFL as the focus. Ignore steps 2 through 6. Stack the current focus in the focus stack.

2. **FOCUS SET COLLECTION**: If there is no CF by the initialization in B above, there is an occurrence of focus sets. When no definite anaphora have appeared in the current sentence, continue collecting focus sets. If an anaphor appeared and it is not in agent position, take its co-specification as focus.

3. **CHOOSING BETWEEN CF and ALFL**: If there are anaphora which co-specify both the CF and some member of ALFL, take as focus whichever is not in agent position. If both are non-agents, retain the CF as focus unless only the ALFL member is mentioned by a pronoun. In that case, move the focus to the ALFL member. (Focus is moved by stacking the CF, setting the CF to the co-specification of the anaphoric term, and then stacking any flagged implicit specs as long as that spec is not the spec to which focus moves.

4. **RETAINING THE CF as FOCUS**: If there are anaphora which co-specify only the CF, retain the CF as focus.

5. **ALFL as FOCUS**: If the anaphora only co-specify a member of ALFL, move the focus to it. If several members of the ALFL are co-specified, choose the focus in the manner suggested by the expected focus algorithm.

6. **FOCUS STACK USE**: If the anaphora only co-specify a member of the focus stack, move the focus to the stack member by popping the stack.

7. **IMPLICIT SPECIFICATION**: If a defnp implicitly specifies an element associated with the focus, retain the CF and flag the defnp as implicit spec. If specification is associated with member of ALFL, move focus to that member and flag the defnp as implicit spec.

8. **LACK OF ANAPHORA**: If there are no anaphora co-specifying any of CF, ALFL or focus stack, but the CF can fill a non-obligatory case[1] in the sentence or if the verb phrase is related to the CF by nominalization, retain the CF.

---

1. Obligatory relations are cases of a verb that must be filled or the sentence is odd as in "John sold." Non-obligatory cases need not be filled; e.g. in "John sold a book," one non-obligatory case is the person to whom the book was sold.

9. **FOCUS SET INITIALIZATION**: If there are no foci mentioned and the sentence is discourse initial, collect focus sets.

10. **NO FOCUS USED**: Otherwise if there are no foci mentioned, retain the CF as focus. For any unspecified pronouns, the non-antecedent pronoun condition holds.[1]

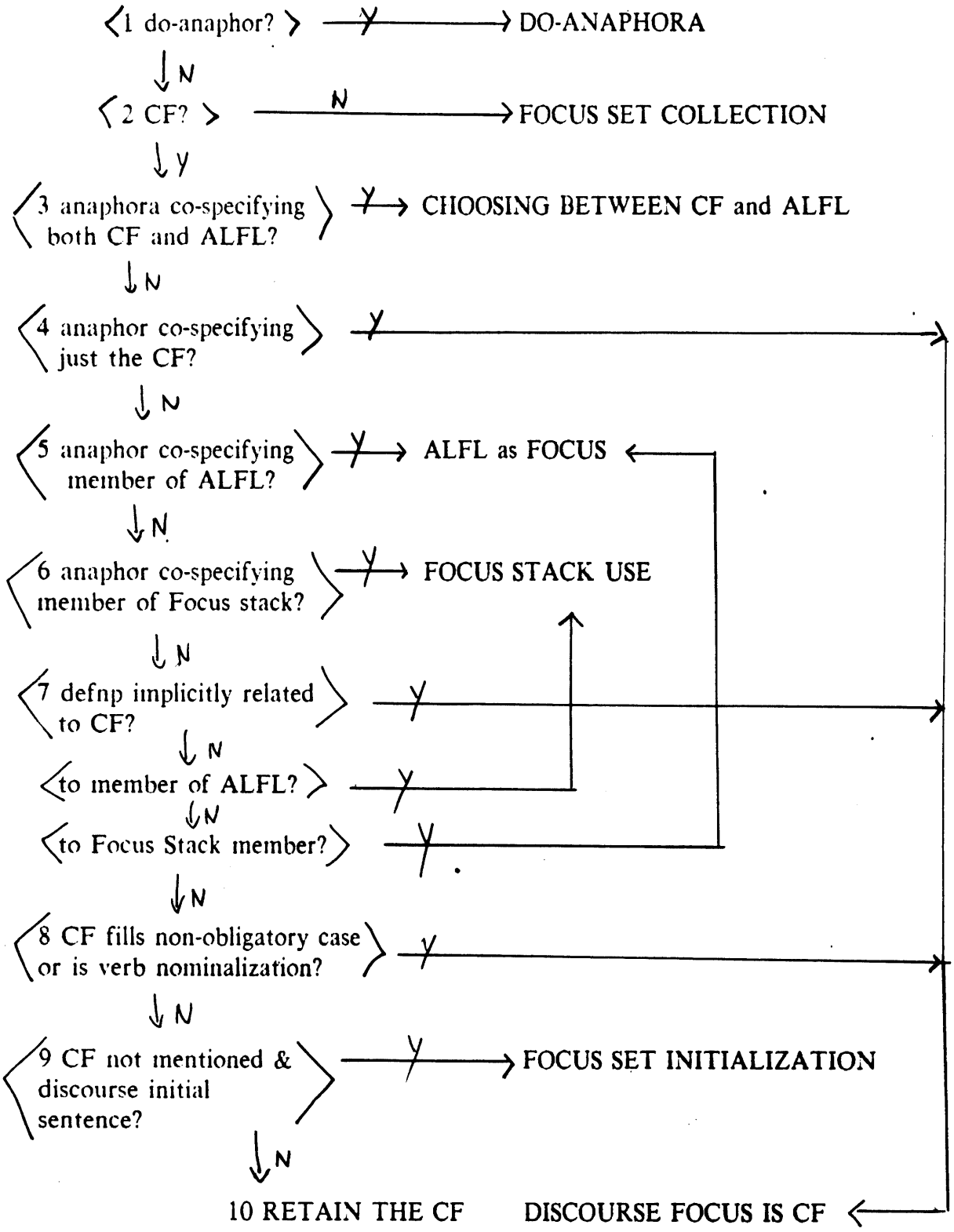
The algorithm for determining the potential focus list (PFL):

1. If a cleft or pseudocleft sentence is used, the potential focus is the cleft item iff the entity in non-clefting position co-specifies the focus. When is does not, the sentence is incoherent.
2. Otherwise order a potential focus list of all the noun phrases filling a thematic relation in the sentence, excluding a noun phrase in agent position and the noun phrase which co-specifies the focus if one exists. The last member of the PFL is the verb phrase of the sentence.

