THE UNIVERSITY *of* EDINBURGH

FOCUS AND COHERENCE IN DISCOURSE

by

L.G.Binet

Department of Artificial Intelligence
&
School of Epistemics

A Thesis submitted for the degree of Master of Philosophy
at the University of Edinburgh, 1984.

CONTENTS

INTRODUCTION

Since Frege, many people have regarded meaning as a relation between an expression in a language and entities in the world, a relation that is essentially fixed in character, so that in a given context the expression always has the same extension. But it is not clear that this approach is really appropriate for cognitive science, and it has been suggested that what we really need is a "procedural semantics", relating utterances to mental representations rather than to the world (see, for example, Isard and Davies 1972, Johnson-Laird 1977, and Woods 1981). At the same time more attention is being paid to pragmatics and the effects of context on meaning. Both these trends are represented, for example, in Johnson-Laird's theory of mental models (Johnson-Laird 1983), in which representations of discourse content, incorporating pragmatic knowledge of the world, guide the process of semantic interpretation in such a way that sense and reference interact dialectically, as it were. In keeping with this sort of approach, I would like to suggest two methodological principles. First, that linguistic structure should always be explained in terms of its function in the cognitive process (and hence that semantic structure is best explained by something like procedural semantics). Second, that a theory of semantics must include some account of pragmatics. In other words, I would like to move away from the notion that language can be formally characterised as an independent abstract structure, and look at how it is used instead. I think computational linguistics may have something special to contribute in this area, and I shall try to illus-

trate this with a discussion of some discourse phenomena.

This paper looks at how the theory of focus can be used to explain some aspects of discourse coherence, particularly those related to the interpretation of anaphora. Part One is a general discussion of such theories, concentrating on two theories of local focus. Part Two looks in detail at a computer model based on one of these. Part Three presents an extended example of discourse analysis based on the theory of focus, and looks at some ways in which it might be extended.

PART ONE THEORIES OF FOCUS

1) COHERENCE

Communication can be seen as a problem in interactive planning, in which each participant is trying to reconstruct the goals and plans of the other, while at the same time trying to reconcile them with his own. These plans will specify a series of subgoals and subplans that range from solving large-scale problems like trying to get a bank loan, down to purely linguistic matters like constructing a sentence, and any part of a discourse will seem coherent in so far as it is possible to construct an interpretation that includes a partial reconstruction of the speaker's plan. This is a matter of fitting utterances into their context, both linguistic and extra-linguistic, and so coherence is a matter of degree - more processing will nearly always produce an interpretation of some sort by finding some less obvious link with what is known about the context, or by making hypotheses that extend it in some way (as, for example, when one tries to "think of a context" for some unusual sentence). Coherence, then, is a measure of how easy it is to interpret what is said, and not an all or nothing affair.

2) RELEVANCE AND KNOWLEDGE ACTIVATION.

One of the most important things AI has taught us about language is that when people communicate they draw on a large pool of shared knowledge about the world and the context (see, for example, Winograd 1972 or Charniak 1972). But if the discourse is to be

coherent, this information must be used in an orderly fashion, according to something like Grice's maxim of relevance (Grice 1975, 1978). The hearer must decide which information to use in interpreting each utterance, and the speaker must help him somehow. Now utterances in a discourse act as instructions to build and manipulate representations of meaning that correspond to knowledge structures in the speaker's mind, and these contain a lot of information that could be used to interpret other parts of the discourse. So if this information remains "active" for some time after an utterance is processed - if the associated meaning representations are stored in some sort of working buffer, for example - then an efficient communication strategy would be for speakers to concentrate on providing unactivated knowledge, leaving hearers to provide the rest themselves.

In less abstract terms, that might mean reserving things like full NPs for newly relevant items, and letting activated knowledge structures take care of anaphora. In this way a balance is struck between informativity and coherence. In accordance with Grice's Maxims (1975), people expect what is said to be relevant to the context, the preceding discourse, the speaker's intentions, etc., and yet to also contain the minimum of redundant information. How the balance is struck between these two goals depends on what effect the speaker is trying to achieve - sometimes it is useful to make the discourse very informative but difficult to process (take James Joyce's "Ulysses", for example), whereas at other times clarity is all important (de Beaugrande and Dressler 1981 describe this as a trade-off between "effectiveness" and "efficiency").

3) FOCUS

In particular, what is said should be relevant to some topic, where "topic" is a loose pretheoretical term indicating some conceptual structure (what the speaker wants to talk about) which in some sense governs or constitutes a step in the discourse plan intermediate between extra-linguistic goals (why he wants to say it) and their linguistic solutions (how he says it). Since the decision to talk about a topic is a sub-goal that may not be directly attainable, topics are often hierarchically arranged so that a given piece of discourse is "about" several things at once, on a number of different levels. The strategies whereby speakers use topics to structure their contributions, and hearers use them to construct their interpretations, are typically part linguistic (the rhetorical conventions of the genre/medium/register, etc.), and part pragmatic (i.e. to reflect the structure of events and entities in the outside world).

The notion of focus is one attempt to provide a processing account of the way in which relevance to topics controls inference in discourse, and so makes it coherent. Although some authors have proposed models in which pragmatic inferences are only made when the normal rules of discourse are violated (e.g. "conversational implicatures"), when other more strictly logical forms of inference fail to produce an interpretation, or when specific linguistic mechanisms signal that they are necessary (e.g. "bridging inferences" - Clark and Haviland 1977), the evidence seems to suggest that they are going on all the time (see, for example, Sanford and Garrod 1981). Such inferences must be guided in some way if they are not to ex-

plode in complexity, and in particular the search for referents required by anaphora must be constrained in some way to make it efficient. To account for this, it has been proposed that only a few representations are kept in active memory storage at any one time, and that only the corresponding entities are readily available as referents. Such entities are "in focus", and seem to correspond to those which have most recently been mentioned and those which are are most closely related to the current topic (c.f. Brown and Yule 1983).

For example, definite NPs are often used in a way that suggests that the entities to which they refer are in some sense "given" or "inferrable" from the text, by virtue of some pragmatic relation to the current topic, even though they have not been explicitly mentioned (Clark and Haviland 1977). This suggests that the process of interpreting the discourse calls up representations for such "implied" entities before they are needed, and that these sit in active storage along with representations for all those things and events explicitly mentioned. Sanford and Garrod (1981) call this "implicit focus", and this sort of automatic inference has led to a number of what they call "scenario-based" models, in which the representations that are focussed include a lot of extra information not drawn from what is said, but gathered from experience of the real world. These representations are generally similar to frames (Minsky 1975) or scripts (Schank and Abelson 1977), and models have been proposed in which they are active data structures that control discourse inference by a process of spreading activation (e.g. Bobrow et al. 1977). There is obviously a link between such focussed

frames and the loose notion of a topic.

In this sort of model, then, the hearer primarily relies on the active part of his knowledge base when trying to interpret what is said, and the speaker "semantically underspecifies" activated items accordingly. For example, he may use a pronoun to refer to some entity that he knows is at the forefront of the hearer's mind, while using a more complex noun phrase for something that is new or has slipped from consciousness, and this speeds up the rate of communication by cutting out redundant information. These theories usually describe activated knowledge representations as being "in focus"- often with the tacit assumption that "focus" is some separate area of memory, although this is not absolutely necessary- and the process of activating and de-activating them is referred to as "focussing". There is also a tendency to call some of the more important items in focus "foci", although this is perhaps a little confusing. I shall use this terminology, albeit rather loosely.

Grosz (1981) points out that the participants of a discourse may have different foci, reflecting the different state of knowledge activation in each, or to put it another way, the different topics that they perceive the discourse to be about. She maintains that coherence relies on the hearer faithfully tracking the speaker's focus, but while this is partly true, it is not the whole story. It is just as important for the speaker to monitor the hearer's focus and tailor what he says accordingly, which means using appropriate linguistic forms to guide the hearer's focus or to get him to use knowledge that is already activated. So each participant has a pool of activated knowledge and some representation of the state of

knowledge activation of the other participant. Furthermore, there may be speaker's models of hearer's models of speaker's activation, and so on ad nauseam, but in normal discourse this would not be necessary. Building embedded models like this is difficult, and people only seem to do it as a last resort. Instead, they seem to mutually agree (in the sense used by Clark and Marshall 1981) that certain structures are to be activated, and these are the ones that are "in focus". Focussed items are thus a subset of activated items, and focus is the product of a communicative contract, not just a psychological process. It is perfectly possible for one or both of the participants in a discourse to have something on his mind without it being an agreed topic of conversation - indeed getting things from activation into focus is surely what discourse planning is all about. Later, I will indicate another notion of what focus is - the idea that it is an abstract feature of a text - and suggest that that too is misleading. If focus is to be properly understood, we must adopt a functional approach, not a purely formal or a purely psychological one.

It is generally supposed that the representations held in focus are arranged to form some coherent structure. This is often thought to consist of a hierarchical network of data structures similar to Minsky's "frames" (Minsky 1975), whose individual structures represent general knowledge about the world, and which are linked together in ways that represent information culled from the discourse. Different sorts of representations have been proposed for different sorts of knowledge (e.g scripts, schemata, scenarios, etc.), but I shall generally refer to them as frames, partly because

Minsky's theory is the most general (and hence the vaguest). Because frames form a hierarachy with each frame having other frames as sub-components, some will be very general and will be associated with the broad structure of the discourse, while others lower down the network will be concerned with its local details - this reflects the hierarchy of topics and subtopics. For the purpose of focussing theory it is usual to divide this continuum into two levels, local and global, and to describe discourse accordingly. This may only be a convenient fiction, but it may also have some psychological validity (perhaps reflecting differences between short term and long term memory, for example). At any rate, it seems useful to distinguish between local focus and global focus, and to look for corresponding local and global structures in discourse.

GLOBAL	LOCAL	AUTHOR
Implicit Focus	Explicit Focus	Sanford and Garrod (1981)
Global Focus (Focus Spaces)	Centering	Grosz et al (1981, 1983)
	Actor Focus Discourse Focus	Sidner (1981)
Context Spaces		Reichmann (1978)
DSEM	DM	Shadbolt and Reichgelt (Forthcoming)
Scripts		Schank and Abelson (1977)

Figure 1 Theories of Focus

A number of theories of focus are listed in figure 1. Global

focus contains high level representations with a great deal of substructure, and these control long stretches of the discourse. The dynamics of global focussing reflect the rhetorical strategies people have for shifting scenes and topics, the effects such shifts have on their expectations, and the overall coherence of the text. A number of processing-oriented models have dealt with this sort of thing, but local focus is perhaps less familiar, because it has usually been described in terms of the corresponding linguistic structures rather than the processes that use them. A theory of local focus should thus try to explain phenomena like "staging" (Grimes 1975), "foregrounding" (Chafe 1972), "thematic structure" (Halliday 1967), and various notions of sentence topic. In particular, local focus seems to be strongly related to the use of pronouns.

For example, Sanford and Garrod (1981) propose a model in which active storage is divided into "implicit focus" and "explicit focus", which correspond roughly to global and local focus. Implicit focus contains frames (Sanford and Garrod call them "scenarios") which provide default information about the scene, to which more details are added as they are mentioned. Representations for the entities and events explicitly mentioned are based on these and stored in explicit focus. Under normal circumstances, an indefinite NP signals the entry of a completely new item into explicit focus, a definite NP demands that something be moved from implicit focus to explicit focus, and pronominal reference acts as an instruction to append extra information to a representation already in explicit focus.

4) GLOBAL FOCUS

Other models of global focus include those of Grosz (1981) and Reichmann (1978). In Grosz's model, the speaker and hearer each represent the discourse in the form of a semantic network, which is divided into regions called focus spaces, and it is these which constrain the search space for anaphoric reference. Cues for the opening and closing of focus spaces are both linguistic and pragmatic, and spaces are often hierarchically arranged. In particular, Grosz discusses task-oriented dialogues, where shifts of focus and topic structure frequently reflect the structure of the task itself.

Reichmann (1978), on the other hand, is more concerned with the more purely linguistic mechanisms used for guiding global focus in ordinary conversation. She talks in terms of "context spaces", which are like focus spaces, but which also contain default information about contexts in the way that frames or scripts do, allowing the hearer's expectations about them to guide him in shifting from one to the next. For example, since we know football matches usually only last about ninety minutes, if we hear a phrase like "three days later" we will not expect to hear about the same football match that was being described in the previous sentence. Context spaces are classified as "issue spaces" or "event spaces", roughly depending on whether what is said is descriptive or narrative (Reichmann's data is all taken from ordinary conversation, i.e. primarily interactional discourse. It is not clear that the same taxonomy would be suitable for transactional discourse such as Grosz's task-oriented dialogues). Reichmann then attempts to formulate general rules about the strategies people have for shifting and tracking focus, the re-

lationships between one space and another, and the linguistic phenomena used to mark boundaries between them. Among these are "clue words", like "but" or "anyway", explicit phrases such as "for example", various deictic phenomena such as tense shift, and various sorts of structural parallelism.

5) LOCAL FOCUS.

In addition to her theory of global focus, Reichmann suggests that the entities referred to in a given context space may be more or less foregrounded (c.f. Chafe 1975), and that this is reflected in pronominalization. Topic entities will often be highly foregrounded, and various effects of recency and emphasis may serve to highlight others as the discourse proceeds. Brown and Yule (1983), for example, point out that narratives tend to use pronouns to refer more frequently to topics, whereas descriptive or transactional discourse tends to use them more for the "current evoked" item (i.e. the last "new" item to have been introduced - the terminology is an extension of that used by Prince 1981, whose "textually evoked" items seem to correspond to foci. In Part Three I will mention the focussing effect of various deictic phenomena that seem to correspond to Prince's "situationally evoked" category.). These phenomena might be related to Reichmann's distinction between issue spaces and event spaces, and seems to reflect the effects of global focus and local focus respectively. At any rate, it seems that some sort of theory of local focus is necessary to explain the latter phenomena, and models based on local focus do seem to offer special insight into the use of pronominal anaphora.

One computationally explicit model of local focus is outlined in Sidner (1981). She further subdivides local focus into "discourse focus" and "actor focus" in order to capture some of the effects of the thematic structure of sentences on foregrounding, and lets each of these regions contain a single most prominent entity (the current focus), a set of "potential foci" suggested by the utterance being processed and ordered by case, and two stacks of foci that have been displaced. Focussing thus consists of moving one of the potential foci to the "current focus" position, and putting the old focus onto the stack.

In this model, pronoun interpretation is done by "hypothesis testing"- possible referents are suggested in a fixed order by the focus system and each such hypothesis is tested against a series of constraints until one is found that satisfies them all. The rules for pronoun interpretation are sketched in figure 2.

If the pronoun is an agent try:

1 The actor focus
2 The discourse focus
3 The potential foci (in order of case)
4 The previous foci
(starting with the most recent)

Otherwise try:

1 The discourse focus
2 The potential foci (in order of case)
3 The actor focus
4 The previous foci
(starting with the most recent)

Figure 2 Pronoun Interpretation in Sidner's Model

The constraints Sidner uses to test these hypotheses cover number

and gender agreement and various other syntactic restrictions, but she points out that a proper theory would include semantic and pragmatic constraints too. Then once the current utterance is fully interpreted the foci are updated. Pronominalization plays an important role here too, because the two foci are generally forced to track pronominalized entities - each staying the same if pronominalised, or otherwise shifting to some other pronominalised entity.

Grosz, Joshi, and Weinstein (1983) outline a model which is in some ways similar to Sidner's, although less complete. Although much of the paper is devoted to making the point that the relation between noun phrases and the entities that they realize is more complex than has often been supposed, and that it is these entities which move in and out of focus during discourse, all this is more in the nature of a metatheoretical manifesto than a definite theory. However, the proposed theory of centering does seem to say something definite about the strategies people have for using pronouns.

They point out that one possible weakness of Sidner's model is that in order to explain examples like

> I haven't seen Jeff for several days.
> Carl thinks he is studying for his exams.
> But I think he went to the Cape with Linda.

in which "he" means Jeff throughout, Sidner finds it necessary to introduce an extra case role based on Halliday's notion of "theme" (this is a case role and is part of his transitivity system, as opposed to the notion of "theme" that is part of the thematic system - the terminology is confusing. See Halliday 1967). So in a sentence like

Carl thinks Pete is studying for his exams.

Carl is the agent and Pete is the theme. Grosz et al see this as a point against Sidner's model, and suggest a more elegant way of doing things.

Essentially, they propose to abandon the idea of actor focus, so that there is just one kind of local focussing, which they call "centering". In their model the current focus is called the "centre", and potential foci are called "f-centres" (strictly speaking, these are abbreviations for "backward looking centre" and "forward looking centres"). As in Sidner's model there is also a stack of previous foci, but it is suggested that those shifts of local focus that rely on it are usually directed by the global focussing system, top-down. Thus the centering system is much the same as Sidner's discourse focus mechanism, although it does not appear to have yet been implemented in computational form, and is correspondingly much vaguer. It is suggested that a number of centering rules will eventually be found which will - according to Joshi and Weinstein (1981) - account for such factors as the effects of "syntactic structure, stress, and discourse context" in order to determine which of the f-centres becomes the centre. However, in the 1983 paper only one such rule is proposed, which is written from the speaker's point of view:

> If the centre of the current utterance is
> the same as the centre of the previous utterance,
> a pronoun should be used.

This is obviously closely related to Sidner's idea that focus should track pronominalized entities, but it is claimed that it explains

examples like (1) more elegantly, because Jeff is the centre throughout and so is pronominalized throughout, and there is apparently no need to invoke case at all. Joshi and Weinstein (1981) have tried to use this model to get a mathematical characterization of coherence, by looking at the complexity of the inferences necessary to incorporate each new utterance into the discourse model. They so far report only limited success.

Sidner's model and the centering model are thus similar in outline, and both make two sorts of empirical claim. The first concerns the way in which items are brought into local focus, and in both cases focus follows those items that are pronominalized. It should be noted that, in addition to the above rule, Grosz et al state that entities other than the centre can be pronominalized "as long as one is used for the centre", implying that, from the hearer's point of view, the corresponding rule should be:

> If pronouns are used, the centre is unchanged and pronominalized. If not, the centre has changed.

This is basically the rule that Sidner uses, except that she has two sorts of local focus. The other kind of empirical claim is about the way that focus then affects the interpretation of pronouns. Sidner gives a complete list of pronoun interpretation rules, listing the order in which various possible interpretations should be tested against her constraints. Grosz et al do not, but the centre rule does suggest that pronouns will most often refer to the current centre. The model described in this paper adds a set of pronoun interpretation rules to produce a working model based on the theory of centering.

PART TWO A MODEL OF LOCAL FOCUS

1) INTRODUCTION.

It is difficult to compare the theory of centering directly with Sidner's model, not only because it has yet to be fully articulated, but also because it appears to be a model of production rather than comprehension. The centre rule is expressed as a restriction on how a speaker may use pronouns, and not a strategy for pronoun interpretation. Grosz et al seem to blur this distinction when they compare the two models, and I suspect this is because they are interested in characterising language as an abstract structure, rather than in modelling the communicative process. For them the centre rule is just a well-formedness condition for utterances in a discourse. Thus, when talking about sentences with several pronouns like:

He thinks he studies too much.

they remark that the two pronouns are allowed to have different referents, because other entities can be pronominalized,just as long as the centre is. Now this is fine, in that it ALLOWS both the following examples:

Jeff is worried.
He thinks he isn't studying enough.

Carl is worried about Jeff.
He thinks he studies too much.

but it doesn't tell you why the pronouns are interpreted in the way they are. Why do the pronouns in the second example refer to dif-

ferent people? Sidner's theory of actor focus is postulated to explain just this sort of thing, whereas the theory of centering doesn't appear to have anything to say about it. I think this is a good example of the way in which the formal approach can be misleading - we have to look at the dynamics of communication, and not just at linguistic structure.

Nevertheless, it may still be possible to reconstruct the theory of centering, or something like it, in a form that can be compared directly can be made with Sidner's model. If we take it seriously as a theory of production, then the centre rule suggests two corresponding strategies for interpretation:

> If pronouns are used, then at least one of them probably refers to the centre, which is unchanged.
>
> If not, the centre has probably shifted.

These are not hard and fast rules, because Grosz et al allow for exceptions - the centre may not be realized in the utterance at all and nonpronominal NP's may be used attributively without shifting the centre.

POCUS (see appendix 1) is a computer program for resolving pronominal anaphora, based on these rules about centering, while adding further pronoun interpretation procedures to produce a more complete model, like Sidner's. It does not tackle the problem of how noun phrases realize entities in discourse, indeed it has no proper semantic component, and so is limited to finding antecedent noun phrases for the pronouns. Such an approach is, of course, unsatisfactory (as a number of authors have stressed, including both Grosz and Sidner - see Brown and Yule 1983 for a review) and may seem out

of keeping with the spirit of Grosz et al's paper, but it does at least allow us to explore some aspects of local focus using a working model and actual examples.