---

1. See Chapter 4 for a discussion of non-antecedent pronoun uses.

**figure I.1. Control Flow for Focussing Algorithm**



An algorithm for determining defnp function in discourse:

1) (Focus Relation) Given a defnp, it co-specifies with the focus if either of the explicit backwards co-specification rules hold. Otherwise the implicit backwards specification rules (except inferred specification) may be applied to determine its specification.

2) (Potential Foci Relation) If no rules apply, the explicit backwards co-specification and implicit backwards specification rules (except inferred specification) may be re-applied to the defnp using each of the available potential foci in place of the focus in the rules.

3) (Stacked Foci Relation) If no rules apply, the explicit backwards co-specification rules may be re-applied for the available stacked foci in place of the focus in the rules.

4) (Closure and Outside Specification) If no rules apply, for a defnp without modifiers, the inferred specification rule may be applied; if it fails to hold, the use is odd. For a defnp with ordinal modifiers, the use is odd while for other defnps containing modifiers, the defnp specifies outside the discourse.

5) (Focus set Use) When no focus has been established, a defnp which lexically generalizes one of the focus sets co-specifies with that set.

**figure I.2. Rule for Third Person Pronoun in Agent Position**

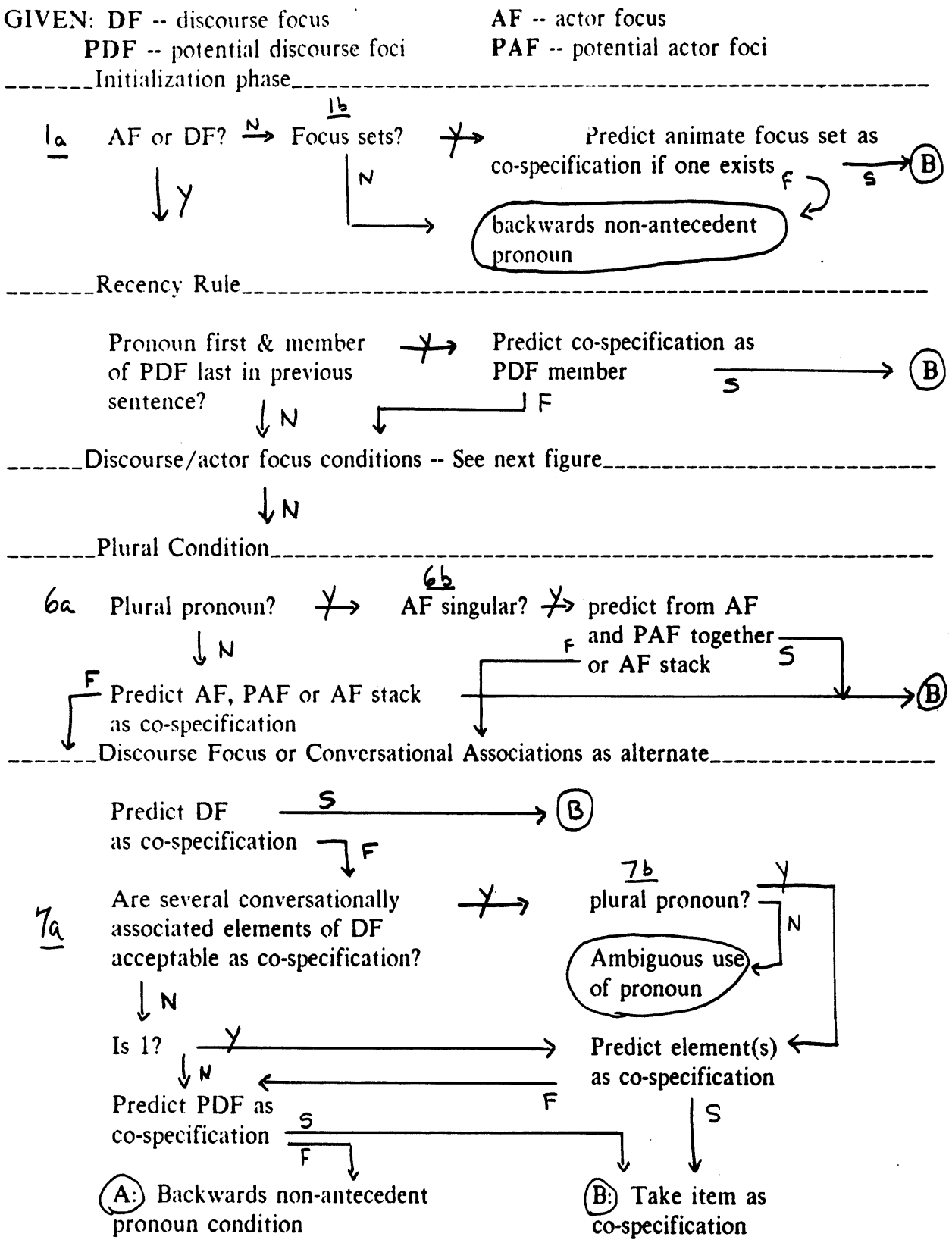


**figure I.3. Discourse/Actor Conditions for Agent Third Person Pronoun**

GIVEN: **DF** -- discourse focus **AF** -- actor focus
**PDF** -- potential discourse foci **PAF** -- potential actor foci

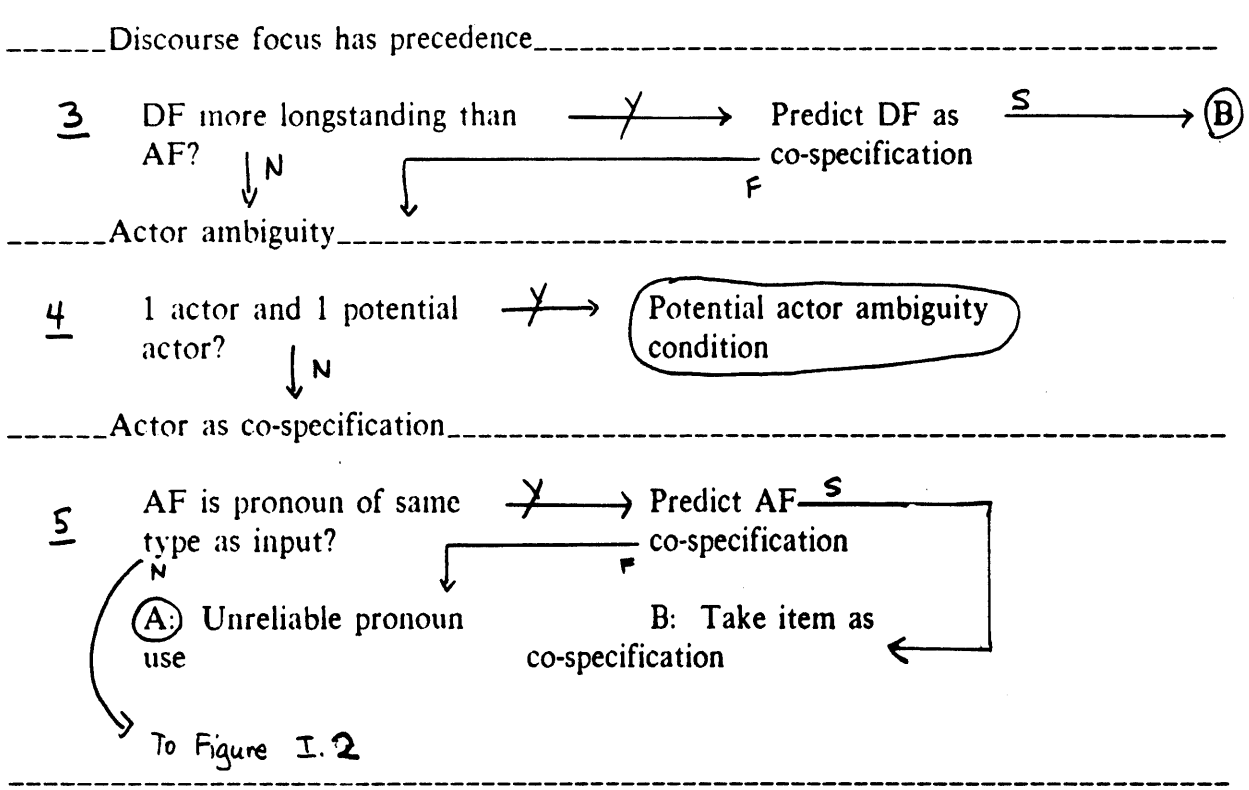


**figure I.4. Rule for Third Person Pronoun in Non-Agent position**

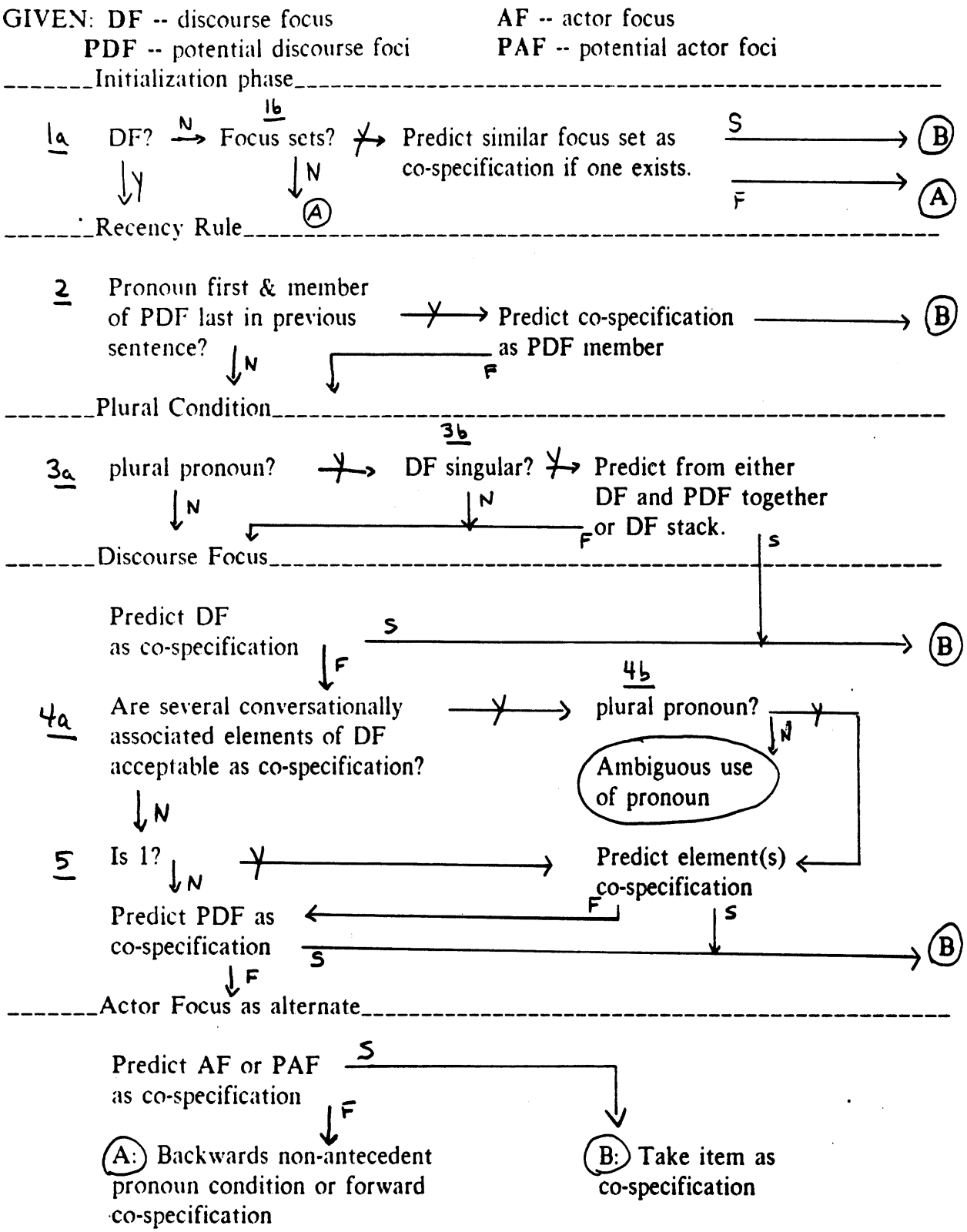


**figure I.5. Rule for Third Person Personal Possessive Pronouns**

GIVEN: DF -- discourse focus AF -- actor focus
PDF -- potential discourse foci PAF -- potential actor foci