Given a short piece of text, consisting of complete sentences, POCUS will parse the sentences and label the pronouns on each parse tree with their antecedent noun phrases, provided it can find them. In the current version of the program these trees are merely used to rewrite the sentences with the pronouns replaced by their antecedents, but a more realistic model would use them as the basis for constructing a representation for the meaning of the text.

The main difference between POCUS and centering theory is that there is no longer a unique centre all the time. This is a necessary consequence of trying to model the hearer rather than the speaker, because although the speaker may have some unique topic entity that he wishes to talk about, the hearer cannot always be expected to latch onto it immediately. For example, a text that began "Mark looked at Alison" could develop in a number of ways:

Mark looked at Alison. She was nervous.
[centre = Alison]

Mark looked at Alison. He was nervous.
[centre = Mark]

Mark looked at Alison. They hadn't met before.
[centre = Mark and Alison]

Mark looked at Alison. She smiled at him.
[centre = ?]

It is clear that the hearer cannot uniquely identify the centre from the first sentence alone, but must at least "focus in" gradually on one. This may be quite an important process in discourse, and may be

reflected in what Grimes calls "staging", the way in which speakers order what they say to "lead" the hearer's focus. But if it is also accepted that what governs the production of discourse is not how the speaker's attention is focussed, but how he believes the hearer's is, then the discourse ceases to have a unique centre in any sense. Instead, a number of items can be in focus at any given instant, and what the speaker does is persuade the listener to focus in on what he thinks is important. Chafe took this sort of stance when he said that foregrounded items are those "assumed to be in the hearer's consciousness" rather than those in the speaker's (Chafe 1972), and I shall follow his example.

Thus POCUS has a list of centered entities, which represent hypotheses about what the speaker is going to focus in on. Since centered items are usually referred to by pronouns once introduced, a useful focussing strategy is to continually prune the list by discarding anything that isn't pronominalized. This is what POCUS does, and it does seem to allow it to home in on a topic entity in those cases where one can be unambiguously identified.

POCUS implements the pronoun interpretation strategies outlined above in a hypothesis-testing procedure like Sidner's. When a pronoun is encountered POCUS tries:

1 The centered items (in order of occurrence)
2 The f-centres of the previous sentence (in order)
3 The whole set of centered items
4 The f-centres of the current sentence (in order)
5 The previous centres (starting with the most recent).

These hypotheses are tested against Chomsky's Matching and Binding conditions (Chomsky 1981). POCUS has been tested on a number of con-

structed "texts", most of which are included in the appendix. It would, of course, be better to use real data, but the limitations imposed by the grammar and the lexicon have prevented this so far

The system includes three new features not included in either Grosz's or Sidner's models. Firstly, hypothesis 3 enables it to deal with some cases where plural anaphora refer to sets of items introduced separately into the discourse, as in:

Mark met Alison in the street.
They decided to go for a walk.

Secondly, hypothesis 4 enables the system to cope with some examples of sentential anaphora, including some examples of cataphora. Thirdly, there is a mechanism that simulates the effects of global shifts of focus on centering by clearing the centre and f-centre lists in response to a "topic-shift" command. None of these features is very sophisticated, however, and they are only included to show what sorts of things might be going on.

2) POCUS

a) PARSING

POCUS uses a top-down active chart parser written by H.Thompson using MCHART, a flexible framework for writing such parsers which provides scheduling mechanisms and data structures without fixing the parsing strategy or grammatical formalism. A simple context-free grammar and a small lexicon have been added to give just enough grammatical coverage to deal with a reasonable range of examples. No claims to elegance or realism are made for the syntactic component - in particular, the grammar overgenerates because it does not make use of case frames or any other features (some features, such as number and gender, are carried through to be used by the pronoun interpreter, but the version of the parser used does not make use of them). However, this is not a problem, because the process of pronoun interpretation is fairly independent of the parsing process.

Although the parser finds all possible parses, it is assumed that most syntactic ambiguities would have been resolved by the time the pronouns are to be interpreted, and so only one parse is used. None of the examples used contain such ambiguities.

When a sentence has been parsed, the parser calls an interface package (see appendix 1) which finds the NPs of the sentence and hands them back to the main program. It searches the parse tree for NPs, and separates out those which are pronominalized. These are returned as they appear on the tree, but the rest are processed

further. Number and gender are found by a routine that locates the head constituent, after which the syntactic structure of the phrase is discarded and the lexical string is merely used as a label for the structure (A realistic model with a semantic component would find or build a representation for the entity that the noun phrase realizes and use that).

b) PRONOUN INTERPRETATION

Having parsed a sentence, and found and sorted all the NPs, POCUS now tries to find antecedents for the pronouns. As in Sidner's model, the focus mechanism suggests a series of possible interpretations for each pronoun, which are then tested against a number of constraints until one is found which obeys them all. POCUS makes these hypotheses about the antecedent in a fixed order, and this ordering constitutes the main empirical content of the model, in the same way that Sidner's pronoun interpretation rules do in her's. Possible antecedents are chosen and tested in order from among the following:

1 The list of centred entities.

As the model stands, it is not always possible to uniquely identify the centre of a sentence, so a list of possible candidates is kept instead. The first choice for the antecedent is always one of these.

2 The f-centres of the previous sentence.

If none of the centred entities is suitable, then each of the f-centres of the last sentence is tested as a possible referent. At present these are tried in a more or less random order, but it seems likely that their positions in the sentence, and in particular their grammatical roles and functions, will prove to be important. For instance, in Sidner's model both potential discourse foci and potential actor foci are ordered by case. This is not possible in POCUS, because the parser does not assign case roles to the NPs of

the sentence.

3 A set of centered entities.

If the pronoun in question is plural, and neither of the first two options yields a referent, then it is possible that it refers to a set composed of entities which were introduced into the discourse separately. An example would be "they" in:

Mark met Alison in London.
They decided to go to the Tate.

4 The f-centres of the current sentence.

POCUS next looks for a referent among the NPs of the current sentence. These are tried in the order in which they occur in the sentence, including those that follow the pronoun, allowing for the possibility of forward pronominalization (or "cataphora").

5 A stack of previous centres.

If all else fails, POCUS searches back through a stack of previous centres until it finds a possible referent, as in Sidner's system. This would normally occur when there is a return to some previous topic, in which case the shift is often signalled by various linguistic mechanisms such as tense shift, the use of special lexical items or phrases, etc. (see Reichmann 1978 for a fuller account). To simulate the action of these mechanisms, it is possible to clear the centre and f-centre lists at any point "by hand", forcing POCUS to either interpret the next pronoun as a sentential anaphor or to look back through the stack, usually inducing a shift of focus in either case. Some examples of the way in which this is used

are included in appendix 3.

Each of these hypotheses must be tested against constraints. In POCUS these constraints are syntactic, but a more realistic model might include semantic and pragmatic constraints as well. The first constraint is the most obvious:

> Matching Condition: anaphor and antecedent must agree in number and gender.

A matching condition for person is not included, as deictic pronouns such as "I" and "you" are not interpreted by POCUS.

Next, the hypothesis must be tested against Chomsky's binding conditions (Chomsky 1981):

> Pronominal binding condition: A pronoun must be free in its governing category, if it has one. (i.e. A pronoun must not be coreferential with any NP in its governing category that c-commands it.)
>
> Anaphor binding condition: A reflexive pronoun must be bound in its governing category.

For each pronoun in the sentence, the governing category is found, and a list of the c-commanding arguments within it constructed (with pronominal NPs replaced by their antecedents, provided they have been found). Then whenever a possible antecedent is found, it is checked against this list to ensure that the binding conditions permit that interpretation.

When an antecedent for the current pronoun has been found, it is entered on the parse tree and the next pronoun is processed. Once an antecedent has been found for a pronoun, it is available for use in checking binding conditions, ensuring that, in an example

like "He hit him", the two pronouns have disjoint reference

c) CENTERING

Having interpreted the pronouns in the sentence, POCUS now identifies the f-centres and looks for candidates for the centre. The f-centres are simply all those entities realized in the sentence by NPs. In this simplified model they are represented by a list containing all the non-pronominal NPs and the antecedents of the pronouns, ignoring all the complexities of 'implicit' reference, "speaker's references", etc. Candidates for the centre are found in accordance with the centre rule - if pronouns are used then the centre is unchanged and pronominalized. By setting the new set of candidates for the centre to be the intersection of the old set and the set of pronominalized entities, POCUS is able to gradually "focus in" if a unique centre is not immediately identifiable. For example, in

Mark gave Alison the gun.
It was making him nervous.
It looked too conspicuous.

the centre could initially be either Mark, Alison, or the gun. By the second sentence we have focussed in on Mark and the gun, and by the end it is clear that the gun is the main topic. This seems reasonable, given that POCUS is is intended as a model of comprehension, whereas Grosz et al take the speaker's point of view - although the speaker may have a unique topic in mind from the beginning, the hearer will not always be able to determine what it is straight away. Indeed, this may be a useful idea for explaining some of the thematic structure of discourse.

If no pronouns are used in the sentence, then it is assumed

that the centre has changed, as Grosz et al suggest, and so all the f-centres of the sentence become candidates for the centre. In a system in which definite NPs were also interpreted, it would also be necessary to determine whether or not any of these were coreferential with the current centre, in order to account for those cases mentioned by Grosz et al such as

> I'm reading "The French Lieutenant's Woman".
> The book, which is Fowles' best ... ,

where the use of the definite NP is attributive rather than referential, i.e. to introduce new information about a referent already in focus (e.g. that "The French Lieutenant's Woman" is a book, and not a play, a poem, or a woman).

Once the list of centre candidates is found, it is pushed onto the stack used by the pronoun interpreter. POCUS then returns its interpretation of the sentence, and moves on to parse the next one.

3) PERFORMANCE

POCUS has been tested on a number of examples with a reasonable degree of success, and a representative sample of these is included in appendix 3. Of particular interest, though, are those examples which the program fails to interpret correctly, since these show up several major weaknesses and suggest how the model might be improved.

a) Syntactic Component

As they stand, the parsing mechanism and the pronoun interpreter are more or less completely independent, so that the details of the syntactic component are fairly unimportant. Of course, the sizes of the grammar and the lexicon determine the range of examples that the program can deal with, but it is simple to extend these. However, the inability of the parser to assign case roles to the NPs of the sentence is a major flaw. Not only would this help in parsing the sentence, but it might also play an important part in the centering process. New centering rules might well be based on case, and f-centres both within and outside the sentence might be ordered for pronominalization on the basis of their case roles, as Sidner's "potential foci" are. Furthermore, role structure could be used as the basis for constructing representations for meaning (as it is in GUS, for example - see Bobrow et al 1977), and, along with the availability of associated features (e.g. animacy), could allow some simple semantic constraints, in the form of selection restrictions, to be added to the system.

There is also some doubt about the way the syntactic component and the pronoun interpreter interact. When a noun phrase has other NPs embedded in it, the interface routine does not list them separately, does not hand them over to the pronoun interpreter, and hence they are not available for reference. This is intended to explain why examples such as

Pete met Graham's mother.
He knew him from Oxford. (him = Graham)

Pete met the girl Graham liked.
He admired his taste. (his = Graham's)

seem odd. In both cases it would be much better to refer to Graham by name in the second sentence, and in the first case one would expect a shift of global focus to occur with Graham as the new topic. However, there do seem to be some counterexamples:

Pete read Alison's book.
He liked her style. (her = Alison's)

Pete read Alison's book.
She had an interesting style. (she = Alison)

It is not yet clear whether these are isolated cases yielding special thematization problems, or whether possessives should be treated separately. More work should be done to determine how people refer to items introduced in this way (see Part Three for some further comments on possessives and embedded NPs).

b) Semantics and Pragmatics.

An obvious class of examples which POCUS can't handle are those where the interpretation it chooses would be ruled out on semantic or pragmatic grounds. Given the example

> Martin and Janet bought some plates.
> They were green.

POCUS interprets "they" as referring to Martin and Janet. A partial solution to to this sort of problem would be (as Sidner suggests) to introduce a set of semantic constraints, in the form of selection restrictions, to accompany the syntactic ones already employed. However, such rules are difficult to formulate and it is often easy to find contexts in which they don't work, so it seems likely that the only satisfactory solution would be one where the sentence is processed up to the level of meaning and handed to a proper reasoning component which, armed with a lot of information about the world, would try to incorporate the sentence into some representation of the discourse as a whole. This is obviously outside the bounds of a project this size.

If POCUS makes any contribution to the theory of centering, it is in the suggested ordering of hypotheses described above, which plays the same sort of role as the pronoun interpretation rules do in Sidner's model. If the idea of centering itself is valid, then perhaps some such ordering could be empirically determined. Yet semantic and pragmatic effects have such a strong influence on the way pronouns are interpreted that they tend to mask any evidence for such an ordering. For example, the following are successfully dealt with by the program, but cannot be used as evidence because of se-

mantic biassing:

> Mark loved Alison.
> He gave her some flowers.
> They were nice.
>
> Mark loved Alison.
> His friends thought they were happy together.

These examples were intended to show, respectively, that reference to an f-centre (the flowers) is a higher priority than reference to a conjoined set (Mark and Alison), and that reference to a conjoined set is a higher priority than reference to something in the same sentence (his friends). It is difficult to find examples that aren't biassed in this way, and those that aren't tend to be genuinely ambiguous, casting doubt on the idea of a clearly ordered set of preferences. Compare the above examples with these, which should demonstrate the same phenomena, but which are more pragmaticlly neutral:

> Mark smiled at Alison.
> He gave her the earrings.
> Everyone watched/stared at/looked at/admired them.
>
> Mark wanted to make Alison happy.
> He threw a party for her.
> Everyone thought they would enjoy it.

In the first example, "them" could mean Mark and Alison or the earrings, and a slight change of wording, or some additional clues from the context (e.g. if we could work out why everyone was looking) could make either interpretation more likely. Yet POCUS would quite definitely favour the second interpretation. Similarly, in the second one, "they" could equally well refer to Mark and Alison or to everyone else. For POCUS it could only mean Mark and Alison, but for people it is at best ambiguous.

Semantics even seems to override the centre rule sometimes. The following is consistent with Grosz's model and could be handled by the program:

> He gave the book to Mike.
> It amused him to see ...

where "him" is the same as "he", and has been identified by the preceding cotext. However, the following example, in which "him" is Mike, cannot be accounted for:

> He gave the book to Mike.
> It amused him.

It seems reasonable to suppose that this is because we assume that Mike is in possession of the book at the time that the second sentence describes, and hence is the one it will amuse. If the second sentence had read "It had amused him." we would not make these assumptions, and so would interpret "him" as referring to whoever gave Mike the book. Similarly, in the next example the interpretation of "he" depends on whether the adjective in the last sentence is "useful" (he = Pete) or grateful (he = Mark).

> Mark loaded up the car.
> Pete helped him.
> He was very useful / grateful.

Furthermore, it is even possible to switch the interpretations round by imagining a context in which Mark's usefulness to Pete, or Pete's gratitude to Mark is being offered as an explanation for Pete's helpfulness. This is in accordance with the idea that one of the most important criteria for determining the reference of an expression is whether or not it leads to a coherent interpretation for the text. Of course, Grosz's rule is only being offered as a maxim, and

it might be suggested that examples like this one, where the rule is ignored, demand more complex inferences to be made and perhaps take more time to process. However, one would at least expect the rule to determine the more semantically ambiguous cases. Yet if the last sentence is changed to the more neutral

He was very careful / cautious,

it becomes difficult to decide who is being talked about, with "careful" slightly suggesting Pete, and "cautious" perhaps favouring Mark!

Semantics also affects the centering of entities not referred to by pronouns. Given

Mark ignored Pete.
He was tired.

POCUS plausibly interprets "he" as Mark, because given two NPs to centre it tests them in the order in which they occur in the sentence. A better account would probably be based on case, but the example seems fairly ambiguous anyway. It is easy to think of contexts for both interpretations (Mark ignored Pete because he was too tired to respond; Mark ignored Pete because Pete's behaviour was caused by tiredness), and once again slight changes in wording can favour one over the other. For example, consider these alternatives to the second sentence:

He was obviously tired	(he = Pete)
He looked tired	(he = Pete).
He was too tired	(he = Mark)
He felt tired	(he = Mark)

At the very least, we need a complete set of centering rules or con-

straints based on semantics to deal with this.

c) Parallel construction.

An equally difficult problem to solve concerns the way in which the construction of a sentence can affect the interpretation. Consider the following:

1 Mark was the leader.
Bill liked him.
He liked the way he treated him.

A sentence like the last one is ambiguous in that the binding rules allow the first pronoun to be coreferential with either (or neither) of the other two (but not both). Hence if there is some preferred order for possible referents we might expect to see evidence for it in such sentences. Given this example, POCUS interprets the last sentence to mean that Mark liked the way he treated Bill, because Mark is the centre and so the preferred choice for both the first two pronouns (the last must then refer to Bill because of the binding rules). This does not seem to be the natural interpretation - it is more likely that Bill likes the way Mark treats him. However, changes in wording can lead to different interpretations:

2 Mark was the leader.
Bill admired him.
He was proud of the way he treated him.

3 Mark was the leader.
He was Bill's friend.
He was proud of the way he treated him.

4 Mark was the leader.
Bill liked him.
He was proud of the way he treated him.

5 Mark was the leader.
He was Bill's friend.
He was proud of the way he helped him.

6 Mark was the leader.
Pete liked him.
He was proud of the way he helped him.

7 Mark was the leader.
Pete admired him.
He was proud of the way he helped him.

8 Mark was the leader.
Pete liked him.
He liked the way he helped him.

9 Mark was the leader.
He was Pete's friend.
He liked the way he helped him.

10 Mark was the leader.
Pete admired him.
He liked the way he helped him.

11 Mark was the leader
Pete admired him.
He liked the way he treated him.

There do seem to be certain regularities in the way one interprets these, but they do not seem to reflect any preferred ordering such as that embodied in POCUS. First, parallelism has a strong effect on the interpretation of the first pronoun. Examples 1,3,5,8,and (to a lesser extent) 11 all show parallel construction between the last two sentences, and in all of them the corresponding pronouns receive corresponding interpretations. Second, the verb used in the subordinate clause affects its interpretation. All the sentences that use "treated" seem to concern Mark's treatment of Bill, whereas "helped" is more ambiguous, slightly favouring a reading in which Bill helps Mark. This is because Mark has been set up as the dominant partner in the relationship, and so is more likely to treat Bill in various ways and be helped by him. If the opening sentence was "Mark was the tea-boy" the opposite might be true. Third, the choice of main verb in the last sentence also affects the interpretation of the subordinate clause. A person is more likely to be "proud" of his own

behaviour, and to "like" that of someone else. This is particularly noticeable when the subordinate clause is the more ambiguous "he helped him".

Once again, these effects can override the centre rule, too. Here is an example where parallelism determines what the centre is:

> Mark was interesting.
> Pete liked him.
> He was .../ he liked ...

In the absence of other contextual clues, whether "he" refers to Mark or to Pete will depend on whether the last sentence is parallel to the first or the second, despite the fact that Mark is unambiguosly centered in both cases.

So, while the connotations of the verbs used and the effects of parallel construction seem to play a major role in determining how pronouns are interpreted, once again there is little evidence for an ordering of hypotheses like that used in POCUS. This may be an important lesson - such ordering effects are the basis of Sidner's model too.

It is interesting to look at one of Sidner's examples in the light of these observations about semantics and parallelism. Sidner finds it necessary to introduce a "theme" case role in order to deal with examples like:

> I haven't seen Jeff for several days.
> Carl thinks he is studying for his exams.
> But I think he went to the cape with Linda.

Grosz et al then go on to point out that this is easily predicted in their model, without resorting to case at all. But let us look at

the example more carefully. For a start, it is very difficult to construct an interpretation in which the second sentence means that Carl thinks Carl is studying, because he would know whether such a thing was true or not, and anyway that would make the sentence seem a non sequitur in most conceivable contexts. Furthermore, the parallel construction of the last two sentences , combined with the use of "but" to suggest a contrast between them, strongly implies that "he" is the same person in both. So these factors alone seem to make the given interpretation inevitable, regardless of any notions of local focus.

This does not mean that the same phenomenon cannot be found in a more neutral example. Here is one which is not quite so bad:

> We decided to take Jeff with us to the Cape.
> Carl said he was studying for his exams.
> I didn't think he would be so foolish.