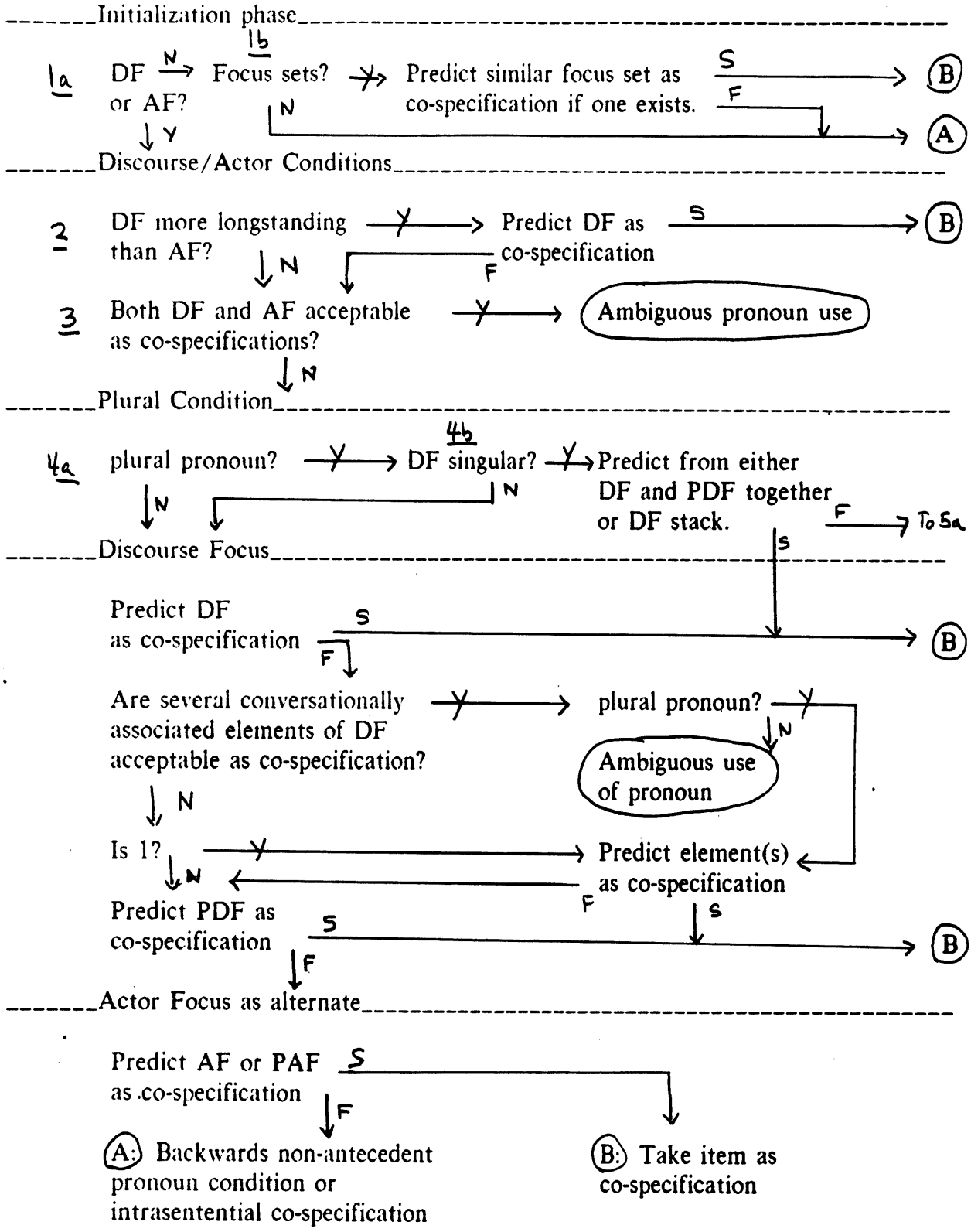


**figure I.6. Rules for Interpretation of *the one...the