The absence of the structural and semantic cues we have discussed makes this text seem less cohesive and more ambiguous when presented out of context. There are at least five possible interpretations:

1 Carl said Jeff was studying,
but I didn't think he would be.

2 Carl said Jeff was studying,
and I was surprised at Jeff's foolishness.

3 Carl said Jeff was studying,
and I was surprised that Carl was so
foolish (as to say such a thing).

4 Carl was studying,
and I thought he was foolish to do so.

5 Carl was studying,
but I didn't think Jeff would be.

Which of these seems more plausible depends on the context, and also

to some extent on intonation (in speech), but the most natural interpretation seems to be either 1 or 2, and both are compatible with the centre rule. Even if intonation, say, suggests that Carl is the centre of the second sentence, as in

> We decided to take Jeff with us to the Cape.
> CARL said he was studying for his exams.
> I didn't think Jeff would be so foolish.

the tendency is to assume that a shift of focus has occurred and go for interpretation 4, so that the only natural way to get back to Jeff in the third sentence would be to refer to him by name. So the centre rule does still seems to work here, provided we ignore the fact that intonation can lead to shifts of focus that seem to violate it, and this might be explained by adding new centering rules of a higher priority, or may even be a matter of global focus rather than centering. However, it is clear that the issue is not as straight- forward as first appears, and it is good practise to look out for the influence of semantics, parallelism, etc. when discussing such examples. People clearly use all the clues available to them when trying to understand one another, and it is therefore difficult (and sometimes perhaps artificial and misleading) to isolate any one interpretation strategy. In particular, it is dangerous to ignore the influence of content and context, as is often the case when linguists and psychologists consider short constructed texts like those handled by POCUS.

d) Centering.

As mentioned in section 2, POCUS is not always able to find a unique centre for every sentence in the text, but rather has to keep a number of possibilities in focus and "home in" on the right one gradually. This seems to be a slight deviation from the theory of Grosz et al, who imply that there would only be one centre. However, as noted in section 1, their theory is in fact written from the speaker's point of view, and the speaker will have much more of an idea of where the discourse is going. So perhaps POCUS is a reasonable model from this point of view. Certainly it explains why, given the initial sentence

Keith met Pauline in a bar.

any of the following sentences would make for a natural continuation:

He was looking for someone else	(centre = Keith)
She was looking lonely	(centre = Pauline)
They started to chat	(centre = Keith and Pauline)
He bought her a drink	(centre = Keith or Pauline)

Furthermore, when centres are being stacked, it does seem necessary to place all possible candidates on the stack, and not just those entities uniquely identified as centres, because there are cases in which the pronoun interpreter has to hark back to them. For example, let us extend the example of section 2:

Mark gave Alison the gun.
It was making him nervous.
It looked too conspicuous.
She could hide it better than he could.

According to Grosz et al, the gun is uniquely identified as the centre of all four sentences, because it is continually referred to by

a pronoun. And yet the pronouns "he" and "she" in the last sentence hark back to entities which were merely possible candidates for the centre. This suggests that "possible centres" might actually be behaving like unique centres. Of course, it may be that more centering rules will make it possible to always find a unique centre, but, if not, a move away from the idea of a single local focus seems necessary (for hearers, at least).

There seem to be some cases where the centre rule is violated. Consider the following text:

> Mark was the leader.
> He had been with them for four years.
> His assistant was called Pete.
> He was rather thin, with grey eyes.
> The others didn't trust him, although
> Mark knew he was the best man in the group.

After the third sentence, Pete becomes the centre, even though the centre continues to be pronominalized, and there are no pragmatic reasons for ruling out Mark as the referent for "he" in either of the subsequent sentences. It seems that something in the construction of the third sentence shifts Pete into focus.

Similarly, there also seem to be cases where the centre rule is violated from the point of view of pronoun interpretation. Consider the following examples:

> Mark was angry.
> Pete did not trust him.
> He disliked the way he treated him.

> Mark was angry.
> He did not trust Pete.
> He disliked the way he treated him.

According to Grosz, Mark should be the centre of the second sentence in both cases, because he is referred to by a pronoun, and (in most

cases) if there is only one pronoun it must refer to the centre. This leads POCUS to interpret the last sentence as meaning that "Mark disliked the way Pete treated Mark" in both cases, since the centre (Mark) is its first choice for the referent of the first pronoun (He). However, this is the wrong interpretation for the first case. So either Mark is not the centre in both cases, in which case the centre rule must come into question, or the strategy POCUS uses is deficient in some other way.

Neither of these is a fatal counterexample to the theory, of course, because the centre rule is offered as a maxim, and not an inviolable rule. It may be that the addition of new centering rules, perhaps based on case, might solve the problem.

e) Reference to sets.

POCUS correctly interprets a number of examples where individually introduced entities are referred to as a set, as is shown in appendix 3. But there are problems here. Such entities would normally have to be in focus to be joined together like this, and there seem to be some restrictions on how things are grouped together. POCUS operates on the principle that only animate entities that are centred can be joined together. It is restricted to animates to explain examples like these:

Mark passed Alison on her bicycle.
They were going quite fast.

Pete looked at the house.
They (both) looked old now.

In the first of these, it would be unusual to interpret "they" as Alison and her bicycle, and the second makes no sense at all unless "both" is included. One of the functions of words such as "both", "all" and "together" (and perhaps also of phrases like "each other") seems to be to indicate that just such a conjunction should occur, as in

Mark and Alison went to a pub.
There they met Jacquie and Pete.
They (all) decided to go to the cinema (together).

where "they" would be interpreted as just referring to Mark and Alison without the "all...together" construction.

So it seems more difficult somehow to refer to sets like this if they include inanimates. But the strategy of restricting conjunction to animates is clearly only a partial solution. Consider the following example:

Pete put the vase next to the clock.
They looked very dusty.

This example indicates that inanimates can be joined to one another quite happily. So perhaps the difficulty is just in joining animates to inanimates. Perhaps this might be explained in terms of semantic restrictions, or in terms of case.

A further example suggests that even POCUS's strategy of restricting this sort of conjunction to centred entities may be wrong:

Pete heard a noise.
He looked at Jacquie.
Someone was following them.

Pete is quite clearly the centre of the second sentence in the model of Grosz et al, and yet becomes joined to Jacquie, who is not centred. Moreover, if additional centering rules make it possible to identify a unique centre for most sentences, rather than focussing in gradually as POCUS does, then the strategy POCUS uses for conjunction will become unworkable because there will be no sets of centered entities to join together. This may be further evidence for the existence of several foci with relative degrees of foregrounding.

Finally, there are some cases where POCUS joins things together when it shouldn't. For example, given

Mark ignored Alison.
They knew she was lying.

POCUS assumes that "they" means Mark and Alison, but most people would tend to assume that it must refer to some other group of people referred to earlier on, unless "both" were used:

Mark ignored Alison
They both knew she was lying.

In terms of the program, this would mean searching the stack rather than joining Mark and Alison together.

So there is obviously a lot of room for improvement in this area. It would be particularly interesting to look at the way the "both" and "all... together" constructions are used to signal this sort of conjunction.

f) Sentential anaphora.

POCUS has not been tested much on sentential anaphora, and only two examples of forward pronominalization have been used, partly because of the limitations imposed by the grammar used. They are:

On his table Pete saw a book.

Although he loved it, Pete sold the Anglia.

both of which are correctly interpreted. In Sidner's model, a pronoun is only interpreted as a cataphor if all else fails, whereas POCUS would prefer to treat it as one rather than have to look through the stack for some previous centre. It is not clear which approach is more satisfactory. Consider the following:

> Helen looked out of the window. She didn't want to watch Bill. He had drunk far too much, and was behaving very badly. Although she was used to that sort of thing, his wife was looking upset.

In the last sentence, "she" could refer to Bill's wife, making the pronoun a cataphor, or it could refer to Helen, a previous centre, or indeed the NP may be attributive and Helen may actually be Bill's wife. POCUS reads it as "although Bill's wife was used to that sort of thing, she was looking upset", but it seems to be somewhat ambiguous - intonation, or additional knowledge about the characters in the story could be used to force either interpretation. Part of the problem may be that some sort of global effect is involved here - because the description of Bill's behaviour is introduced to explain Helen's, the reader is expecting to return to Helen (c.f. Reichmann). Such directed shifts are probably much more likely than the sort of arbitrary return to the last topic that POCUS uses when it gets into trouble, so it may be that if POCUS could be extended to

deal with them by adding a proper global focussing mechanism there would be no need for a centre stack anyway, and the problem would not arise.

Once again, no account has been taken of possible case effects in sentential anaphora.

g) Global focus

POCUS includes a first attempt to integrate the theory of centering with other work, such as that of Grosz (1981) and Reichmann (1978), on global focus. In POCUS, changes of focus can be induced "by hand", simulating the action of a separate global focus mechanism. For example, compare the difference that the addition of the word "similarly" makes to the interpretation of this text:

> Mike was a doctor.
> Michelle loved him.
> (Similarly,) Susan liked the way he treated her.

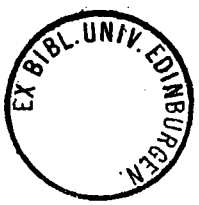
In one version "her" seems to refer to Michelle, in the other to Susan. By supposing that "similarly" acts like one of Reichmann's "clue words", POCUS is able to get the right interpretation in both cases. Similarly,

the example quoted previously

> We decided to take Jeff with us to the cape.
> CARL said he was studying for his exams.
> I didn't think he would be so foolish.

in which intonation signals a shift of focus onto Carl, would be correctly dealt with by this system. Although this simple stack-based approach is obviously not a realistic account of the way in which shifts of global focus occur - both Grosz and Reichmann agree that specific linguistic strategies are used to direct shifts of focus towards particular entities - and has not been extensively tested, there does seem to be considerable scope for progress in this area. As well as clarifying the relationship between local and global focus, further work on "clue words" may unearth results analogous to Reichmann's at the local level (e.g. in the use of con-

junctions like "and" and "but"). The centering approach may make it relatively easy to formalize some of these notions.



h) Comparison with Sidner's Model.

The main question is: how well does POCUS cope with sentences with multiple pronouns? These are the sorts of examples that Sidner's theory of actor focus was postulated to explain, and the main task in constructing POCUS was to come up with an alternative explanation within the theory of centering, given that Grosz et al do not suggest one as such. A partial solution is provided by the ordering of the hypotheses - POCUS will always try to interpret pronouns as centred items if it can, and those left over will tend to be interpreted as f-centres. All that remains is to sort out what happens when a choice has to be made between several centres or f-centres. Here the solution is crude but moderately effective: in both cases they are chosen in the order in which they occur in the discourse. This seems to capture some of the parallelism of examples like:

Pete looked at Bill.
He hated him.

which Sidner needs to use case to explain. However, there are still many examples that Sidner's model can handle and POCUS can't. The example cited above,

Carl is worried about Jeff.
He thinks he studies too much.

would be misinterpreted as meaning "Carl thinks Carl studies too much", and there are other similar examples like:

Mark was the leader.
Bill liked him.
He liked the way he treated him.

where the last sentence would be interpreted as "Pete liked the way

Pete treated Bill". So, although we can now treat examples like (1) fairly simply, we have not yet solved all the problems associated with this type of example. However, before rejecting centering theory as inadequate, let us look more closely at the last example.

There are two separate problems here. The first is the interpretation of the pronoun in the main clause, and here Sidner's model seems to have the advantage. It is difficult to construct examples of this form where the subject of the last sentence is not taken to be the same as that of the previous one, although there are examples involving "point of view" phenomena (Chafe 1975) where Sidner gets this wrong too, such as

Pete looked at Bill.
He was obviously tired. [both get it wrong]

Pete looked at Bill.
He felt tired. [both get it right]

So it is clear that we need some theory of how a pronoun's position in the sentence affects its interpretation, and that Sidner's theory of actor focus is probably better than one based on order of occurrence, such as that embodied in POCUS. However, the literature of discourse analysis contains a number of alternative theories, such as Halliday's theory of "thematic structure" (Halliday 1967), or Chafe's ideas about "packaging phenomena" (Chafe 1975) using a variety of properties such as subject, theme, topic, point of view, etc rather than case, and it may be that something even better could be constructed within the confines of centering theory using these ideas.

The second problem concerns the interpretation of pronouns within the embedded clause, and it is not so clear that Sidner has

the right answer here. Examples like

> Mark was the leader.
> Bill admired him.
> He was proud of the way he helped him.

can be constructed where the effects of pragmatics or parallel construction can change the interpretation. The effects of pragmatics have been noted before, e.g.

> The councilmen refused to give the women a permit for demonstration because they feared violence / advocated revolution (Winograd 1972).

in which the meaning of "they" is different in each case, and

> Mother made some cookies and left one out on a plate. She put the plate on the kitchen table and went into the living room. "I'm sure Janet will like it", thought Mother. (Charniak 1972)

in which "it" is generally taken to mean the cookie for pragmatic reasons, although a context could be found in which "it" might mean the plate (note that the cookie is not centered by the time the pronoun occurs). These effects are quite insidious. Example (1), which posed a slight problem to Sidner, is a case in point.

> I haven't seen Jeff for several days.
> Carl thinks he is studying for his exams.
> But I think he went to the Cape with Linda.

There are two problems here. First, the interpretation "Carl thinks Carl is studying for his exams" is ruled out on pragmatic grounds. Second, the parallel construction of the last two sentences, combined with the use of "but" suggests a contrast and hence a parallel interpretation of the pronouns. So it is dubious to conlude anything from this example as it stands. A more neutral example with a similar form might be

We decided to take Jeff with us to the Cape.
Carl said he was studying for his exams.
I didn't think he would be so foolish.

but this seems to be ambiguous at the very least.

These effects make it difficult to test models like POCUS or Sidner's reliably. For instance, one of the main empirical predictions made by a serial "hypothesis-testing" model is that certain interpretations will always be preferred in certain otherwise ambiguous cases. For example, the order of hypotheses in POCUS suggests correctly that in

Mark smiled at Alison.
He gave her the earrings.
Everyone admired them.

"them" is interpreted as "the earrings" and not "Mark and Alison". But slight changes of wording can change the interpretation, e.g.

Mark smiled at Alison.
He gave her the earrings.
Everyone watched them.

and the most semantically neutral examples just seem ambiguous:

Mark smiled at Alison.
He gave her the earrings.
Everyone stared at them.

It is tempting to conclude that there is no real order of hypotheses, and that all possible alternatives are considered in parallel, with semantic and pragmatic information eventually being used to make a choice. Certainly, people do seem to delay these decisions sometimes, so that "he hit him" is often just interpreted as "someone hit someone else". Whatever shape a decent theory takes, though, it must take account of these pragmatic effects.

4) CONCLUSION

POCUS shows that it is possible to build a reasonably successful working system based around the theory of centering, and its performance indicates how a more complete model might be built. More work could be done on reference to embedded NPs and to conjoined sets, on the effects of thematic structure on focussing, and on interfacing the local focus system with one for global focus. A more radical change would be to extend POCUS to try search for referents for non-pronominal NPs. This would necessitate abandoning the naive approach used so far, in which "finding a referent" really means finding an antecedent NP, and attempting to formalize the suggestions of Grosz et al about the relation between NPs and the entities that they realize, with all the problems of "speaker's references", "implied entities", etc.A proper semantic component would have to be introduced, with access to knowledge about the real world, and sentences would have to be processed for meaning and incorporated into a full representation for the discourse. This would then open the way for a more sophisticated approach to semantic constraints - for instance, a frame-based reasoning system might be used to check the likelihood of possible interpretations, rather than merely using selection restrictions. The interaction between the parser and the centering mechanism would also have to change in character - the latter would now be dealing with semantic representations rather than NPs and parse trees.

This interaction itself demands more careful consideration. It has been assumed that pronouns are only interpreted after the

sentence has been parsed, and probably after it has been processed for meaning as well, yet people can obviously understand pronouns when they are used in incomplete and elliptical sentences, probably often make decisions about their meaning before the sentence has been parsed, and may even use top-down information gained from reference resolution to guide parsing. Perhaps a more realistic model would have syntactic, semantic and focussing components running together in parallel. This needs further investigation. A related problem of interest concerns the treatment of ambiguity. What happens to pronoun interpretation when the meaning of a sentence is ambiguous, and how is a sentence incorporated into the discourse representation when some of its NPs are not unambiguously interpreted? Unlike POCUS, people often seem to simply reserve judgement about reference in the more ambiguous cases, and then tie up the loose ends when more information arrives. A model in which such sentences are partly processed, put aside in some stack, and then fully incorporated into the discourse a while later might be more realistic.

But before any of these changes would be worth making, more work must be done to determine whether the very idea of centering has any value at all. Is it a viable alternative to, say, Sidner's theory? In so far as POCUS is successful, it is probably because it shares many of the features of Sidner's model. And the sort of centering used in POCUS could perhaps be regarded as a deviation from the original theory of Grosz et al. Finally, because of the difficulties that the effects of pragmatic inference and parallel construction present, it is difficult to find any really convincing

evidence for the pronoun interpretation rules offered here, and indeed at least one of the established examples from the literature seems to come under suspicion for the same reasons, so no really decisive tests or comparisons have been possible. Perhaps the real lesson here is that no model that ignores the role of pragmatics in determining the reference of an expression can ever be called complete.

PART THREE USING FOCUS IN DISCOURSE ANALYSIS

1) INTRODUCTION.

POCUS shows that centering theory can be recast in a procedural form and demonstrates the potential value of local focus as a model of the way people interpret pronouns. But clearly there's a lot missing from the model. There are only the very crudest attempts to account for thematic and global effects, and no account at all is taken of pragmatics or structural parallelism. These missing elements make it difficult to test the model conclusively, since each time a counterexample is met they can always be held responsible. As a result, the examples shown in Appendix 3 can do little more than show off the range of problems that POCUS was designed to tackle, and say little about its value as a realistic model of discourse processing.

This may be a reasonable criticism of all computational linguistics, and indeed of cognitive science in general, that when our models fail to work we can always "pass the buck" onto some as yet unspecified process (hence the syndrome in AI where large amounts of time are spent discussing yet to be written programs - c.f. McDermott 1976). Hence it seems that computational models, while useful, do not constitute falsifiable hypotheses about human cognition. There has been a lot of talk over the years about whether or not there is such a thing as "the cognitive paradigm" (see, for example, Dresher and Hornstein 1976 & 1977, Winograd 1976), but whatever that means (and Kuhn himself uses the word "paradigm" in at

least twenty-two different ways in his book according to Masterman 1970), there is one important sense in which we clearly don't have one: we don't have any sort of agreement about the overall structure of the human language processor. And without that infrastructure we are in the same position as the pre-Newtonian astronomers - we can always add another epicycle to our models to "save the phenomena". In a sense, then, Dresher and Hornstein (1976) may be right when they say that the aims of AI are technological rather than scientific, although that does not mean that it's not a valuable pursuit, because the tools that we are developing are intellectual ones that will hopefully provide a vocabulary that will one day enable us to frame mature scientific hypotheses about cognition.

In this light, there are a number of ways in which we could proceed. POCUS could be extended to cope with other sorts of anaphora, taking account of things like pragmatics, parallelism, thematic structure, and global focus in some limited way. But that would mean adding a semantic component, a knowledge representation system, a reasoning component, etc., which is clearly beyond the scope of a project this size, and besides there is a good chance that in trying to solve those problems we would lose sight of the original issues. Another option would be to try to extend POCUS a little, so as to run it on some natural data, but that would still be difficult and the points raised above suggest that it would not necessarily be of any great value. So instead I propose to try to extend the theory of focus, but in an informal way, in the hope that this will suggest some new focussing strategies that could be implemented in a programme like POCUS. I will look at both local and glo-

bal effects, and in particular look at the way they interact.

So far, POCUS has only been tested on a small corpus of about thirty examples, consisting of specially-written "stories" of two or three lines each. This is obviously not natural data in any sense. For a start, the syntax used is very simple (partly because of the simple parser used), and all the examples consist of complete sentences. This is particularly disagreeable to those who take the view that spontaneous speech is the primary source of data for discourse analysis, since speech is often elliptical and fragmentary, consisting as often as not of sequences of phrases rather than sentences. Furthermore, because the examples are mostly stories about people, the foci are usually drawn from a very small cast of characters, often introduced by name and distinguishable by gender. Longer texts of other types, interactional, transactional or ideational, present us with more complex problems and may reveal new focussing strategies. Indeed the very length of a text might be expected to determine, say, how prominent global effects are.

So in this part of the paper I will examine some real data in the light of ideas about focus. Because of the difficulty and dubious value of extending POCUS to cope with such data, the analysis will perhaps be less formal from here on, but I hope this will be offset by being able to look at a richer range of focussing phenomena.

The text to be considered was transcribed from a T.V. programme about gardening. There is no agreed convention for representing speech, and in a sense the choice of representation is always

subjective, since each researcher tends to choose one that includes all the features relevant to his own style of analysis without fogging the issue with irrelevant complexities. As Kuhn says, there are no facts without theories. All the ideas about focus I have presented so far have been about sentences, so in transcribing the data I have as far as possible tried to use rhythm, intonation, syntax and sense to divide it into sentences and clauses. Where this was not possible, I have simply marked apparent boundaries between constituents (e.g. phrases) with a dash (-). Where one speaker overlaps another, ellipsis marks are used (...). Prosodic features are not marked, although I will make some remarks about stress in the analysis. Some important visual events related to deixis are indicated.

The programme takes the form of an informal interview, with the presenter, B, acting as the interviewer. The two participants seem to speak fairly spontaneously, although some topics may well be prepared in advance, with B guiding the discourse. This pre-planning may affect the global structure of the text, but no discourse is ever completely free of it (no coherent discourse, that is). In even the most casual conversations people are constantly planning ahead to introduce new topics, and this gives them much of their coherence. However, the amount of planning may have some affect the focussing strategies used, as I will discuss later when comparing this spoken text with part of a written one.

What follows is a preliminary analysis of the text made on the basis of a few simple rules pretty much akin to those used by POCUS. Thereafter, I will look at various new focussing phenomena in

the light of this analysis, suggest some new rules, and finally present a complete analysis of the text in terms of them.

2) PRELIMINARY ANALYSIS.

a) Potential Foci.

In POCUS, the potential foci were simply those items referred to by noun phrases (strictly speaking, of course, it only dealt with the NPs, not their referrents). But in a real text the situation is more complex, because not all the NPs refer to items that can be focussed, and not all the potential foci are explicitly mentioned with NPs.

Non-referential NPs include attributive usages that do not introduce separate potential foci, such as:

A: 'That's THE HONEY DEW THAT THEY EXCRETE.

A: 'Those are SOME BUTTERFLY EGGS.'

B: 'Yes, Quosia's rather A GOOD ONE.'

B: 'Is this called A CHARM CHRYSANTHEMUM?'

Recognising such NPs is not a trivial problem of course, but in the preliminary analysis I shall assume that they have been identified and do not contribute additional potential foci to the discourse.

Just as we can no longer equate NPs directly with the potential foci, so the relationship between pronouns and the focus system becomes more complex. Some don't refer at all, or only do so vaguely:

'THEY say IT will rain today'

'IT's hard to say if it will.'

Other pronouns refer to states of affairs or actions:

'I saw you rob the bank.
Two other people saw IT as well.

'I saw a bank get robbed.
You see THAT quite often here.'

'I robbed a bank today.'
'Why did you do THAT?'

Again, I suspect that many of these usages have nothing to do with focus, and will not regard them as presenting potential foci for the time being.

A further category of utterances that do not seem to play an obvious focussing role are those in which the speaker expresses his attitude to something just said, or to some state of affairs. These "metastatements" range from simple acknowledgements like "yes" to prolonged discussions, often involving anaphoric expression referring to bits of the discourse itself, such as:

'THAT ARGUMENT doesn't hold water.'

'THE POINTS I HAVE JUST RAISED illustrate something else.'

'THAT's an interesting argument.'

These sorts of sentences abound in academic writing, of course, and may be a sort of global focus shift, but I shall regard them as out of the range of the present discussion and assume that they do not play a focussing role, an assumption which seems to work fairly well, in this text at least, where they are used sparingly and always as small digressions from the main topics.

Similarly, I will assume that personal pronouns referring deictically to the two speakers are not potential foci either. This may seem a strange assumption, but once again it works well in a

text like this where such pronouns are only used to identify a viewpoint, and where the speakers themselves do not seem to be the topic of conversation. In a sense things like the speaker's identity, the time and place of utterance, etc. are always in focus - they form the context for what is said, but these "point of view phenomena" (c.f. Chafe 1975) seem to operate separately from the local focus system.

Finally, just as not all NPs are potential foci, so not all potential foci are explicitly represented by NPs. Elliptical expressions may be used to focus some items, as I will discuss later on.

b) Rules For Analysis.

For the initial analysis I will assume that the pronoun interpretation rules used in POCUS are more or less correct, with two reservations. Firstly, I will no longer use the rule whereby all the items in local focus can be referred to as a set. Much of the time in this analysis a unique focus exists, and there are no examples of such reference to sets in the text, besides which there are too many problems with this rule, which might be better dealt with as a global strategy. Secondly, I will assume that the "Hark-back" strategy, whereby the stack is searched is a last resort, and that before doing that the hearer tries the global focus. So the new rules are:

When a pronoun is found try:

a) The local focus.
b) The potential foci of the last utterance.
c) Those of the current utterance.
d) The global focus.
e) Previous foci.

The terms "local focus" and "potential local focus" replace "centre" and "f-centre". "Utterance" is used to allow for the possibility of elliptical phrases, etc. and I will assume that any two parts of a sentence separated by an interjection from the other speaker (such as an acknowledgement) constitute separate utterances (there is some evidence on the basis of this that an analysis based on clauses, not sentences would be better).

Foci are then assigned to utterances by these rules:

To find the local focus, try:

f) Any item referred to via rule 1.
g) Any other item referred to anaphorically.
h) Any other potential local focus.

In addition, we add this rule for finding global foci:

> i) When an item not recently mentioned or in local focus is referred to anaphorically, it may be the global focus.

Using these rules I will try to assign foci to the utterances of the text, bearing in mind that, as mentioned in Part One, although it is tempting to think of some utterance as "having a focus", it is actually the speakers themselves who do the focussing, and they may occasionally disagree about what's in focus. Some simple disagreements do occurr in the text, and they must be recognised in order for foci to be assigned correctly. For example, A here tries to move on to a new focus, but B continues to talk about the old one:

A: 'Shan't have to take THIS home. Um - there's a little blackfly ...'
B: 'You're going to leave IT here with us, are you?'

This simple example would confuse a program like POCUS that treats focus as a property of the text rather than of the participants - it would assume that "it" meant the blackfly. It is easy to think of more difficult examples, where the speakers are at cross-purposes for long sections of the discourse, and maybe have to construct multiply embedded models of each other's foci in order to communicate properly.

c) Analysis.

With these precautions in mind we can begin the analysis. The following is an extract from "Gardening Time", an STV television programme. Using the rules given above, it has been marked as follows:

Local foci are capitalised thus:	GREENFLY
Global foci are marked thus:	/greenfly/
Non-referring expressions and "metastatements" are marked:	<greenfly>

The two participants, A and B, are talking about chrysanthemums and examining one that is standing in front of them. Greenfly are referred to by a pronoun later on, so must have been in focus in the first utterance:

A: 'and then if you look inside the leaf there you can see perhaps a hundred GREENFLY even in that little ... ' [points to leaf]
B: 'Where does the STICKINESS come from?'
A: 'THAT's <the honeydew that /they/ excrete>.'
B: '<I see.>'
A: 'And IT just collects over the plant.'
B: '<Yes.>'

There is only one potential focus when honeydew is first introduced, and it is pronominalised on the third mention, so is probably in local focus throughout the above section by rules a and f. Harking back to greenfly sets them up as global focus. Honeydew seems to stay in focus by interpretation rule a and focussing rule a in what follows, although later we will re-examine this claim:

A: 'And you get fungus growing on that as well. And some of the trees you see that - a black fungus growing on IT.'
B: '<Do you really? Very nasty indeed.>'
A: 'which you notice. And also LADYBIRDS are very useful. There's a little LADYBIRD there look.'
B: 'Now what does SHE do? SHE eats /them/, does SHE?'

A: 'Yeah, SHE eats /them/. Some people think that they've got to kill everything, but ...'

Initially ladybirds are the only possible foci. Pronominalization keeps the ladybird in local focus by rules a and f, while harking back to the greenfly puts them in global focus. They in turn then move into local focus by the rules d and f:

B: 'Ants milk THEM ...'
A: 'Yes, ants milk ...'
B: 'And ladybirds eat THEM.'
A: '<Yes. Yes, it is strange, isn't it?> But the less that you kill LADYBIRDS AND THEIR LARVAE the better, because THEY'll certainly eat their way through a few. <So that's quite helpful.>'
B: 'Well, let's just hope that SHE won't fly away home.'

Ladybirds are reintroduced into local focus by a full NP and subsequently pronominalized. In what follows the eggs are initially the only foci:

A: 'Oh look, there's ...'
B: '<What>'s THAT YELLOW THING there?'
[points to eggs]
A: 'Yes, I hadn't seen THOSE.
THOSE are <some butterfly eggs>.'
B: '<Oh, I see.>'

Subsequent pronominalization suggests the eggs are in focus here too, not the cabbage white:

A: 'Yes, THOSE are from the cabbage white.'
B: '<Uh-huh.>'
A: 'THEY'll get killed before THEY hatch out.'
B: '<Yes.>'
A: 'Otherwise THEY would cause a bit of a problem.'
B: '<Uh-huh.>'
A: 'And there's a little BLACKFLY'
[points to blackfly]
'We've got everything on this one, haven't we?'
B: '<Yes, we have, haven't we? Yes.>'
A: 'Shan't have to take THIS home.
Um - there's a little BLACKFLY ...'
B: 'You're going to leave IT here with us, are you?'

B fails to follow A's new focus. His pronominal reference to the

plant suggests it was the focus when A last mentioned it. A series of utterances follow with only one potential focus apiece:

A: '<Yes, I wouldn't be very popular if I did that, would I?> But - um - BLACKFLY AND GREENFLY are <the two commonest problems I think that you get>.
B: 'Yes. Now what's THE CURE then?'
A: 'Any of the INSECTICIDES that are recommended at your local garden centre.'
B: '<Yes.>'
A: 'But there's a good one here - MALATHION - THAT's fairly safe.'
[holds up a bottle of Malathion]
B: '<Yes.>'
A: 'I don't think you want anything too powerful.'
B: 'No, I like THE PYRETHRUM ONES for <that reason>.
A: 'Yes, PYRETHRUM. Yes, DERRIS as well, THAT's er - and we use NICOSOAP - NICOTINE.'
B: '<Oh, do you? Yes. Yes.>'
A: 'We find THAT's quite good.'
B: 'and QUOSIA? Do you use THAT?'
A: 'No, we haven't got round to .
B: '<No.>'
A: 'I'll have to try THAT ONE.'
B: 'Yes, QUOSIA's <rather a good one>.'
A: 'But <it>'s also important to change THEM round.'
B: '<Hmm.>'

The foci are not well-defined immediately after this. Tentatively we might say that "it" refers to Malathion and so puts it into focus, while harking back to blackfly and greenfly put them in global focus. But clearly this section is intended to be vague, meaning pests (in general) get used to an insecticide (in general).

A: 'I mean <it>'s no good saying "Oh I always use Malathion. As soon as you see any ..."'
B: 'Because they get used to it.'
A: 'Yes THEY do. And eventually you get some that even thrive on IT.'
B: '<Oh, I say.>'
A: '<It's like the Super-rats and things like that.> But THAT's <a good standard one.>'
[Holds up the bottle again.]

Twice immediately hereafter a demonstrative followed by ordinary pronouns seems to be in focus:

B: 'Now how - let's just lift THIS up onto the table so that we can see what a very large and magnificent plant IT is, apart from the greenfly of course. [A puts the chrysanthemum on the table] 'Now I want you to tell us, John, when you first took the cutting for THIS, how old IT is, and what IT's going to be.'

A: 'Yes well THIS is <one of the large ones - what we used to call the Japs or the Giant Exhibition.>'

B: '<Yes.>'

A: 'So we will end up with one bloom, as large as we can get, on the top.'

B: 'Uh-huh. A sort of big mop-head type of thing.'

A: 'Yes, <they> do call THEM mop-heads.'

The focus seems to be on whatever they call mop-heads, although it is not clear whether these are the plants or the blooms (possibly both) - perhaps an example of the role of pragmatics. Either way, the plant seems to move back into local focus:

B: '<Yes.>'

A: 'So because IT's <rather a specialised thing>, we take the cutting fairly early.'

B: 'What, January?'

A: 'Yes, or even before we go away for the christmas break.'

B: '<Oh really?>'

A: 'THE FIRST FEW go in then.'

B: '<Yes.>'

A: 'There's always some exhibitors that we sell to that want THEM.'

There might seem to be a sort of ambiguity between "the first few cuttings" and "the fist few plants" here. It does not yield any real ambiguity of sense because the representations for the plants will be the same as those for the cuttings at this point - it is the representations that get focussed, not the referring expressions.

B: 'And THESE are <the tiny cuttings that you get off the stools of the old plants>.' [points]

A: '<That's right, yes.> Yes, THEY're growing up, and you need to take ONE that's growing quite nicely, hasn't gone hard or ...'

B: 'And take THE BIGGEST ONE, I suppose.'

A: 'Yes - well - yes, THE BEST-LOOKING ONE.'

B: '<Yes.>'

A: '<It's very important> - if you don't start off with the right thing, who knows what you're going to grow into?'
B: 'Well you don't want to start off with the runt of the litter, do you?'
A: 'No, no. No, we soon get rid of those.'

In the next section focus is on either the chrysanthemum or on chrysanthemums in general. Once again, a demonstrative followed by ordinary pronouns seems to carry the focus:

B: 'So THIS then is as old as from January
A: 'Yes. Yes, IT starts very slowly ...'
B: 'And when will IT be blooming?'
A: 'This will be blooming in the first week of November - last weekend of October.'
B: '<Yes.>'
A: 'We try and time OURS. We have an open weekend last weekend of October - and we try and get THEM for then.'
B: 'Yes. Now what was the PROCEDURE? You've not kept this in the greenhouse all the time.'
A: 'No, THIS is standing outside at the moment.
B: '<Yes.>'
A: 'IT starts off in the greenhouse obviously at that time of year, grows very slowly - keep THEM as cool as possible.'
B: 'Yes. Put THEM out when? End of May?'
A: 'Yes. Yes, once the frost has gone. You put THEM in the cold frame before then.'
B: '<Yes. Yes, marvellous.>'
A: 'So - and then IT's stopped. <They>'ve pinched THE TOP out.'
B: '<Yes.>'

Rules a and f suggest "stopping" as a new focus, and then the chrysanthemum.

A: 'STOPPING worries some people, but IT shouldn't really.'
B: '<No.>'
A: 'IT's quite a simple thing. If you leave THE CHRYSANTHEMUM, IT will go on and grow normally anyway.'
B: 'and become <a spray>.'
A: '<Yes. Yes.>'
B: '<What they call a spray chrysanthemum>.'
A: 'But if you want THEM for a particular date, you often ...'

A series of utterances with only one potential focus follows:

B: 'Tell us then about THESE LITTLE CHAPS here.'
[points to three small plants off to one side]
A: '<Yes.>'
B: 'Which one do you want to talk about first?'
A: 'well, THE ONE ON THE RIGHT.'
B: 'THAT ONE.' [points]
A: 'THIS ONE.' [points and nods in agreement]
'With this hot weather we've been having .
B: 'Is THIS called <a Charm chrysanthemum>?'
A: 'Yes, THAT's <a Charm>. THEY're all three <Charms>.
And I think one of the problems people have had is
that <it>'s got hot and dry ...'
B: '<Yes.>'
A: 'And some of these go to bud too early.'
B: '<Uh-huh.>'
A: 'So you get THESE EARLY FLOWERS which -
THEY don't look very good.'
B: 'Yes, some of them there, like THAT for instance
has gone off completely, hasn't IT?'
[points to flower]
A: 'Yes. So THIS was flowering - IT started flowering
about the middle of July, which is far too early
B: 'When do you want THEM to be in full bloom?'
A: 'Again, really in October, but THEY slow down at
the end, so if THEY're coming in in September ..
B: 'I've forgotten to ask you, what do you
fertilise your chrysanthemums with?'
A: 'Um - we just use A GENERAL FERTILISER.
We sell OUR OWN, and we use THAT obviously, but

By rules a and f, local focus in B's anecdote seems to start on his grandfather, then to move onto sheep manure before returning. Later I will suggest that focus is on sheep manure throughout. Reference to chrysanthemums makes them the global focus by these rules.

B: 'MY GRANDFATHER used to collect a bit of sheep
manure and press the lumps in with his thumb.
<I can see him doing it now.>
And HE used to think that that was the most
wonderful fertiliser for /them/.'
A: 'I should think IT must have worked fairly well.'
B: 'Yes. Why SHEEP, I don't know.'
A: 'Yes, I haven't tried THAT ONE either.
<I've obviously got something to learn
when I get back, haven't I?>'
B: 'Well he insisted on having SHEEP.

The foci are not well-defined in the next section, which almost has

the character of a prolonged series of metastatements.

A: 'Yes, but ALL THESE THINGS work.
Some people feel that the more feed you put in,
the bigger and better plant you get.'
B: '<Yes.>'
A: '<But this doesn't - this isn't true.>
<It>'s far too easy to overfeed, and then you
get a lush plant that's susceptible to disease.'
B: 'And <it>'s very wrong, I understand, to feed AN
AILING PLANT. <They> say <it>'s like giving a
baby who's not very well caviar or something like that,
and therefore not doing very well.
Now what about THE DAHLIAS there?'

The dahlias are unambiguously the only possible focus. Focus then shifts between the specific plants and dahlias in general:

A: 'Um - THE DAHLIAS - yes.
THESE are again <a few specimens> that I -
normally dahlias would grow in the open ground.'
B: '<Yes, of course.>'
A: 'And THESE have been dug up recently, and THEY're
suffering a bit. But ...'
B: 'You've just brought THEM to demonstrate to us.'
A: 'Yes, yes. THEY - THEY - THESE have also had not
quite enough light. When you're growing dahlias
<it>'s important to have a good open area - get
plenty of light in.'
B: '<Yes.>'
A: 'THEY like plenty of feed .
B: '<Yes>'
A: '...DAHLIAS do, and THEY don't want to dry out
at all.'
B: '<No.>'
A: '<It>'s very important with DAHLIAS that THEY don't
dry out too much. But if you take this one here,
you can see that there are a central bud and then
a bud either side.'
B: 'Would you take THE SIDE ONES OFF off?'
A: 'Yes, I would just take THOSE off.
Some people worry about taking things off ...'
B: '<Yes, well, that's grand. Thank you very much.>'

3) Further Analysis.

Using the rules given, foci have been assigned to about three-quarters of the utterances in the text. Hence it is possible to explain how nearly all of the pronouns get interpreted. One problematic class of examples is where a plural pronoun is used to refer to a general class, as in:

A: 'Yes. So this was flowering - IT started flowering about the middle of July, which is far too early.'
B: 'When do you want THEM to be in full bloom?'

More will be said about this later. Few global foci have been assigned so far, but those that have seem to make good sense. For instance, honeydew replaces greenfly as local focus, but greenfly become the global focus, so that honeydew is seen as a sub-topic of greenfly, which is intuitively appealing. I shall now use the analysis to draw some conclusions about other focussing phenomena. Appendices 5 and 6 summarize the findings of this section, and all references to rule numbers refer to those given in Appendix 5.

a) Demonstratives and Deixis.

POCUS only deals with pronouns like "he", "she", and "it", but hopefully a full theory of focus will tell us something about all sorts of anaphora. For instance, this text contains fifty or so demonstratives, many of them anaphoric, and it would be interesting to see if they can be incorporated into the theory somehow. Of course, we can't expect to explain them all. A number are of the sort mentioned above that refer to actions, states of affairs, or parts of the discourse:

B: 'And ladybirds eat them.'
A: 'Yes. Yes, IT is strange, isn't it?
But the less that you kill ladybirds and their larvae the better, because they'll certainly eat their way through a few. So THAT's quite helpful.'

A: 'Yes, I wouldn't be very popular if I did THAT, would I?'

But if we put these aside, there are still about forty with well-defined referents, which are usually physical objects. These include noun phrases with demonstratives as determiners like "this one", and demonstrative pronouns like "this". Some are used deictically, some anaphorically, but it is not always clear which is which, and indeed there may be no clear dividing line. It may be that deixis is a separate system working in parallel with the focus system, so that some demonstratives can be interpreted correctly by both systems. For instance, deixis may be used to reinforce focus by occasionally gesturing towards the object being talked about, using some deictic expression, etc., even though it is in focus and a simple pronoun would do. When A says

A: 'Yes, I hadn't seen those.
Those are some butterfly eggs.'

he may be doing just that. If we assume both the demonstratives in this example are deictic, and that the first puts the eggs in focus by rule 17, then we would expect an ordinary pronoun in the second sentence ('They're butterfly eggs'). This would certainly do, but the additional deictic expression seems to buttress the focussing mechanism somehow. Even more common are cases where deixis is used to guide the return to a previous topic, rather than rely on a search back through previous foci or some other global process:

A: 'But there's a good one here - Malathion -
THAT's fairly safe.'
[holds up a bottle of Malathion]
|
There then follows a long discussion
of various insecticides, followed by:
|
A: 'But THAT's a good standard one.'
[Holds up the bottle.]

A: 'So because it's rather a specialised thing,
we take the cutting fairly early.'
B: 'What, January?'
|
There is a discussion of the cuttings followed by:
|
B: 'So THIS then is as old as from January.'

This use of deixis to back up focus makes sense, because it places less of a load on the hearer's memory. I shall return to deixis shortly, but first let us look at how demonstratives can be used anaphorically.