other***

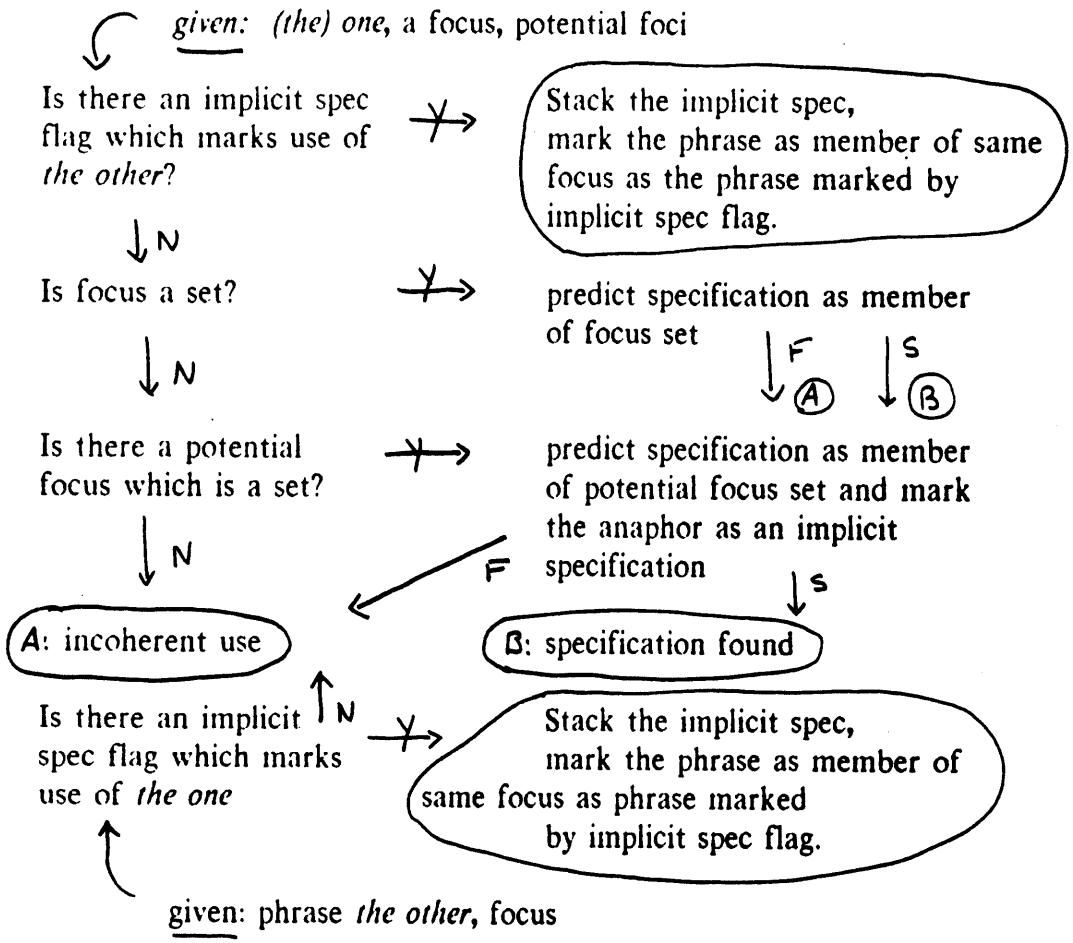


**figure I.7. Interpretation of *This* Noun