1) Anaphoric Demonstratives.

These account for a good proportion of the anaphora in the text, and most of them are pronouns, so perhaps our rules can be applied to them. Certainly it looks as if our rules for pronoun interpretation may be of use. Of the twenty-five demonstrative pronouns that seem to be anaphoric, eighteen refer to the current local focus, and five refer to items mentioned in the previous utterance when a focus has not been assigned to it. So it seems as if rules 4 and 5 apply to demonstrative pronouns too. Two examples occurr where a demonstrative pronoun refers to something in the same sentence, but both are examples of the same sort of idiom ("super-rats and things like that", "caviar or something like that") and it is difficult to think of other such examples that do not involve some sort of coordination:

> 'A Renault 5 was parked outside, and the man who drove THAT had to be a maniac, because there were dents all down the side.'

It seems that rule 6 is not used except in these special constructions, many of which would not count if we adopted a clause-based analysis as suggested earlier. The rules therefore seem to be:

When a demonstrative pronoun is encountered try:

1) The local focus.
2) The potential foci of the last utterance.

The rule for focussing looks very simple too. All but three of the demonstrative pronouns are definitely in local focus, and those three occurr in utterances where a focus has not been assigned. It seems that demonstrative pronouns are always in local focus (with one qualification to be given later). So our focussing rules become:

For the local focus, try:

a) Any item referred to by a demonstrative pronoun.
b) Any item referred to by rule 4.
c) Any other item referred to anaphorically.
d) Any other potential focus.

The main use to which to which these new rules are put seems to be to reinforce the focus when some new item has just been focussed. Typically, the new focus is mentioned as a full NP or deictic expression, then referred to by a demonstrative pronoun to reaffirm its new status as local focus, after which simple pronouns are used. For example:

B: 'Where does THE STICKINESS come from?'
A: 'THAT's the honeydew that they excrete.'
B: 'I see.'
A: 'And IT just collects over the plant.'

A: 'We sell OUR OWN, and we use THAT obviously.

This strategy seems particularly useful for moving a series of items rapidly through focus:

A: 'Yes, Pyrethrum. Yes, Derris as well, THAT's er - and we use nicosoap - nicotine.'
B: 'Oh, do you? Yes. Yes.'
A: 'We find THAT's quite good.'
B: 'and Quosia? Do you use THAT?'

So it seems that using a demonstrative pronoun can tell the hearer that a shift of local focus has recently occurred. Not all examples fall into this category - some may be deictic or contrastive, as will discuss below - but a rough rule would be:

When a demonstrative pronoun is used anaphorically, local focus may have recently shifted to an item just mentioned.

The other class of anaphoric demonstratives to be considered are those consisting of an NP with a demonstrative as a determiner,

like "this one". I will discuss their interpretation later on, but in the meantime note that the few examples found in this text seem to be in focus:

B: 'and Quosia? Do you use that?'
A: 'No, we haven't got round to ...'
B: 'No.'
A: 'I'll have to try THAT ONE.'

A: 'I should think it must have worked fairly well.'
B: 'Yes. Why sheep, I don't know.'
A: 'Yes, I haven't tried THAT ONE either.

ii) Deixis.

One of the most obvious ways in which demonstratives show their power is when they are used deictically, and because this a transcript from a T.V. programme in which the participants are often talking about physical objects visible to both them and the audience, many examples of deixis occurr. Deixis is obviously closely related to focus - after all, one of the simplest and most effective ways of getting an object into focus is to point to it. Children learn to do this at an early age, and it is reasonable to suggest (as Lyons 1981 does) that it is one of the fundamental processes on which reference is based.

As mentioned above, it is not always clear when deixis is being used. In a number of cases a possibly deictic expression would also receive the correct interpretation if treated as an anaphor, but there are still some cases where the speaker uses expressions like "there" or is gesturing nicely to indicate that deixis is at work:

A: 'There's a little LADYBIRD there look.'

B: 'What's THAT YELLOW THING there?' [points]

A: 'And there's a little BLACKFLY' [points]

A: 'But there's a good one here - MALATHION - that's fairly safe.' [holds up the bottle]

A: 'But THAT's a good standard one.' [Holds up the bottle.]

B: 'Now how - let's just lift THIS up onto the table so that we can see what a very large and magnificent plant it is, apart from the greenfly of course.

B: 'And THESE are the tiny cuttings that you get off the stools of the old plants.' [points]

B: 'So THIS then is as old as from January.'

B: 'Tell us then about THESE LITTLE CHAPS here.'
[points to three small plants off to one side]
A: 'Yes.'
B: 'Which one do you want to talk about first?'
A: 'well, THE ONE ON THE RIGHT.'
B: 'THAT ONE.' [points]
A: 'THIS ONE.' [points and nods in agreement]

A: 'So you get THESE EARLY FLOWERS which - they don't look very good.'
B: 'Yes, some of them there, like THAT for instance has gone off completely, hasn't it?' [points]

B: 'Now what about THE DAHLIAS there?'

A: 'But if you take THIS ONE here, you can see that there are a central bud and then a bud either side.'

In all these cases (apart from the last one where a focus has yet to be assigned) the item referred to is in local focus according to the initial analysis, so that it may be that deixis always focusses things. As with the demonstrative rule, I shall offer one slight qualification to this later, but for the time being the rules seem to be:

When a demonstrative pronoun is encountered try:
1) The local focus.
2) A potential local focus from the last utterance.
3) Deixis (especially if there are other clues like gestures).

For the local focus try:
14) Any item referred to deictically, or by a demonstrative.
15) Any item referred to anaphorically by rule 4.
16) Any other item referred to anaphorically.
17) Any other potential local focus.

These rules take care of those cases where a demonstrative pronoun may either be deictic or anaphoric, because as far as focussing is concerned the result is the same. The first pronoun in this example is probably deictic, and puts the plant into focus:

'Yes. Now what was the procedure? You've not kept

THIS in the greenhouse all the time.'
A: 'No, THIS is standing outside at the moment.'
B: 'Yes.'
A: 'IT starts off in the greenhouse obviously at that time of year, grows very slowly ...'

The second demonstrative may be anaphoric or deictic, either way it receives the same interpretation and remains in focus. Then, once the new focus is well-established, simple pronouns can be used. In this way deixis and focus can be seen to reinforce one another. So it seems that demonstratives and deictic expressions have the strongest focussing effect, and are particularly useful when shifts of local focus take place. I will now suggest the first qualification on these rules.

b) Stress and Contrast.

It might be thought that when a demonstrative is used to refer to the local focus an ordinary pronoun would do just as well. After all, the interpretation and focussing would be the same. Yet we find that if we replace the demonstrative by a simple pronoun in such cases the sense changes somehow:

A: 'We've got everything on this one, haven't we?'
B: 'Yes, we have, haven't we? Yes.'
A: 'Shan't have to take THIS/IT home.'

The difference lies in the way the last sentence would be stressed in the two cases. In the original, the stress was on the demonstrative, but in the new version a more natural reading would be

'Shan't have to take it HOME.'

Similarly, the stress would change in these examples:

A: 'But there's a good one here - Malathion
`THAT's fairly safe.'

or: `It's fairly SAFE.'
or: `It's FAIRLY safe.'

B: 'and Quosia? Do you use that?'

or: `Do YOU use it?'
or: `Do you USE it?'

The problem seems to be that the demonstrative can carry the stress but "it" can't. And this really does seem to be a problem that is restricted to "it", because we can devise other examples containing all of the following:

HIM
HER
THEM
'Shan't have to take THOSE home'
THIS
THAT

So it seems that the demonstrative is used instead of "it" when stress is required. Given what we now know about demonstratives, this suggests that stress interacts with focus in some way. According to some authors (e.g. Halliday 1967), pronouns are only stressed when used contrastively, but this is not quite true. For a start, a stressed pronoun may be deictic:

'Look at HIM.'
'SHE's pretty.'
'what are THOSE.'
'THAT's a nice car.'
'Where did you get THEM?' etc.

Then there are the genuinely contrastive examples. The quoted example seems to be one of these:

'Shan't have to take THIS home.

where the stress seems to say "this-as-opposed-to-any-other". Similarly there are examples where the contrast is between two specific items:

'I know about THEM.
It's the OTHERS I'm worried about.

'SHE bought the gun, but HE pulled the trigger.'

But there is also a third class of examples which, while contrastive in a sense, are somewhat different. Look at:

A: 'There's a little ladybird there look.'
B: 'Now what does SHE do?'

Speaker B puts stress on the pronoun - why? It should be noted that this is a new focus, and that if the referent were treated as neuter a demonstrative would be the natural choice:

'Now what does THAT do?'

This seems to be one of those cases where an anaphoric demonstraive signals a shift of focus. It seems that stressed pronouns in these case play the same role as demonstratives. It is even possible to construct examples where stressing the pronoun changes the interpretation and shifts the focus:

'Jacquie wrote a story. She showed it to Sally.
She wasn't sure about it.'

If the pronoun in the last sentence was stressed ('SHE wasn't sure...) we would think it referred to Sally, otherwise it might refer to Jacquie. The same effects can be observed with demonstratives:

'The white ball crossed the table.
It hit the black.
It/THAT went into the pocket.'

All this could be explained by changing the rule given above to:

11) When an anaphoric demonstrative or stressed pronoun is used, focus may have recently shifted to an item just mentioned. Otherwise a contrast may be implied.

This is not a hard and fast rule, because of the possibility of contrastive usages (indeed, this is a sort of contrast itself - between foci). No definite conclusions can be drawn from this text about the effects of stress on NPs generally, but it may be that it still has some sort of focussing effect, even when contrastive. Certainly in those cases where an item is contrastive it also seems to be in focus, and that corroborates my remarks in Part Two about the possible effects of stress in shifting focus. Perhaps all stressed NPs are focussed, in which case using a demonstrative seems to be simply another way of giving the hearer the same information that would be carried by stress.

Incidentally, just as deixis may be a system that operates in parallel with the anaphoric one, the same may be true of contrast. In this example, the demonstrative in the second sentence may mark a shift to a new focus, but it also may be contrastive, and the two are not mutually exclusive:

B: 'What's that yellow thing there?'
A: 'Yes, I hadn't seen THOSE.'
(as opposed to the others)

And in this example the pronoun may be both deictic and contrastive

A: 'Um - the dahlias - yes.
These are again a few specimens that I
normally dahlias would grow in the open ground.
B: 'Yes, of course.'
A: 'And THESE have been dug up recently,
and they're suffering a bit.'

a conclusion that gets some support from the fact that a contrastive "but" would seem at home at the beginning of the last line.

c) Global Structure.

So far we have discussed the text purely in terms of local focus, but one of the advantages of studying an extended text like this is that it gives us more scope for analysing global effects. Now that we have a reasonable idea of the local structure it is time to concentrate on these.

A rough definition of a global effect might be that whenever anaphoric expressions are used to refer to things that are not in local focus, or that have not been recently mentioned, some kind of global process is involved. So in the opening lines of the text, the local foci are as marked:

A: 'and then if you look inside the leaf there you can see perhaps a hundred GREENFLY even in that little. [points to leaf]
B: 'Where does the STICKINESS come from?'
A: 'THAT's the honeydew that they excrete.'

and so to interpret "they" in the last line, the hearer is forced to search back through the previous foci until he finds the greenfly. By the global focussing rule given above (rule 18) greenfly become the global focus here, and this accords nicely with our intuitions, as honeydew is introduced as a sub-topic of greenfly, to illustrate one of the reasons why they are undesirable. Similarly, a little later we encounter the line:

B: 'Now what does she do? She eats them, does she?'

where the ladybird seems to be the local focus, and once again the greenfly have neither been in local focus nor recently mentioned, so that they become the global focus here too, which is appealing be-

cause ladybirds are introduced as another sub-topic of greenfly, this time to illustrate how they may be controlled.

POCUS would correctly interpret both these examples (provided we could rule out an interpretation in which the ladybird eats ladybirds on pragmatic grounds) by searching back through the previous foci (the "hark-back stack"), but there is more to the global structure of a text than a list of foci, just as there is more to the text itself than a list of sentences. The foci are related to one another in ways that reflect the topic structure of the discourse, and a theory of global focus must take account of such structures, and if possible say something about how they get built and used. The rule that POCUS uses ignores a fact that is obvious when one looks at any real linguistic data: that when we do return to some previous focus we often find ourselves doing so several times in a short space of time, so that the item in question seems to play some special role as topic for that part of the discourse, even though it may not be locally focussed. So, in the part of the text where the ladybird is in focus, there are three references to greenfly, which seems to be the global focus for that section.

The pattern seems to be that once an item moves out of local focus it either ceases to be focussed altogether, or it becomes the new global focus and thereafter forms an important part of the context in which what follows is to be interpreted. This kind of idea is not new, in fact it's much like Grosz's or Reichmann's ideas about global focus. In both cases we have hierarchical structures that constrain the interpretation of what is said, an idea that is also common in computational models that use knowledge representa-

tion structures to provide contextual clues to language processing (such as Bobrow et al 1977). But what is perhaps interesting about rules 7, 8, and 18 is that, whereas Grosz, Reichmann and others have talked about the way in which pragmatic and rhetorical strategies drive the global system, it is local focus that does the work in this case. Items move into global focus from local focus, and it is the failures of the local focus system to deal with anaphora that drive the global system to construct links between different parts of the text.

Looking again at the first few lines of the text, we see the following succession of foci in the final analysis:

LOCAL	GLOBAL
Greenfly	?
Honeydew	Greenfly
Fungus	Honeydew
Ladybirds	?
Ladybird	Greenfly
Greenfly	?
Ladybirds and Larvae	Greenfly

(some of these assignments require further explnanation, to be given in subsequent sections). So if we say that the global focus is a topic, and the local focus one its sub-topics, then we can represent the topic structure by a sort of tree diagram, thus:

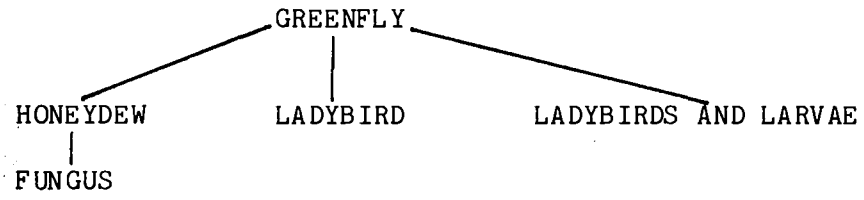


Fig. 3 Topic Structure.

This sort of structure could then become the basis of a knowledge

representation structure - a sort of conceptual dependency diagram that would hopefully tell us something about how people remember and use the information in a text.

d) Connectives and Shifts of Focus.

Reichmann (1978) mentions that certain "clue words" seem to trigger shifts to new topics or returns to old ones, and several examples of this sort of thing are seen in the text. All the following utterances mark shifts of either local focus, global focus, or both:

B: 'Yes. Now what's the cure then?' [new local focus]

B: 'Now I want you to tell us, John, when you first took the cutting for this, how old it is, and what it's going to be.' [new local focus]

B: 'Now what was the procedure?' [new local focus]

B: 'Now what about the dahlias there?' [new local focus]

A: 'But the less that you kill ladybirds and their larvae the better, because they'll certainly eat their way through a few.' [return to previous foci]

A: 'But - um - blackfly and greenfly are the two commonest problems I think that you get.' [new local focus]

A: 'But there's a good one here - Malathion - that's fairly safe.' [new foci]

A: 'But it's also important to change them round.' [return to previous focus]

A: 'But that's a good standard one.' [return to previous focus]

A: But if you take this one here, you can see that there are a central bud and then a bud either side.' [new foci]

A: 'So we will end up with one bloom, as large as we can get, on the top.' [new local focus]

A: 'So because it's rather a specialised thing, we take the cutting fairly early.' [return to previous focus]

B: 'So this then is as old as from January.' [return to previous focus]

A: 'So - and then it's stopped.'
[return to previous focus]

A: 'So this was flowering - it started flowering about the middle of July, which is far too early.'
[return to previous focus]

Roughly, it seems that the connectives "now", "then", "so", and "but" can all be used to mark shifts of focus (apart from their other uses), with "now" apparently reserved for the introduction of completely new local foci. A typical strategy is to introduce a new topic with "now", digress into a sequence of sub-topics (perhaps with "so" or "but"), and then return to it with "so" or "but". For example:

B: 'Yes. NOW what's the cure then?'
A: 'Any of the insecticides that are recommended at your local garden centre.'
B: 'Yes.'
A: 'BUT there's a good one here - Malathion - that's fairly safe.'
|
[There is then a long discussion of various insecticides]
|
A: 'But it's also important to change them round.'

It should be noted in passing that these connectives do not just prompt the shifts on their own. The interpretation would generally be the same without them, because the rules for pronoun interpretation and deixis produce shifts of focus on their own, but the connectives do give the hearer additional warning, and so perhaps speed up things by allowing him to skip possible interpretations that would not involve any shift of focus.

It should also be noted that some of these shifts occurr in Wh-questions, which themselves may have a focussing effect. A question like

'Now what's the cure then?'

seems like a good way of introducing a new focus, particularly in an interview-style discussion like this. Such questions typically receive elliptical answers consisting of a single noun phrase, such as

'Any of the insecticides that are recommended
at your local garden centre.'

and in the case of "what" and "which" questions this tends to be the new focus, although not for "when" and "where" questions, where as we will see later the focus tends to be contained in the missing fragment of the response. To understand why, we must look at the role that locatives and time phrases play in local and global focus, the subject of the next section.

e) Prepositional Phrases.

This text contains more than thirty examples of prepositional phrases, many of them in sentences where we can now be fairly sure what the foci are, so it now makes sense to see how they affect the focussing system, if at all. For instance, in a sentence like 'I found a box in the shed.' we might wonder if the shed can be the focus. As far as local focus is concerned, the answer seems to be "no". If we follow such a sentence with 'It's made of wood.', then the natural interpretation is that the box is made of wood, and not the shed, suggesting that the box is more likely to be the focus of the first sentence. And in this text virtually no examples occurr of items inside PPs being local foci. There are some ambiguous cases, though. In

A: 'We've got everything on this one, haven't we?'

we might think that our demonstrative rules put "this one" into focus, and the same is true of

A: 'and then if you look inside the leaf there you can see perhaps a hundred greenfly even in that little ... ' [points to leaf]

and

A: 'And you get fungus growing on that as well.'

But in both these examples there are good reasons for choosing another item as the focus. First, the greenfly are later referred to in a way that suggests they must have been in focus at this point - because the "hark-back" rules only deal with previous foci. Second, the fungus is immediately afterwards referred to by a demonstrative that suggests it has just moved into focus by rule 11:

A: 'And some of the trees you see THAT'

If the honeydew referred to inside the PP had really been in local focus, then our pronoun interpretation rules predict that this sentence would mean that you see honeydew on the trees, not fungus. So it looks as if

> Items referred to within prepositional phrases are not potential local foci.

This is the qualification on our demonstrative rules hinted at earlier, that demonstratives do not become local foci when they occurr inside PPs. That means that in the cited example

> 'we've got everything on this one, haven't we?'

the local focus is "everything", and there is one slender piece of evidence for this claim: the NP is stressed, whereas "this one" isn't.

The new rule also explains why "when" and "where" questions don't seem to have the same focussing effect as "what" and "which" questions: the answers to such questions tend to be PPs. As we shall see in the next section, this forces the hearer to look for the local foci in the missing fragment of the answer when it is elliptical. And it may be that the rule extends to all locatives and time phrases, although there is insufficient evidence in this text to be certain.

However, I do believe that prepositions can have a focussing effect, although it is global, not local. After all, there is no reason why something should be barred from global focus just because

it can't be the local focus. In fact, our rules seem to suggest it. In the fungus example,

A: 'And you get FUNGUS growing on that as well.
And some of the trees you see THAT -
a black FUNGUS growing on it.'

the local foci are now as marked, with fungus being introduced by a full NP followed by a demonstrative pronoun in the stereotyped manner used to mark a shift of focus. To find referent for "it", though, we are forced back to the previous local focus, and rule 18 says that that makes honeydew the global focus, as indicated earlier when discussing figure 3. It seems, then, that an anaphor inside a PP can refer to the global focus. Indeed, it should be noted that the two references to honeydew in the above example follow a pattern we have seen elsewhere: a demonstrative pronoun marks a shift of focus, after which reference is made using simple pronouns - only this time the shift is a global one. This suggest that rule 11 extends to global foci as well.

It may be that this strategy of using a PP containing an anaphor to set up the global focus is quite common, but unfortunately there are no other really convincing examples in this text. So let us at this point consider another text, this time an extract from a short story called "Sredni Vashtar" by Saki:

> 'In the dull and cheerless garden, overlooked by so many windows that were ready to open with a message not to do this or that, or a reminder that medicines were due, HE found little attraction. The few FRUIT TREES that it contained were set jealously apart from his plucking, as if THEY were rare specimens of their kind blooming in an arid waste; it would probably have been difficult to find a market gardener who would have offered ten shillings for their yearly produce. In a forgotten corner, however, almost hidden behind a dismal shrubbery, was a disused TOOL-SHED of respectable proportions, and within its walls CONRADIN found

> a haven, something that took on the varying aspects of a playroom and a cathedral. HE had peopled it with a legion of familiar phantoms evoked partly from history and partly from his own brain, but it also boasted two inmates of flesh and blood. In one corner lived a ragged-plumaged HOUDAN HEN, on which the boy lavished an affection that scarcely knew another outlet. Further back in the gloom stood a large HUTCH ...'

Local foci have been marked on the basis of the rule given so far. Initially, it seems local focus is on Conradin. The garden is not a potential local focus, because it appears in a PP, but two subsequent references to it, one by a pronoun, the other by the elliptical "In a forgotten corner". If we assume that our pronoun interpretation rule are correct, and that they perhaps extend to ellipsis (both big assumptions, of course) then the garden must be the global focus of the first sentence, because it's certainly not the local focus, a potential local focus, or a previous focus (this is the first mention), and our rules won't let us refer back to anything else. This assumption is also borne out by the long gap between the second and third references to the garden, typical of a global focus.

The same trick, of using a fronted PP containing an anaphor ("within its walls"), seems to put the shed into global focus too. Immediately after this, the reintroduction and subsequent pronominalisation of Conradin suggests that he returns to local focus, leaving the shed as global focus by rule 7. Then once again we have an elliptical PP ("In one corner") that would make the shed the global focus if ellipsis behaves at all like pronoun interpretation (as I shall suggest it does below). This even suggests that in the sentence 'but it also boasted two inmates of flesh and blood' it may be the inmates that are in local focus, since the shed seems to be in

global focus before and after. Finally the phrase "further back in the gloom" might be read as meaning something like "further back from the hen in the gloom of the shed", so that the definite NP "the gloom" is seen as containing a sort of elliptical PP that gets interpreted with respect to the global focus.

This is straying into the sort of territory Sandford and Garrod (1981) were exploring when they suggested that having "the house" in explicit focus puts things like "the door", "the windows", etc. into implicit focus. In our terms, noun phrases like "the door" are interpreted as being related to the local or global focus, and can put it into global focus, as in this example:

> 'The house was old and draughty. The windows needed mending, the door had no handle, and the roof leaked. But it was still home.'

Here local focus starts on the house, but moves in turn to the windows, the door, and the roof. Because these NPs are definite, the hearer expects them to be "anaphoric" in some sense, but none of them are previously mentioned items. So he concludes that the windows relate to some previous focus in some way, looks at the previous local focus (the house) and finds that in his representation for it he has included the possibility that it may have windows. This puts the house into global focus - windows are seen as a sub-topic of the house. Then the same process is used to interpret "the door" and "the roof" with respect to this global focus. Finally, "but" signals a shift back to the original focus. This is the sort of thing that Grosz (1981) deals with in her global focus analysis of task-oriented dialogues, and some possible examples crop up in our T.V. programme, but not enough to draw any interesting conclusions

from. Suffice it to say that, we may have a rule says:

20) When a definite NP refers to an "implied" entity, like "the door", the "implier" (e.g. "the house") is probably one of the foci, and moves into global focus.

In this light, the foci in the Saki extract seem to be:

LOCAL	GLOBAL
Conradin	Garden
Trees	Garden
Trees?	Garden
Shed	Garden
Conradin	Shed
Conradin	Shed
Inmates?	Shed
Hen	Shed
Hutch	Shed

and so, once again we can produce a tree diagram of the topic structure that is intuitively appealing:

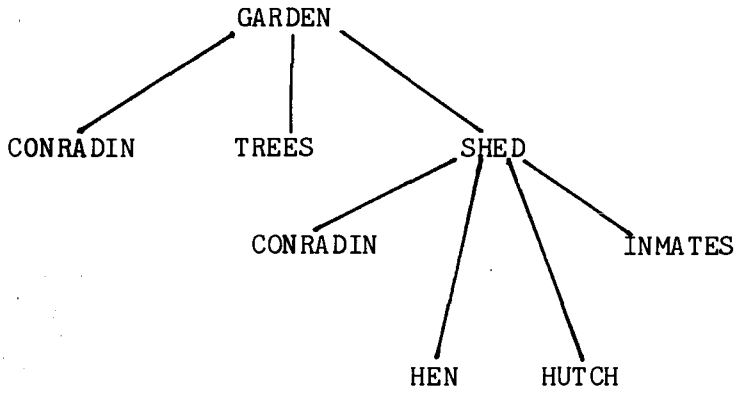


Fig. 4 Topic Structure.

This analysis of the structure is backed up by one or two other pointers in the text. The trees are only mentioned in order to contrast them with the shed as two aspects of Conradin's attitude to the garden (the contrast being emphasised by "however"), and as such the garden does seem a natural topic for the early part of the paragraph. Similarly, the point of focussing on the hen and the hutch is

to explain what Conradin found so interesting in the shed, and in particular to explain who the "two inmates" are, so the shed makes a reasonable topic for the second part.

So a focussed-based analysis gives us some insight into the way Saki leads us up the garden path, as it were, gradually focussing in from the garden, to the shed, to the hutch, and finally to the creature that lives inside. This long drawn-out process tantalises the reader (What does he mean when he says the shed is like a cathedral? Who are the two inmates? What lives in the hutch?) so that when it is finished the final focus seems particularly significant, and indeed it is, being the pet ferret that gives the story its title. All this also illustrates how much tighter the global structure of a written text tends to be. Local focus seems a less useful tool for analysis in such texts, and it may be that this is a consequence of the amount of advanced planning that goes on, as hinted at in the introduction to Part Three.

So to return to the original topic, it looks as if another plausible rule may be:

19) If a previous local focus is referred to in a PP,
    it becomes the global focus.

a rule that may be particularly important for fronted PPs like those in the Saki extract.

f) Ellipsis.

As explained in Part One, the whole point of focus is that it allows us to semantically underspecify things that are currently easy to access, which speeds up communication. Now the most extreme form of underspecification is to leave something out altogether - ellipsis. I have already suggested that this may be related to focus, so let us now look at the phenomenon in detail.

There are a few examples in this text where it certainly looks as if the local focus is referred to elliptically. Consider:

B: 'Yes. Put THEM out when? * End of May?'
A: 'Yes. Yes, * once the frost has gone. You put THEM in the cold frame before then.'

B: 'When do you want THEM to be in full bloom?'
A: 'Again, really * in October, but THEY slow down at the end, so if THEY're coming in in September...'

The local foci have been marked for some of the sentences, but in the others we have elliptical answers to "when" questions, consisting of time phrases. The missing fragments are marked "*". Now the rule presented in the last section suggests that such phrases cannot be or contain the local focus, so perhaps it is contained in the missing part. If we reconstruct the full sentences, we find that each contains a "hidden pronoun" that could carry the focus (e.g. 'You really want THEM to be in full bloom in October'). And it would certainly be appealing to be able to say that this was the case, because then the local foci would be the same throughout each of the above examples, thus explaining the pronominal references to them after the elliptical sections ("but they slow down" and "you put them in the cold-frame").

Another obvious class of examples is where a series of clauses has the same subject, as in:

A: 'It's quite a simple thing. If you leave THE CHRYSANTHEMUM, IT will go on and * grow normally anyway.'
B: 'and * become a spray.'
A: 'Yes. Yes.'
B: '* What they call a spray chrysanthemum.'

Once again ellipsis has been marked where it occurrs, and once again the local focus must be inside the missing fragment, because NPs like "a spray" and "what they call a spray chrysanthemum" are attributive and therefore not potential foci. So it seems that when an utterance consists of a fragment that is an elliptical version of a full clause, then the missing part may contain something that refers to the local focus.

Not all elliptical phrases can carry local focus, of course. We would not expect sentences with elliptical PPs like "John walked in" to contain the local focus in the PP, because PPs can't carry local focus, although we have already seen from the Saki piece that they may carry the global focus:

'In a forgotten corner *, however,

'In one corner *, lived...'

In the next section I will suggest that elliptical quantifier phrases and possessives like "a few" and "our own" may have a similar effect. Furthermore I will suggest that ellipsis is interpreted by the focussing system in much the same way as a pronoun is. Brown and Yule (1983) describe ellipsis as a sort of "null anaphor", and in the case of phrases with an elliptical NP, a good strategy for interpreting it seems to be to treat it like a pronoun, and to try

substituting either the local or global focus. As we will see, the situation is slightly more complicated than this in the case of things like "a few", where the missing NP does not actually refer, and neither does it not seem to be the best strategy for interpreting elliptical answers to questions, which seem to use the original question as a pattern for constructing the full answer (see Bobrow et al 1977, for example). But for clauses where the subject is elliptical, or elliptical PPs, it does seem to be an efficient strategy - the missing NP does seem to refer to either the local or the global focus. Once again, however, there are too few examples in this text to really justify the hypothesis, and more research is needed.

g) The Particular and the General.

In her 1981 article, Sidner points out that a theory of anaphor must explain those cases where there is a shift from talking about some specific item to talking about the whole class of such items, or some prototypical member of it, as in

`I bought a Renault. They're nice little cars.'

All the sentences on the left below could be followed by all those on the right, producing a range of such shifts:

I bought a Renault.	I saw one yesterday.
I want a Renault.	I liked yours so much.
I like the Renault 5.	They're nice little cars.
I like Renaults.	I've seen several today.

Some of these sentences introduce specific items (`I bought a Renault'), some prototypical ones (`I want a Renault'), some sets (`I've seen several recently'), while others talk about whole classes (`They're nice little cars'). A number of examples of this sort of thing crop up in the text, which constantly shifts from examining particular plants and their problems, to making general points about all such plants, and back to the original examples again, and this is typical of a discussion like this that uses specific examples to illustrate general points. In this extract

A: 'THIS will be blooming in the first week of November - last weekend of October.'
B: 'Yes.'
A: 'We try and time OURS. We have an open weekend last weekend of October - and we try and get THEM for then.'
B: 'Yes. Now what was the procedure? You've not kept THIS in the greenhouse all the time.'
A: 'No, THIS is standing outside at the moment.'
B: 'Yes.'
A: 'IT starts off in the greenhouse obviously at that

time of year, grows very slowly - keep THEM as cool as possible.'
B: 'Yes. Put THEM out when? End of May?'
A: 'Yes. Yes, once the frost has gone. You put THEM in the cold frame before then.'
B: 'Yes. Yes, marvellous.'
A: 'So - and then IT's stopped.'

the local foci have been marked, and we see how focus moves from one particular plant, to those grown at B's nursery, then back to the original plant, then on to Chrysanthemums in general, and finally back to the original one again.

The main strategy used to get from the particular to the general in this text is to use a plural pronoun. In most of the examples the specific and general items are in local focus, before and after the shift respectively. For example:

A: 'IT starts off in the greenhouse obviously at that time of year, grows very slowly - keep THEM as cool as possible.'

A: 'If you leave THE CHRYSANTHEMUM, IT will go on and grow normally anyway ...
But if you want THEM for a particular date...'

A: 'IT started flowering about the middle of July...
B: 'When do you want THEM to be in full bloom?'

So the rule seems to be:

9) When a plural pronoun is encountered, try generalisingfrom the local focus. The general class becomes the new local focus.

This process is often accompanied by other clues, such as changes in mood or tense (the imperative in the quoted extract may be a clue: `keep them as cool as possible'), quantifiers ("all of them"), pragmatic effects, etc. It should be noted that this process has its parallel in the "reference to sets" strategy used by POCUS, although that was for shifts from individual foci to focussed sets of indivi-

duals, and there too there were clues to guide the hearer ("they both went in", "they all went in",etc.). The nearest we find to an example of that sort of thing is when A says

A: `That's a charm. They're all three charms.'

but in that case he is in fact referring back to a previous local focus, one which, as we will shortly see, seems to be in global focus at this point, and it is quite possible that all such shifts are global in nature.

Shifts from the general to the particular follow a different strategy. Although simple plural pronouns are acceptable for making generalizations by the above rule, singular pronouns don't seem to be used to focus in on specific examples to illustrate general points, unless this marks a return to a previous focus, in which case the shift is often marked by deixis, and perhaps a connective like "so" or "but":

A: 'IT starts off in the greenhouse obviously
at that time of year, grows very slowly -
keep THEM as cool as possible.'
B: 'Yes. Put THEM out when? End of May?'
A: 'Yes. Yes, once the frost has gone.
You put THEM in the cold frame before then.'
B: 'Yes. Yes, marvellous.'
A: 'SO - and then IT's stopped.'

But a common way of shifting from the general the particular is to use one of these sorts of expressions:

`I like YOURS'
`I saw SOME yesterday'
`What about THIS ONE'

In these examples, the NP used is either elliptical or contains the vague pronoun "one" (or "ones"). Such NPs can also be used to intro-

duce prototypical items or sets of such items. For example:

A: `Because they'll certainly eat
their way through A FEW.'

A: `They're growing up, and you need to take
ONE that's growing quite nicely.'

The pronoun or elliptical noun in such expressions does not refer back to any previous item exactly, but is nevertheless related to something mentioned beforehand - when I say "I bought a Renault 5. I liked yours so much", the Renault I bought is not a potential local focus in the second sentence, but the phrase "yours" obviously depends on it in some way for its interpretation. What seems to happen is that the pronoun or elliptical NP is treated rather like an ordinary anaphoric pronoun, except that once the antecedent has been found, instead of letting the pronoun refer to it, it is instead linked to the general class of such items. In other words, the search procedure is the same as for ordinary pronouns, but instead of coming back with a copy of the representation for a specific item, it returns a "blank frame", ready to be filled in with information that will define a new specific item, a prototype, a set of items, or whatever. Lets take a concrete example. The first sentence here puts a specific plant into focus:

A: 'THIS will be blooming in the first week
of November - last weekend of October.'
B: 'Yes.'
A: 'We try and time OURS.

When this happens, a representation of the plant is created and filed in active storage, in the area reserved for local foci. Then, when the elliptical NP "ours" is encountered, a search is initiated, much like that initiated by an ordinary pronoun. The first item to

be tried is the local focus, so the representation of the plant is retrieved. If the search had been initiated by an ordinary pronoun, this representation would then be used to stand for the referent, updated with whatever new information is contained in the sentence, and filed back in active storage. But because we have an elliptical possessive, what gets used is a copy of the "chrysanthemum" frame used to create the representation originally. Such a frame is a representation structure containing a lot of information about prototypical chrysanthemums, but as yet representing no particular plant. However, when the information contained in the new sentence (that a set of items belonging to some group of people of which A is a member is being talked about) is added to the frame, a representation of the new referent ("our chrysanthemums") is created, and this then gets filed in active storage as the new local focus, with the old frame being filed elsewhere. This sort of copying procedure is discussed by Minsky (1975) and Bobrow and Winograd (1977).

In most cases the antecedent seems to be the local focus:

B: 'Now what's the Cure then?'
A: 'Any of the INSECTICIDES that are recommended at your local garden centre.'
B: 'Yes.'
A: 'But there's A GOOD ONE here...'

But not all the examples found conform to this pattern. If we continue the above example, we find several such expressions harking back to a previous focus:

B: 'Yes. Now what's the cure then?'
A: 'Any of the INSECTICIDES that are recommended at your local garden centre.'
B: 'Yes.'
A: 'But there's A GOOD ONE here - Malathion - that's fairly safe.'

[holds up a bottle of Malathion]
B: 'Yes.'
A: 'I don't think you want ANYTHING TOO POWERFUL.'
B: 'No, I like THE PYRETHRUM ONES for that reason.'
A: 'Yes, Pyrethrum. Yes, Derris as well, that's er - and we use nicosoap - nicotine.'
B: 'Oh, do you? Yes. Yes.'
A: 'We find that's quite good.'
B: 'and Quosia? Do you use that?'
A: 'No, we haven't got round to ...'
B: 'No.'
A: 'I'll have to try THAT ONE.'
B: 'Yes, Quosia's rather A GOOD ONE.'
A: 'But it's also important to change them round.'

The phrases marked all seem to point back to "insecticides", a previous local focus. And, just as when ordinary pronouns refer back to such an item several times in a short space, so in this case "insecticides" seems to be the overall topic, indeed the last line marks a return to insecticides in general as the local focus. So it looks as if insecticides is in global focus in this extract. There is thus a parallel with anaphoric PPs: both sorts of phrases put the antecedent item into global focus, but not into local focus. So the rule seem to be:

10) When an elliptical possessive or quantifier, or a phrase containing the vague pronoun "one(s)", is found, try the following as prototypes for the new referent:

a) The global focus.
b) The local focus.
c) A previous focus.

And put the prototype into global focus.

Once again, this helps us to explain the topic structure of the extract. Initially the focus is on insecticides in general, but then we move to a discussion of specific examples, with "insecticides" as the global focus and controlling topic, after which we move back to the general discussion with a "so" shift. A further advantage of

this approach is that it may help us to understand how vague lexical items like "anything" in "anything too powerful" get interpreted. If such items are treated in the same way as pronouns like "one", then they may be correctly interpreted - "anything" just stands as an instruction to take the global focus as prototype, and the phrase is interpreted as meaning "any insecticide that is too powerful". Similarly, in this extract the use of "one" sets up the cuttings as global foci whilst discussing cuttings in general, and when we encounter the phrase "the right thing" it gets interpreted by taking them as prototype:

B: 'And THESE are the tiny cuttings that you get off the stools of the old plants.' [points]
A: 'That's right, yes. Yes, THEY're growing up, and you need to take ONE that's growing quite nicely, hasn't gone hard or
B: 'And take the biggest ONE, I suppose.'
A: 'Yes - well - yes, the best looking ONE.'
B: 'Yes.'
A: 'It's very important - if you don't start off with THE RIGHT THING, who knows what you're going to grow into?'

This is a good illustration of the way that focus constrains our search for an interpretation, and so speeds things up - although the phrase is vague, it must be related to the foci somehow, it must be relevant. These ideas are not new, of course - Reichmann discusses shifts between particular and general herself - but what is perhaps interesting about this approach is that some insight is gained into the possible communicative function of such shifts, and into the cognitive processes that may be involved - in particular the role that local focus plays, as well as global focus.

h) Sentences or Clauses?

Most of the analysis presented so far has been based on whole sentences, although in this text there are a good number of utterances that do not fall into this category. We have already seen that some sorts of coordination produce shifts of focus, and if we look at the text we see a few cases where coordinated clauses are separated by an interjection from the other speaker, such as an acknowledgement. Now if we take these as signs that the utterance has been processed and understood, it may be that foci are assigned to clauses rather than sentences. Some linguists have suggested the clause as the basic unit of language processing (e.g. Halliday 1967)), and it might be interesting to see how such an approach would affect focus theory. In this light, for example, the sort of contrast discussed earlier between two stressed NPs can be seen as a contrast between two local foci:

`SHE bought the gun, but HE pulled the trigger.'

where stress puts the contrasted items into focus and "but" signals a shift from one to the other. Similarly, it might be possible to say something interesting about "If ... then" sentences and other sorts of coordination. One example from the text is suggestive:

A: `But if you take this one here, you can see
there are a central bud and then bud either side.'

The phrase "this one", which is deictic, focusses us in on a particular dahlia, whilst putting dahlias in general into global focus by the process described in the previous section, and this shift from general to specific is marked by "but". Now if we assign foci to

clauses, this admits the possibility that the buds move into focus next, which is backed up by the next line:

B: `would you take the side ones off?'

It may be that rhetorical structures like "If ... then", "not only ... but also", etc. are designed to forge conceptual links between different foci, so that maybe what happens in the above example is that the dahlia moves into global focus when the buds move into local focus, producing a topic structure like this:

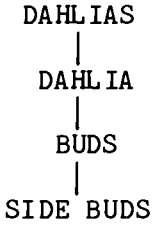


Fig. 5 Topic Structure.

(the buds move into global focus because of the phrase "side ONES", as described in the last section). This sort of approach may also have something illuminating to say about sentences where a clause or phrase is fronted to "set the scene" for the main clause, such as:

`Because he hated work, he decided ...'
`Having finished the job, he set off...'
`In one corner lived a Houdan hen...'
`Afterwards he realised that ...'

We have already seen fronted PPs play a role in setting up global foci in the Saki extract, and it may be that a similar analysis can be found for other examples of this class, although I suspect that this is a more literary gambit, and not typical of a text like the one under consideration.