Phrases**

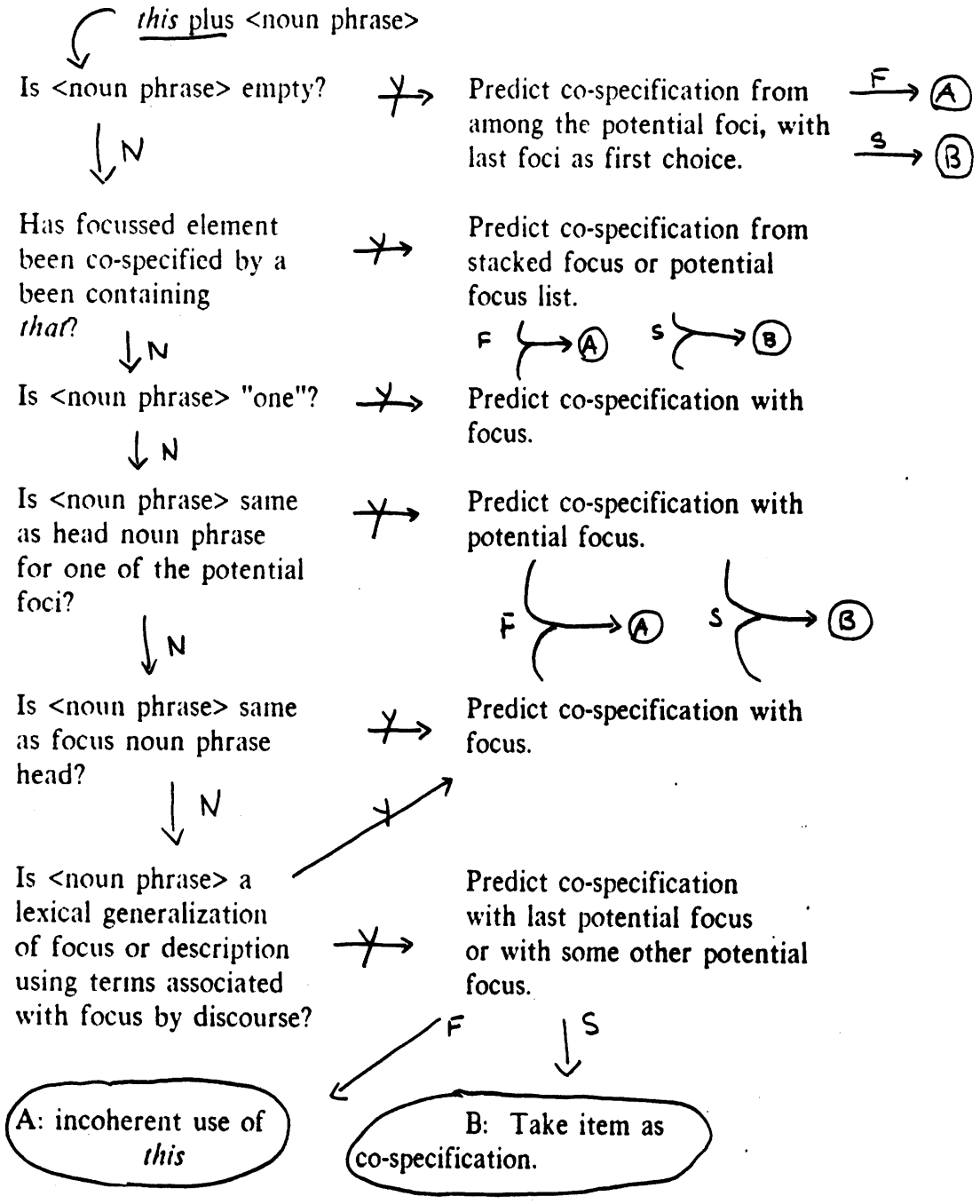
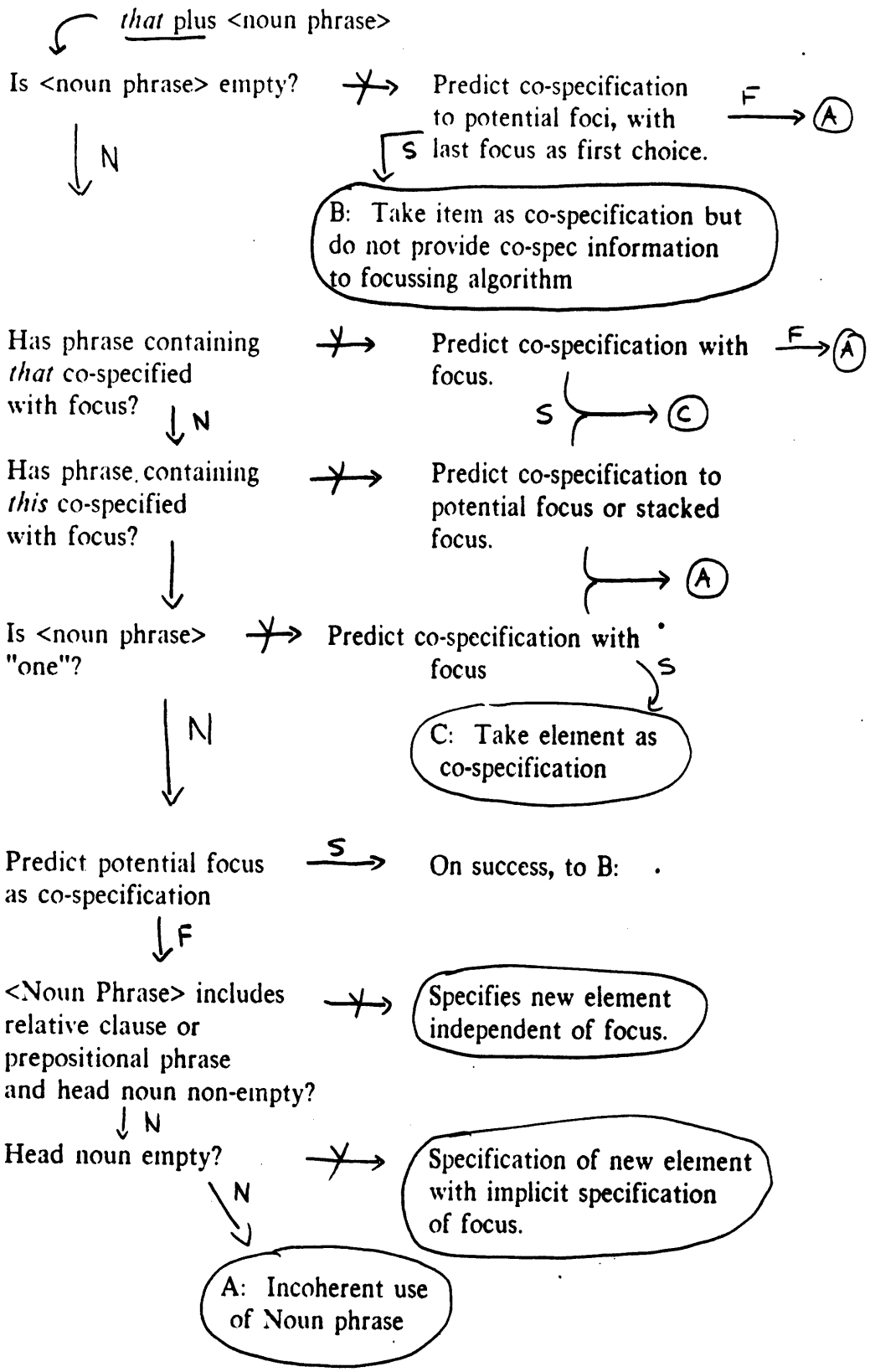