Another advantage of an analysis in which focussing takes

place within the sentence is that it may help us to understand subordination. Perhaps subordinate clauses have foci too, with the focus of the main clause bearing the same relationship to that of the subordinate clause as global to local. There isn't much subordination in this extract, but one example suggests a possible pattern:

A: `Thats the honeydew that they excrete.'

where the honeydew is the local focus and the greenfly the global focus, and it may be that this sort of thing is generally used to set up global foci. That would perhaps explain why Grosz's example is so problematic

Carl thinks he is studying for his exams.

because Carl bears a special relationship to the subordinate clause, perhaps like that of a global focus. It is certainly very common in fiction for a lot of sentences beginning "He thought...", "He felt ...", etc. to refer back globally to some character not in local focus, a character who is priveleged in that the author tells the story from his or her point of view. For example, Conradin plays such a role in "Sredni Vashtar", and there are numerous sentences in the story that are prefaced with phrases like "In his eyes, ...", or "To Conradin, ...". Such references seem to play much the same sort of role as the pronouns "I" and "you" do in the gardening text - they set up a point of view. This may in a sense be another aspect of global focus, although a proper discussion of point of view phenomena is outside the scope of this discussion. Suffice it to say that a clause-based analysis, coupled with more research into the

effects of coordination and subordination on focus, might prove interesting, particularly in written texts, where the clause structure is tighter and more complex.

5) Conclusion

I have painted a picture of local focus and global focus working together to interpret anaphora, structure topics, and ultimately help produce representations for the content of the text, with systems like deixis, contrast, and viewpoint working in parallel and interacting with them. The rules suggested are, of course, just a start. A full theory of focus must look at all aspects of thematic structure, particularly at subordination, coordination, topicalization, and the roles of subject and object (or agent and patient). Global focus is bound to be affected by changes of tense, aspect, and mood, and probably by various sorts of modality and point-of-view phenomena. Sandford and Garrod (1981), Grosz (1981), and others have looked at the way pragmatic effects guide global focus, and I have touched on this very briefly above. Reichmann (1978) concentrates more on rhetorical strategies, and written texts in particular look to be a fruitful source of such data, while speech provides all sorts of prosodic clues which may have a focussing effect (c.f. my remarks on stress and various theories about intonation, e.g Brown 1977 on "paratones"). Eventually it would be nice to have a proper procedural theory in which the items that move in and out of focus are complex representation structures carrying lots of information about time, place, viewpoint, speaker, modality, and topic structure - structures that would provide a full context for discourse interpretation, a basis for pragmatic inferences, and an explanation of the way we remember and use the information in a text. Meanwhile, I hope I have shown that the concept of focus is a useful tool for discourse analysis, and sheds a little light on the

notion of "relevance".

REFERENCES

de Beaugrande, R. and Dressler, W. (1981)
"Introduction to Text Linguistics", Longmann.

Bobrow, D.G., Kaplan, R.M., Kay, M., Norman, D.A., Thompson, H.S., and Winograd, T. (1977)
"GUS-1, A Frame-Based Dialog System",
Artificial Intelligence 8(1).

Bobrow, D.G. and Winograd, T. (1977)
"An overview of KRL, a knowledge representation language",
Cognitive Science Vol.1 No. 1

Brown, G. (1977) "Listening to spoken English", Longman.

Brown, G. and Yule, G. (1983)
"Discourse Analysis", CUP.

Chafe, W. (1972) "Discourse Structure and Human Knowledge"
[in (eds.) J.B.Carroll and R.O.Freedle
"Language Comprehension and the Acquisition of Knowledge", V.H.Winston and sons.]

Chafe, W. (1975) "Givenness, Contrastiveness, Definiteness, Subjects, Topics, and Point of View"
[in C.N.Li (ed.) "Subject and Topic", Academic Press].

Charniak, E. (1972) "Towards a Model of Children's Story Comprehension", MIT AI Lab.

Chomsky, N.C. (1981) "Lectures on Government and Binding", Foris.

Clark, H.H. and Marshall, C.R. (1981)
"Definite Reference and Mutual Knowledge"
in Joshi, Webber and Sag 1981

Clark, H.H. and Haviland, S.E. (1977)
"Comprehension and the Given-New Contract",
in (ed.) R.O.Freedle "Discourse Production and Comprehension", Ablex.

Dresher, E. and Hornstein, N. (1976)
"On some supposed contributions of AI to the scientific study of language", Cognition 4 .

Dresher, E. and Hornstein, N. (1977)
"Reply to Winograd", Cogntion 5.

Grice, H.P. (1975) "Logic and Conversation"
in (eds.) P.Cole and J.Morgan
"Syntax and Semantics 3: Speech Acts",
Academic Press.

Grice, H.P. (1978) "Further Notes on Logic and Conversation" [in P.Cole (ed) "Syntax and Semantics 9: Pragmatics", Academic Press.]

Grimes, J.E. (1975) "The Thread of Discourse", Mouton.

Grosz, B.J. (1981) "Focusing and Description in Natural Language Dialogues" [in Joshi, Webber, and Sag.]

Grosz, B.J., Joshi, A.K., and Weinstein, S. (1983) "Providing a Unified Account of Definite Noun Phrases in Discourse", ACL proceedings 1983.

Halliday, M.A.K. (1967) "Notes on Transitivity and Theme in English (Part 2)" Journal of Linguistics.

Isard, S.D., and Davies, D.J.M. (1972) "Utterances as Programs" [in D.Michie "Machine Intelligence 7" Edinburgh University Press].

Johnson-Laird, P.N. (1977) "Procedural Semantics", Cognition Vol 5.

Johnson-Laird, P.N. (1983) "Mental Models".

Joshi, A.K., Webber, B.L., and Sag, I.A. (eds), (1981) "Elements of Discourse Processes", C.U.P.

Joshi, A.K. and Weinstein, S. (1981) "Control of Inference: Role of some Aspects of Discourse Structure - Centering", IJCAI-81 Vol. 1.

Masterman, M.G. (1970) "The Nature of a Paradigm" in I. Lakatos and A. Musgrave (eds) "Criticism and the Growth of Knowledge", Cambridge.

McDermott, D. (1976) "Artificial Intelligence meets Natural Stupidity", SIGART Newsletter No. 57.

Minsky, M. (1975) "A Framework for Representation Knowledge" in (ed.) P.H.Winston "The Psychology of Computer Vision", McGraw-Hill.

Prince, E.F. (1981) "Toward a taxonomy of Given-New information" in (ed.) P.Cole "Radical Pragmatics", Academic Press.

Reichmann, R. (1978) "Conversational Coherency", Cog. Sci., Vol. 2 1978.

Sanford, A.J. and Garrod, S.C. (1981) "Understanding Written Language", Wiley.

Schank, R.C. and Abelson, R. (1977) "Scripts, Plans, Goals, and Understanding", Lawrence Erlbaum.

Sidner, C.L. (1981) "Focusing for Interpretation of Pronouns", American Journal of Computational Linguistics, Vol. 7 No. 4.

Shadbolt, N. and Reichgelt, H. (Forthcoming) "Reference in Discourse", Journal of Pragmatics.

Winograd, T. (1972) "Understanding Natural Language", Academic Press.

Winograd, T. (1976) "On some contested suppositions about the scientific study of language", Cognition 5.

Woods, W.A. (1981) "Procedural Semantics as a Theory of Meaning" [in Joshi, Webber and Sag].

Appendix 1 POCUS

The following program was written in Franz Lisp (Opus 38.22), and was run under UNIX on a VAX 11/750. This appendix includes the main routines for POCUS, as well as those used to interface with MCHART, which is not included. MCHART is called by the command "parse", and when it has finished it calls the interface package via the signal "parsecomplete".

```
; "stories" is the TOP-LEVEL command. It allows you to choose whether or not
; to display the following diagnostics:

; in interface:  the parse tree in a nice format,
;                the NPs (including pronouns) of the sentence in tree form,

; in pocus:      NPs in lex/feature form (otherwise just words),
;                pronouns in tree form (otherwise just words)
;                referents in lex/feature form,
;                centre in lex/feature form (otherwise just words),
;                fcentres in lex/feature form (otherwise just words),
;                the parse tree unformatted.

; It then calls "story" which is the main loop. When one "story" is finished
; (by typing "the end") control passes back to "stories", which asks you if you
; want to tell it another story or not. Unless the answer is "no", it goes
; round again.
(def stories (lambda () (prog (interswitch pocusswitch)
(patom 'diagnostics-)(print 'interface?)(terpr)
(cond ((eq (read) 'y)(setq interswitch t)) )
(patom 'diagnostics-)(print 'pocus?)(terpr)
(cond ((eq (read) 'y)(setq pocusswitch t)) )
loop (print (story)) (terpr)(terpr)
(print 'another?) (terpr)
(cond ((eq (read) 'n) (return)) )
(go loop)]
```

```
; This is the MAIN LOOP.
; The command "parse" calls a version of the MCHART parser (see PARSER)
; which parses a single sentence and then calls the functions in INTERFACE,
; which set the following variables:
;    tree - the parse tree for the sentence,
;    pronouns - the pronouns in tree form,
;    nps - the nps in lex/feature form,
;    sentence - the words in the sentence,

; In addition to parsing the input as a sentence, INTERFACE will also recognise
; single NPs, and these are used to send messages to the pronoun interpreter.
;  "the end" tells it that the current "story" is finished, whereupon it
;  returns the contents of the stack and hands control back to "stories".
;  "a topic" tells it that there has been a shift of topic, so that it clears
;  the centre and the list of f-centres before looping back to parse the next
;  sentence.

;  The variable "centre" is not strictly the same as Grosz's centre, because
; as the theory stands it is not always possible to uniquely identify the
; centre immediately. Instead, "centre" contains a list of possible centres,
; and this list is pruned as more information arrives - the system "focuses in"
; on one particular entity. Also note that the "entities" used here are really
; only NPs. A proper system would use frames or something.

;  If there are no pronouns in the sentence, then it is assumed that a new
; topic has been introduced, and so all the NPs of the sentence are centred.
;  However, if there are pronouns then the centre is assumed to be one of the
; pronominalized entities, in accordance with Grosz's rule. The referents are
; found and collected in a list called "referents", and "centre" is set to
; the intersection of the sets of centred and pronominalized entities.
; If none of the referents are centred, then a change of topic (probably back
; to some previous topic) is assumed to have occurred, and the centre is set
; to the list of referents.

;  Once "centre" is found, it is added to a stack which may be used to "hark
; back" to previous topics. Then the f-centres are set to be the list of
; entities referred to by NPs or pronouns in the sentence.


; Having found the referents, the centre, and the f-centres,
; the program prints out:
;    the sentence,
;    the NPs,
;    the pronouns,
;    their referents (if asked),
;    the centre,
;    the fcentres,
;    the parse tree (if asked),
;    messages announcing any change of topic or search of the stack.

; Finally, it prints out its interpretation of the sentence, with all pronouns
; replaced by their referents, and loops back to parse the next sentence.

(def story (lambda ()
(prog  (sentence tree pronouns nps fcentres
```

```
                stack flag flag2 referents centre)
(cent nil)
loop (parse)
(print sentence)(terpr)
(cond ((equal sentence '(the end)) (terpr)(patom 'stack_)(return (reverse stack)))
((equal sentence '(a topic)) (cent nil) (setq fcentres nil) (go loop)) )
(patom 'nps-)(print (cond (pocusswitch nps)
                          (t (words nps)) ))(terpr)
(patom 'pronouns-)(print (cond  (pocusswitch pronouns)
                                (t (words pronouns)) ))(terpr)
(cond ((null pronouns) (cent nps) (setq flag2 t) (go a) ))
(setq referents (bindings pronouns))
(cond (pocusswitch (patom 'referents-)(print referents)(terpr) ))
(cond ((null (intersect centre referents)) (setq flag2 t) (cent referents))
( t (cent (intersect centre referents) )) )
a (patom 'centre-)(print (cond (pocusswitch centre)
                                (t (words centre)) ))(terpr)
(stackit)
(setq fcentres (fcent referents nps))
(patom 'fcentres)(print (cond (pocusswitch fcentres)
                                (t (words fcentres)) ))(terpr)
(cond (pocusswitch (terpr)(patom 'tree-)(print tree)(terpr)) )
(cond (flag (print 'lookstack) (terpr)) )
(cond (flag2 (print 'topicshift) (terpr)) )
(terpr)(patom 'interpretation-)(print (interpretation tree)) (terpr)
(setq flag nil flag2 nil)
(go loop)]
```

```
; This sets the list of centred entities.
(def cent (lambda (a)(setq centre a)]

; This returns the intersection of two sets, but allowing for the possibility
; that the whole of the first set may be an element in the second set. This
; was used for centre= (jack jill), they= (jack jill) cases, but is now
; unnecessary.
(def intersect (lambda (a b) (cond
((null a) nil)
((memq a b) a)
((memq (car a) b) (cons (car a) (intersect (cdr a) b) ))
(t (intersect (cdr a) b)) ]

; This puts the centre onto the hark-back stack.
(def stackit (lambda () (setq stack (cons centre stack))]

; This just ensures that if something is referred to by both an NP and a
; pronoun it is only added to f-centres once.
(def fcent (lambda (referents nps)(cond
((null referents) nps)
((null nps) referents)
((memq (car referents) nps) (fcent (cdr referents) nps))
(t      (fcent (cdr referents) (cons (car referents) nps)))]
```

```
; This returns the words of the sentence, but with the pronouns replaced
; by their referents. A slight bug is that, if the antecedent NP has a pronoun
; in it, it will turn up again in the interpretation, but a proper approach
; would use frames anyway. The routine works by going through the parse tree,
; and collecting up the bindings of the pronouns, as well as the lexical entries
; for all other nodes. Both of these are entered on the tree in the same way as
; features are. This routine is an adaption of the INTERFACE function "words",
; and should be replaced by some sort of semantics in a realistic model.
(def interpretation (lambda (tree)
(cond ((atom tree) nil)
((atom (car tree))(cond
        ((eq (car tree) 'lex) (cdr tree))
        ((eq (car tree) 'pro) (words (binding tree)))
        (t (interpretation (cdr tree))) ))
((eq (car tree) 'pro) (binding tree))
(t (append (interpretation (car tree)) (interpretation (cdr tree)) ]
```

```
; PRONOUN INTERPRETATION


; This returns a list of referents for the pronouns.
; The first pronoun is taken, and the arguments in the sentence which c-command
; it are found using the "government" program. Then the following are tested as
; possible referents, within the confines of Chomsky's Matching and Binding
; conditions:
;       one of the centred entities,
;       one of the f-centres of the last sentence,
;       a set of centred entities (e.g. Jack and Jill),
;       something in the current sentence,
;       something previously centred from the stack.

; The order in which these alternatives are tried is crucial.
; If a possible referent is found, it is entered on the tree as a feature,
; as well as being added to the list "bindings". If none is found, "??" is
; entered on the tree. When this process has been done for all the pronouns,
; the list of referents is returned.
(def bindings (lambda (pronouns)
(prog (pronoun rest cargs)
(setq pronoun (car pronouns) rest (cdr pronouns))
(setq cargs  (caddr (government tree pronoun)) )
(return (cond
((null pronouns)   nil)
((centrebinds centre)      (cons (binding pronoun) (bindings rest)))
((subbind fcentres)(cons (binding pronoun) (bindings rest)))
((conjbind centre) (cons (binding pronoun) (bindings rest)))
((sentbind nps)    (cons (binding pronoun) (bindings rest)))
((stackbind stack)
      (setq flag t)(cons (binding pronoun) (bindings rest)))
(t (setq tree (hung tree pronoun
        (list 'pro (list 'binding (list(list 'lex '??))) (cdr pronoun)) ))
              (bindings rest) )]



; This attempts to bind a pronoun to some centred entity.
; If a referent can be found in the centre, it is returned. Otherwise nil.
(def centrebinds (lambda (centre) (prog (c)
(setq c centre)
loop (return (cond
((null c) nil)
((binds (car c)) (car c))
(t (setq c (cdr c)) (go loop)) ]

; This attempts to bind a pronoun to a forward centre in the previous sentence.
; Same as centbind really.
(def subbind (lambda (fcentres) (cond
((null fcentres) nil)
((binds (car fcentres)) (car fcentres) )
(t (subbind (cdr fcentres)) ) ]

; This attempts to bind the pronoun to a set of centred entities.
```

```
; A slight fudge is necessary to ensure that only animates are selected:
; only those NPs that are masculine/feminine are used.
(def conjbind (lambda (centre)(prog (agents possagents)
(setq possagents centre)
loop (cond
        ((null possagents) (go b))
        ((null (gender (car possagents))) (go a)) )
(setq agents (cons (car possagents) agents))
a (setq possagents (cdr possagents))
(go loop)
b (return (cond
        ((null agents) nil)
        ((null (cdr agents)) nil)
        (t (binds (list (cons 'lex (conjoin agents)) '(number pl)) )))]

; This takes a set like (jack jill) and returns (jack and jill).
(def conjoin (lambda (centre)(cond
((null centre) nil)
((null (cdr centre)) (words centre))
(t (append (words (car centre)) (cons 'and (conjoin (cdr centre))) ))]

; This looks for the referent of the pronoun within the sentence.
(def sentbind (lambda (x) (cond
((null x) nil)
((binds (car x)) (car x))
(t (sentbind (cdr x)) ) ]

; This looks in the stack for the referent.
(def stackbind (lambda (stack) (cond
((null stack) nil)
((centrebinds (car stack)) (car stack))
(t (stackbind (cdr stack)) ) ]
```

```
; CONSTRAINTS


; This returns the binding if anaphor and antecedent obey
; Chomsky's Matching and Binding conditions, as well as hanging the referent
; on the tree. Other constraints, such as semantic ones, should be added, but
; it is not clear that they would necessarily fit in here easily.
; (Remember, pronoun arrives in tree form.)
; Note that different binding conditions are imposed for ordinary pronouns
; and reflexive ones.
(def binds (lambda (np)(cond
((and (agree pronoun np)(eq (ccommands np cargs) (reflexive pronoun)))
(setq tree (hung tree pronoun (list 'pro (list 'binding np) (cdr pronoun))))
  (setq pronoun (list 'pro (list 'binding np) (cdr pronoun))) )
(t nil) ]


; This replaces a given node on a tree by some new node. It is used for
; marking the referents of pronouns.
(def hung (lambda (tree node newnode)
(cadr (hangtree tree node newnode))]

; A function for altering a parse tree-
; it searches out the leftmost node matching the description
; given, and then replaces it with the new node. It then
; returns a flag to say whether or not it found the node,
; followed by the new tree.
        (def hangtree (lambda (tree node newnode)(prog (a)
(cond   ((atom tree) (return (list nil tree)))
        ((equal tree node)  (return (list 't newnode)) ))
(setq a (hangtree (car tree) node newnode))
(cond ((car a) (return (list 't (cons (cadr a) (cdr tree)))) ))
(setq a (hangtree (cdr tree) node newnode))
(return (list (car a) (cons (car tree) (cadr a)))) ]
```

```
; MATCHING CONDITIONS

; Now a series of functions used to check that two items agree in number
; and gender.
(def agree (lambda (a b)
(and (numbagree a b) (gendagree a b) ]

(def numbagree (lambda (a b)
(eq (number a) (number b)]

(def number (lambda (constituent)
(car (getfeature constituent 'number 'sing))]

(def gendagree (lambda (a b)
(eq (gender a) (gender b))]

(def gender (lambda (constituent)
(car (getfeature constituent 'gender 'nil))]

; This takes a constituent as it would occur on the tree, and returns the
; value of a given feature. There is an option for allowing some features to
; have defalt values - for example, NPs are assumed to be singular and neuter
; unless otherwise marked.
; The result comes back in brackets.
(def getfeature (lambda (constituent feature default)(cond
((null constituent)  (list default) )
((atom (car constituent)) (getfeature (cdr constituent) feature default))
((eq (caar constituent) feature) (cdar constituent))
(t (getfeature (cdr constituent) feature default))]

; This returns the referent of a pronoun if there is one. Referents are hung on
; the tree in the same way as features, so "getfeature" is used.
(def binding (lambda (pronoun)
(car (getfeature pronoun 'binding 'nil))]

; This returns t if a pronoun is reflexive.
(def reflexive (lambda (pronoun)
(car (getfeature pronoun 'reflexive 'nil))]

; Another function for finding the words in a constituent.
; Could use 'words', but this may be quicker.
(def lexitems (lambda (constituent)
(getfeature constituent 'lex 'nil)]
```

```
; BINDING CONDITIONS

; A function to check whether or not an np is one of the c-arguments of
; the pronoun's governing category. Note that the cargs must be "interpreted",
; since they are in tree form, whereas the np will be in lex/feature form.
(def ccommands (lambda (np cargs)(cond
((null cargs) nil)
((equal (words np) (interpretation (car cargs)) ) t)
(t (ccommands np (cdr cargs)))]


; This does a version of Chomsky's theory of government, although it only
; looks at pronouns, rather than NPs in general. Binding rules are provided
; by the function "binds" above. A slight difference with Chomsky's theory
; is that he talks about binding NPs (some of which are pronouns), whereas
; I just look at pronouns. I would ultimately prefer to do it his way.
; This returns:
;     a switch- nil= pronoun not found
;               1  = pronoun found but no governing element found yet
;               2  = governing element found, but no governing category yet
;               3  = no governing element because of sbar or np barrier
;               4  = governing category found.
;     the governing category for the pronoun.
;     the c-commanding arguments of the governing category.
;     the governing element (all last three in tree form).
; It searches the tree for the pronoun, and when it finds it returns '1'
; to say that the constituent under consideration contains it.
; Once '1' has been signalled, it looks for the governing element,
; collecting c-arguments as it goes, searching progressively higher up the
; tree until it finds it (signalling '2') or is stopped by an sbar
; or np barrier (signalling '3').
; Once '2' has been signalled, it looks for the next s or np that contains
; the governing element, still collecting c-arguments as it goes,
; and passes on the tree that has that as top node as the governing element.
; Once 3 or 4 are signalled, the results are just passed on up and out.

(def government (lambda (tree pronoun)
(prog (nodes nodesleft current switch govcat cargs governor x )
(cond ((equal tree pronoun)            (return (list '1 nil nil nil)))
      ((equal tree (list 'np pronoun))  (return (list '1 nil nil nil))) )
(setq nodes (cdr tree))
(setq nodesleft nodes)
loop1 (cond ((null nodesleft) (return)) )
(setq current (car nodesleft))
(cond ((atom current) (return)))
(setq x (government current pronoun))
(setq switch (car x))
(setq govcat (cadr x))
(setq cargs  (caddr x))
(setq governor (cadddr x))

(cond ((null switch) (setq nodesleft (cdr nodesleft))
                     (go loop1))

      ((eq switch '1) (return (prog (y)
```

```
                (setq y nodes)
          loop2 (cond
                 ((null y)          (go loop3))
                 ((and (eq (caar y) 'np) (not(equal (cadar y) pronoun)))
                        (setq cargs (cons (car y) cargs)))
                 ((eq (caar y) 'v) (setq governor (car y)))
                 ((eq (caar y) 'n) (setq governor (car y)))
                 ((eq (caar y) 'Adj) (setq governor (car y)))
                 ((eq (caar y) 'Tense) (setq governor (car y)))
                 ((eq (caar y) 'Prep) (setq governor (car y))) )
                (setq y (cdr y))
                (go loop2)
          loop3 (return (cond
                 (governor (cond
        ((eq (car tree) 's) (list '4 tree cargs governor))
        ((eq (car tree) 'np)(list '4 tree cargs governor))
        (t                  (list '2  nil cargs governor)) ))
                 ((eq (car tree) 'sbar) (list '3 nil nil governor))
                 ((eq (car tree) 'np)   (list '3 nil nil governor))
                 (t                     (list '1 nil nil governor)) )))))

((eq switch '2) (return (prog (y)
                (setq y nodes)
          loop4 (cond
                 ((null y) (go loop5))
                 ((and (eq (caar y) 'np) (not(equal (cadar y) pronoun)))
                        (setq cargs (cons(car y) cargs))) )
                (setq y (cdr y))
                (go loop4)
          loop5 (return (cond
                ((eq (car tree) 's) (list '4 tree cargs governor))
                ((eq (car tree) 'np)(list '4 tree cargs governor))
                ( t                 (list '2  nil cargs governor)) )) )))

(  t            (return (list switch govcat cargs governor)))]
```

```
{;; INTERFACE- The pronoun interpreter calls the parser, and at the
                end of the parse the parser signals 'parsecomplete',
                which in turn calls various functions which search
                the parse tree for noun phrases and pronouns and set
                appropriate variables at the top level. The pronoun
                interpreter goes on to use these, eventually calling
                the parser again for the next sentence.


{;; At the end of the parse the parser signals 'parsecomplete',
       and this in turn calls 'outparse'.}
(setq signaltable (cons '(parsecomplete outparse) signaltable))

{;; This accepts labels for the left and right hand ends of the chart,
       the distinguished category (S), and the type of edge (inactive)
       signalled by 'parsecomplete' which calls it. It returns nil,
       but applies 'infocus' to all the parse trees for NP or S.
       'infocus' sets variables in the main program, so if there is
       more than one parse only the last one will be used.
       The parse trees may be displayed at this point as part of the
       interface diagnostics.}
(def outparse (lambda (left right topcat type)
(prog ()
(cond (interswitch (showedges left right topcat type)(terpr)))
(mapedges left right 'np 'inactive 'infocus)
(mapedges left right topcat type 'infocus)]

{;; 'infocus' is called by 'mapedges', accepts an edge corresponding
       to a parse (together with a test predicate which is not used),
       and returns nil.
     On the way it sets these variables:
       'tree'- the parse tree for the sentence (or NP),
       'sentence'- the words of the sentence,
       'nps'- the non-pronominal NPs (but not those embedded in other NPs-
              right or wrong?) in the form: ((lex the man) (gender masc))
       'pronouns'- the pronouns in tree form.}
(def infocus (lambda (edge test)
(prog ()
(setq tree (normalize <edge:label>))
(cond (interswitch (patom 'nptrees-)(print (subtree tree 'np))(terpr)))
(setq nps (skimsearch tree 'np))
(setq pronouns (subtree tree 'pro))
(setq nps (chop pronouns nps))
(setq sentence  (words tree) ) ]

{;; This takes a constituent in tree form and returns a list of its features
       e.g. ((gender masc)(reflexive t)) .It is designed for NPs, and works
       by inheriting features from the head. NPs of the form (NP conj NP)
       are just assumed to be plural.}
(def features (lambda (constit)(prog (topnode)
(cond ((atom constit) (return)) )
(setq topnode (car constit))
(cond
((eq topnode 'conj) (return '((number pl)) ))
((eq topnode 'n) (return (cddr constit) ))
```

```
((eq topnode 'propn) (return (cddr constit) ))
((eq topnode 'pro) (return (cddr constit) ))
((eq topnode 'np) (return (features (head constit))))
(t (return))]

{;; Takes an NP and returns the head by looking for the first daughter
       node which is a proper noun, noun, conjunction, NP, or pronoun
       in that order.}
(def head (lambda (np)(prog (a)
(cond ((atom np)(return)) )
(setq a (daughter np 'propn))
(cond (a (return a)) )
(setq a (daughter np 'n))
(cond (a (return a)) )
(setq a (daughter np 'conj))
(cond (a (return a)) )
(setq a (head (daughter np 'np)))
(cond (a (return a)) )
(setq a (daughter np 'pro))
(cond (a (return a)) )
(return)]

{;; This re-orders a parse tree into conventional form, since the parser
       provides 'backward' trees}
(def normalize (lambda (parsetree)(cond
((atom parsetree) parsetree)
(t (cons (car parsetree) (reverse (mapcar 'normalize (cdr parsetree))) ]

{;; This finds the words in a constituent by searching for lists
       of the form: (lex the man) }
(def words (lambda (x)
(cond ((atom x) nil)
((atom (car x))(cond
       ((eq (car x) 'lex) (cdr x))
       (t (words (cdr x))) ))
(t (append (words (car x)) (words (cdr x)) ]

{;; This takes a constituent in tree form and returns the first immediate
       daughter node of the given category.}
(def daughter (lambda (con cat)(cond
((atom con) nil)
((atom (car con)) (daughter (cdr con) cat))
((eq (caar con) cat) (car con))
(t (daughter (cdr con) cat))]

{;; This chops the pronouns out of the NP list- not quite just set intersection
       since pronouns are in tree form and NPs are in lex/feature form.}
(def chop (lambda (pronouns nps)(cond
((null nps) nil)
((null pronouns) nps)
(t (chop (cdr pronouns) (remove (cdr (car pronouns)) nps))]

{;; This finds all the subconstituents of a given category in tree form.}
(def subtree (lambda (constit cat)(cond
((atom constit)        nil)
```

```
((atom (car constit)) (cond
        ((eq (car constit) cat) (cons constit
                                  (subtree (cdr constit) cat) ))
        (t (subtree (cdr constit) cat)) ))
(t (append (subtree (car constit) cat) (subtree (cdr constit) cat)]

{;; This finds subconstituents of a given category in lex/feature form}
(def deepsearch (lambda (constit cat)(cond
((atom constit)         nil)
((atom (car constit)) (cond
        ((eq (car constit) cat) (cons (cons (cons 'lex (words constit))
                                            (features constit))
                                  (deepsearch (cdr constit) cat) ))
        (t (deepsearch (cdr constit) cat)) ))
(t (append (deepsearch (car constit) cat) (deepsearch (cdr constit) cat)]


{;; This finds subconstituents of a given category in lex/feature form,
        but at the top level only.}
(def skimsearch (lambda (constit cat)(cond
((atom constit)         nil)
((atom (car constit)) (cond
        ((eq (car constit) cat) (cons (cons (cons 'lex (words constit))
                                            (features constit)) nil) )
        (t (skimsearch (cdr constit) cat)) ))
(t (append (skimsearch (car constit) cat) (skimsearch (cdr constit) cat)]
```