**figure I.8. Interpretation of *that* Noun Phrases**

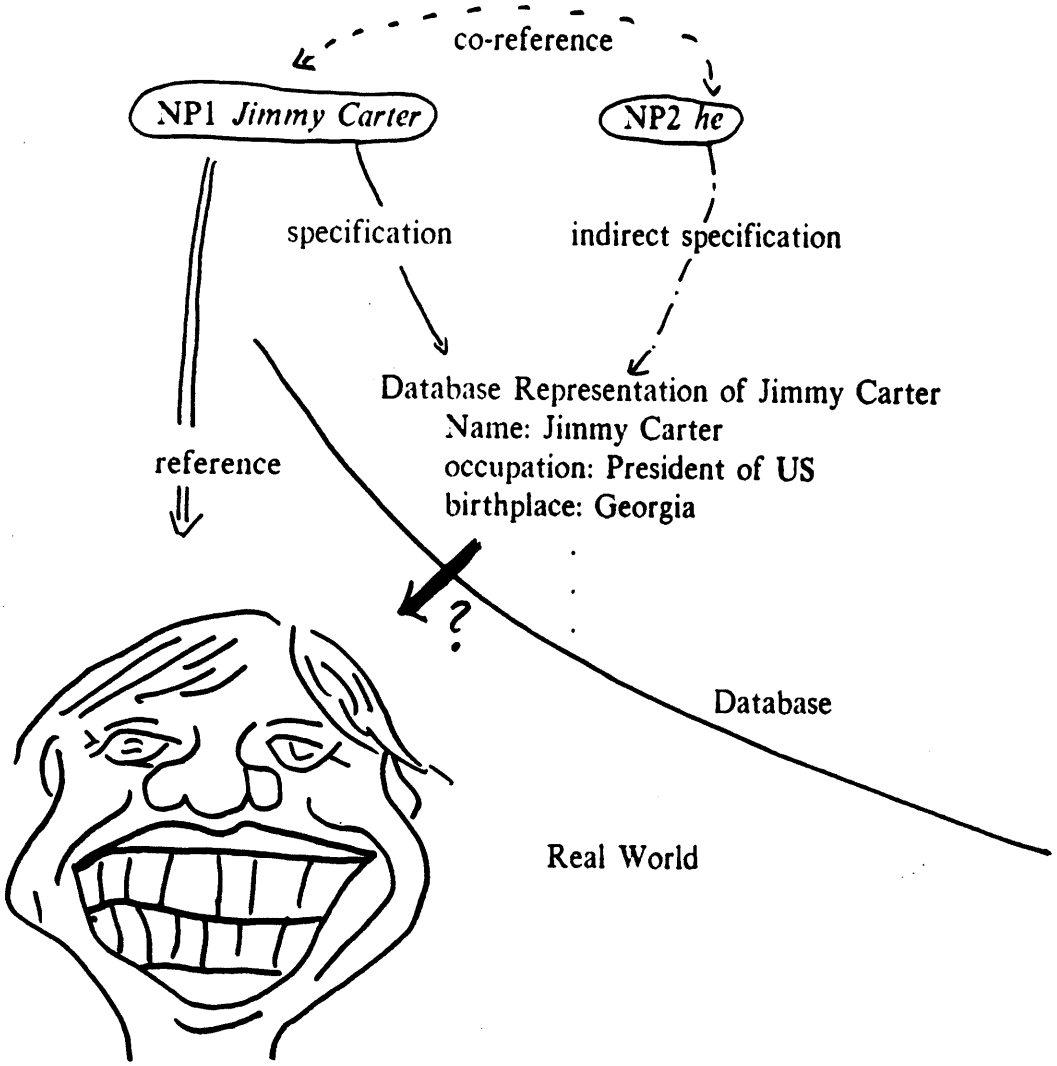


## Appendix II - Co-specification and Reference

When we informally talk of noun phrases, we call them "referring expressions." In fact, noun phrases can be distinguished by two different uses. First, a noun phrase may be used refer to some entity in the real, or some imaginery, world. The noun phrase *Jimmy Carter* can be used to refer to Jimmy Carter, who is the current president of the United States. In figure 1 below, this relationship is expressed by the reference arrow. Names are a common way of referring, and definite anaphora are another. If a speaker talks about Jimmy Carter, the speaker may continue to speak using the pronouns *he* or *him*, or a definite noun phrase such as *the President*. In this case, the

---

**figure II.1. Reference Links Between Noun Phrases**



definite anaphor is said to co-refer with the noun phrase *Jimmy Carter*, that is, both the anaphor and the noun phrase refer, and are being used to refer to the same entity in the world. A sample use of co-reference, one between a name and a pronoun, is depicted by the dashed link from NP2 to NP1 in figure 1.

When we speak of the comprehension of a pronoun, or some other definite anaphora, we might want to define comprehension as finding the co-reference of the pronoun, in the context of some other noun phrases which are used to refer. This definition is a troublesome one because comprehension by a hearer is possible even when the hearer does not know who is referred to by the first referring noun phrases.

The trouble with co-reference as a definition of antecedence can be eliminated by a computational framework; rather than consider the relation between a string of words and entities in the real world, I will define some relations between words, their interpretation in the sentence in which they occur, and elements of a database. Informally, I call the noun phrase and its interpretation based on syntax and semantics the bundle of a noun phrase. Specification is the relation between a noun phrase, its syntactic and semantic interpretation in a sentence, and some database object. In other words, the bundle of a noun phrase specifies a database element which represents the real world object. The first use of *Jimmy Carter* specifies the database element shown in figure 1 as the database representation of Jimmy Carter. The database object is a representation of the real world person Jimmy Carter. By analogy to reference one can ask how the database object which is the specification refers to the real world. This question is, to my knowledge, an open question in the theory of cognitive science and will remain so in this report; the relationship in question is depicted by the filled arrow, which is labelled with a question mark, between the database representation and the person Jimmy Carter.

What can be said about the definite anaphor *he*? The bundle of the noun phrase *Jimmy Carter*, and its specification can be used to determine that *he* specifies the same database element that the *Jimmy Carter* bundle does. Put another way, *he* and its bundle co-specifies with the *Jimmy Carter* bundle to the specification of Jimmy Carter in the database. Because the specification of *he* depends on *Jimmy Carter* and its specification, the specification of *he* is obtained indirectly. This indirect specification of *he* is shown in figure 1 by the dot-dash link.

What is the advantage of this framework? When no representation exists in the database for Jimmy Carter, we can still speak of the relation between the *Jimmy Carter* bundle and the *he* bundle. The *he* bundle still co-specifies with the *Jimmy Carter* bundle, and the bundles alone can be used to establish the co-specification relation. If a specification is ever found for *Jimmy Carter*, the one for *he* is found by virtue of the co-specification link. Furthermore, this definition sets aside the classical problem in logic of the denotation of referring expressions in English (see Kripke [1972], Russell [1905] and Frege [1896]).[1]

There is a second advantage in the computational definitions. Noun phrases which we informally call referring expressions are not always used to refer! They are sometimes used a second way, namely, to construct something which can be talked about. An example will make plain this fuzzy talk.

D1-1 Mary has a dog.
2 He's quite friendly
3 because he wags his tail a lot and wants to play.

In D1, the noun phrase *a dog* is not used to refer in the same way as a name, that is, *a dog* does not denote an entity in the world; instead the speaker in saying D1-1 constructs an entity in the discourse about which more is said in D1-2 and 3. Informally people talk about the relation between *he* and *a dog* as being one of antecedence, but since *a dog* does not denote, the co-reference link fails to describe the informal notion of antecedence. However, an antecedent for *he* can be constructed from the previous sentence. The antecedent is a description like "the dog which Mary owns." Evans [1977] speaks of such pronouns (he calls them E-type pronouns) as having the reference fixed by a description which is formed from the antecedent sentence by a conjunction of the main clause into which the antecedent is inserted, by the common noun in the antecedent expression, and by any relative clause restricting the antecedent.[2]

---

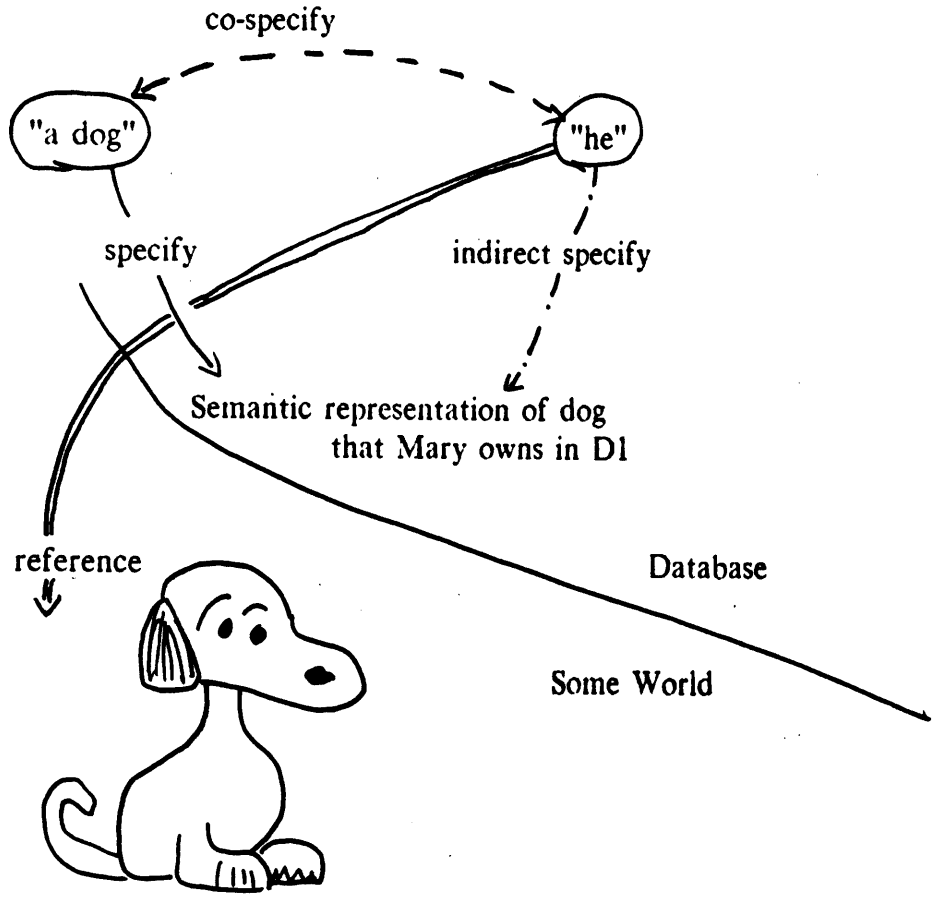
1. Fodor [1978] and Putnam [1978] indicate that this is not a problem for attack by psychologists and artificial intelligence reseachers. Smith [1978] explores a computational view of the concept of reference using the notion of interpretation to define reference.
2. Evans, *op. cit.*, page 535.

The description of Mary's dog is what people intuitively understand as the antecedent of *he* in D1; the co-specification relation captures exactly this intuition. *He* specifies the same element as the bundle of *a dog*, i.e., the dog that Mary owns. In this case, the specification is just a semantic representation of the dog which can be constructed from the discourse context. Thus the bundle of *he* co-specifies with the bundle of *a dog*, to the dog that Mary owns, as shown in figure 2.

The co-specification relation has the advantage of allowing antecedence to be defined in a consistent fashion whether or not the noun phrase which is part of the co-specification relation refers. In fact there is no need to talk of reference or co-reference at all. Instead, the antecedence relation between a noun phrase and a definite anaphor can be defined in terms of the specification relation. We can leave the

---

**figure II.2. Co-specification for indefinite noun phrases and definite anaphors**



question of whether the pronoun refers[1] for philosophical inquiry.

To summarize, for noun phrases that are discourse constructs, the specification of the noun phrase bundle is a semantic representation in the database derived from the discourse. A definite anaphor and its bundle will co-specify with the noun phrase bundle to the semantic representation. For noun phrases which denote, co-specification will function much like co-reference. For the case of *Jimmy Carter* in figure 1, NP2 *he* co-specifies with NP1's bundle to the database representation of Jimmy Carter. As long as there is a person to whom the noun phrase *Jimmy Carter* refers, and that person is represented in the database, the co-specification relation between an anaphor and an antecedent chooses the represented person. Said another way, the relation of antecedence between two discourse phrase bundles can be defined as the co-specification with relation.

---

1. A current debate in philosophy centers on whether pronouns can be thought of as co-referring or as bound variables. Geach [1962] holds the latter position and claims that the pronouns in (a) and (b) function in the same way.

(a) If some man eats Rice Krispies, he will be happy.

(b) If John eats Rice Krispies, he will be happy.

In the first case the pronoun is a bound variable and does not refer (because its antecedent does not), so similarly in the second case it does not either. Evans [1977] argues, however, that (b) type pronouns function in co-reference because the truth-theoretic semantics of these sentences can be interpreted differently. Evans points out in his paper the argument which is most critical for this report: if pronouns are described as co-referring rather than as bound variables, only a small change in the referential semantics of pronouns is needed for handling pronouns which refer across sentence boundaries, while on Geach's account, the semantics are quite complex if any account can be made at all. Furthermore, co-referring treatment is more in line with the E-type cases mentioned earlier.