Appendix 2 Grammar and Lexicon

These two programs contain the grammar and lexicon used by the parser. Because the parser does not use case frames, verbs with different numbers of objects are classified separately, e.g. "Vtwo" for a verb like "give". "clue words" are classified under "control".

```
{;; GRAMMAR - This sets up a simple context-free grammar for the parser to use}

(SETQ GRAMCOMS (QUOTE ((VARS Grammar))))

(SETQ Grammar (QUOTE (  (S S2)
                        (S control S2)
                        (S condition S2)
                        (condition cond S2)
                        (S2 NP VP)
                        (S2 PP NP VP)
                        (Sbar Comp S)
                        (Sbar S)
                        (VP VP Adv)
                        (VP Adv VP)
                        (VP Vtwo NP NP)
                        (VP Vint)
                        (VP be Adj)
                        (VP be NP)
                        (VP be part)
                        (VP Vepi Sbar)
                        (VP Vloc PP)
                        (VP V NP PP)
                        (VP V NP)
                        (NP NP Conj NP)
                        (NP NP Sbar)
                        (NP Det N)
                        (NP Det AP N)
                        (NP PropN)
                        (NP pro)
                        (AP Adj)
                        (Det Quant)
                        (Det pro)
                        (Det NP /'s)
                        (Det Art)
                        (PP Prep NP))))
```

```
{;; LEXICON - This is the dictionary for the parser. Because there are no
        case frames etc. restrictions on the number of objects have to be
        expressed by putting verbs into different classes. Also note that
        because only one parse is handed to the pronoun interpreter,
        problems arise if words are form-class ambiguous.}

(SETQ DICTCOMS (QUOTE ((VARS Dict Redun))))

(SETQ Dict (QUOTE (
                        (thinks (Vepi))
                        (think (Vepi))
                        (thought (Vepi))
                        (says (Vepi))
                        (said (Vepi))
                        (knew (Vepi))
                        (gives (Vtwo))
                        (gave (Vtwo))
                        (offered (Vtwo))
                        (runs (Vint))
                        (bought (V))
                        (had (V))
                        (has (V))
                        (went (Vloc))
                        (saw (V))
                        (like (V))
                        (likes (V))
                        (liked (V))
                        (met (V))
                        (treats (V))
                        (treat (V))
                        (treated (V))
                        (ignored (V))
                        (watched (V))
                        (drove (V))
                        (loves (V))
                        (love (V))
                        (loved (V))
                        (hated (V))
                        (was (be))
                        (were (be))
                        (is (be))
                        (are (be))
                        (pretending (part))
                        (lying (part))
                        (and (Conj))
                        (that (Comp))
                        (with (Prep))
                        (to (Prep))
                        (from (Prep))
                        (in (Prep))
                        (at (Prep))
                        (on (Prep))
                        (i (NP))
                        (you (NP))
```

```
(he (pro (gender m)))
(she (pro (gender f)))
(it (pro))
(this (pro))
(that (pro))
(they (pro (number pl)))
(him (pro (gender m)))
(her (pro (gender f)))
(them (pro (number pl)))
(himself (pro (gender m)(reflexive t)))
(herself (pro (gender f)(reflexive t)))
(itself (pro (reflexive t)))
(themselves (pro (number pl)(reflexive t)))
(his (pro (gender m) (possessive t)))
(her (pro (gender f)(possessive t)))
(their (pro (number pl)(possessive t)))
(its (pro (possessive t)))
(Jack (PropN (gender m)))
(Jill (PropN (gender f)))
(Jim (PropN (gender m)))
(Janet (PropN (gender f)))
(John (PropN (gender m)))
(some (Quant))
(two (Quant))
(a (Art))
(the (Art))
(/'s (Pos))
(end (N))
(topic (N))
(way (N))
(cold (N))
(car (N))
(bar (N))
(table (N))
(book (N))
(drink (N))
(college (N))
(cinema (N))
(friends (N (number pl)))
(friend (N))
(plates (N (number pl)))
(flowers (N (number pl)))
(men (N (number pl)))
(man (N (gender m)))
(sailor (N (gender m)))
(good (Adj))
(happy (Adj))
(nice (Adj))
(ill (Adj))
(lonely (Adj))
(tired (Adj))
(green (Adj))
(cruel (Adj))
(today (Adv))
(last-night (Adv))
```

```
(tonight (Adv))
(slowly (Adv))
(obviously (Adv))
(too (Adv))
(together (Adv))
(all (Adv))
(both (Adv))
(although (cond))
(and (control))
(but (control))
(however (control))
(anyway (control))
(similarly (control))
(meanwhile (control))
(by-the-way (control))
                        )))
```

appendix 3 EXAMPLES

The following examples have been used to test POCUS. They constitute a representative sample of the sort of data the program can and can't handle. The symbol " --> " is used to denote that a topic shift is introduced at this point by the operator.

	EXAMPLE	POCUS'S INTERPRETATION
1	john likes himself.	john likes john.
2	john likes him.	john likes ??
3	jack thinks he is ill.	jack thinks jack is ill.
4	i saw john today. jack thought he was ill. i thought he was pretending.	i saw john today. jack thought john was ill. i thought john was pretending.
5	i saw john today. jack said he had a cold. i thought he was pretending.	i saw john today. jack said john had a cold. i thought john was pretending.
6	i saw john today. --> by the way, jack says he has a cold. i think he is pretending.	i saw john today. by the way, jack says jack has a cold. i think jack is pretending.
7	i saw john today. jack said he had a cold. i think he was lying.	i saw john today. jack said john had a cold. i think john was lying.
8	jack was a sailor. jill liked the way he treated her.	jack was a sailor. jill liked the way jack treated her.
9	jack was a sailor. jill loved him. janet liked the way he treated her.	jack was a sailor. jill loved jack. janet liked the way jack treated jill
10	jack was a sailor. jill loved him. --> similarly, janet liked the way he treated her.	jack was a sailor jill loved jack similarly janet liked the way jack treated janet.
11	jack was a sailor. john liked him. he liked the way he treated him.	jack was a sailor. john liked jack. jack liked the way jack treated john.
12	jack was a sailor. john liked him. he liked the way jack treated him.	jack was a sailor. john liked jack. jack liked the way jack treated john.

13 jack was a sailor. he was john's friend. he liked the way he treated him.	jack was a sailor. jack was john's friend. jack liked the way jack treated john.
14 jack met jill in a bar. they went to the cinema.	jack met jill in a bar. jack and jill went to the cinema.
15 jack and jill bought a car. they drove it to the cinema.	jack and jill bought a car. jack and jill drove a car to the cinema
16 john bought some plates. they were green.	john bought some plates. some plates were green.
17 jack and janet bought some plates. they were green.	jack and janet bought some plates. jack and janet were green.
18 jim and jill bought some plates. they liked them.	jim and jill bought some plates. jim and jill liked some plates.
19 jack and jill met janet and john. they saw them at the college. they all went to the cinema together.	jack and jill met janet and john. jack and jill saw janet and john at the college. jack and jill all went to the cinema together.
20 jill was cruel. john hated the way she treated him.	jill was cruel. john hated the way jill treated john.
21 jack and jill went to the cinema. they met john on the way. they knew he was lonely.	jack and jill went to the cinema. jack and jill met john on the way. jack and jill knew john was lonely
22 jack and jill met john in a bar. they offered him a drink.	jack and jill met john in a bar. jack and jill offered john a drink.
23 jack ignored jill. they both knew she was lying.	jack ignored jill. jack and jill both knew jill was lying.
24 john met jill in the bar. they saw his friend at a table.	john met jill in the bar. john and jill saw john's friend at a table.
25 jack met jill at the college. they both liked his friend.	jack met jill at the college. jack and jill both liked jack's friend.
26 jack liked jill. he thought they were good friends.	jack liked jill. jack thought jack and jill were good friends.
27 jack watched john. he was tired.	jack watched john. jack was tired.
28 jack ignored john. he was tired.	jack ignored john. jack was tired.
29 jack loved jill.	jack loved jill.

he gave her some flowers. they were nice.	jack gave jill some flowers. some flowers were nice.
30 jack loved jill. his friends thought they were happy together.	jack loved jill. jack's friends thought jack and jill were happy together.
31 on his table john saw a book.	on john table john saw a book.
32 although he hated it john drove the car.	although john hated the car john drove the car.

Appendix 4 A Sample Session.

The following is a sample of the sort of input and output POCUS uses. This session involves the processing of examples the following examples:

John likes himself.

john likes him.

Jim and Jill bought some plates.
They liked them.

Jack and Jill went to the cinema.
They met John on the way.
They knew he was lonely.

Jack loved Jill.
He gave her some flowers.
They were nice.

Jack loved Jill.
His friends thought they were happy together.

```
Franz Lisp, Opus 38.22
Hi leslie
-> (load 'baby)
[load baby]
[load pocus]
[fasl /u1/ht/lisp/myuci.o]
load debugger - now or never? n
load tracer - now or never? n
[load myuci2.l]
[fasl /u1/ht/lisp/clisp.o]
[fasl /u1/ht/lisp/if.o]
[fasl /u1/ht/lisp/recm.o]
[fasl /u1/ht/lisp/record.o]
[fasl /u1/ht/lisp/myio.o]
[fasl /u1/ht/lisp/agenda.o]
[fasl /u1/ht/lisp/tracer.o]
[fasl /u1/ht/lisp/chart.o]
[fasl /u1/ht/lisp/fns.o]
[fasl /u1/ht/lisp/dictfns.o]
[fasl /u1/ht/lisp/td.o]
[load parser]
[load grammar]
[load lexicon]
[load interface]
t
-> (stories)
diagnostics-interface? n
diagnostics-pocus? n

Parse as s: john likes himself
```

```
nps-(john)
pronouns-(himself)
centre-(john)
fcentres(john)
topicshift

interpretation-(john likes john)

Parse as s: the end

stack_((((lex john) (gender m) (pn s3))))

another?
y

Parse as s: john likes him
nps-(john)
pronouns-(him)
centre-nil
fcentres(john)
topicshift

interpretation-(john likes ??)

Parse as s: the end

stack_(nil)

another?
y

Parse as s: jim and jill bought some plates
nps-(jim and jill some plates)
pronouns-nil
centre-(jim and jill some plates)
fcentres(jim and jill some plates)
topicshift

interpretation-(jim and jill bought some plates)

Parse as s: they liked them
nps-nil
pronouns-(they them)
centre-(jim and jill some plates)
fcentres(jim and jill some plates)

interpretation-(jim and jill liked some plates)

Parse as s: the end

stack_((((lex jim and jill) (number pl))
        ((lex some plates) (number pl) (pn s3)))
       (((lex jim and jill) (number pl))
        ((lex some plates) (number pl) (pn s3))))
```

```
another?
y

Parse as s: jack and jill went to the cinema
nps-(jack and jill the cinema)
pronouns-nil
centre-(jack and jill the cinema)
fcentres(jack and jill the cinema)
topicshift

interpretation-(jack and jill went to the cinema)

Parse as s: they met john on the way
nps-(john the way)
pronouns-(they)
centre-(jack and jill)
fcentres(jack and jill john the way)

interpretation-(jack and jill met john on the way)

Parse as s: they knew he was lonely
nps-nil
pronouns-(they he)
centre-(jack and jill)
fcentres(jack and jill john)

interpretation-(jack and jill knew john was lonely)

Parse as s: the end

stack_((((lex jack and jill) (number pl))
        ((lex the cinema) (pn s3)))
       (((lex jack and jill) (number pl)))
       (((lex jack and jill) (number pl))))

another?
y

Parse as s: jack loved jill
nps-(jack jill)
pronouns-nil
centre-(jack jill)
fcentres(jack jill)
topicshift

interpretation-(jack loved jill)

Parse as s: he gave her some flowers
nps-(some flowers)
pronouns-(he her)
centre-(jack jill)
fcentres(jill jack some flowers)

interpretation-(jack gave jill some flowers)
```

```
Parse as s: they were nice
nps-nil
pronouns-(they)
centre-(some flowers)
fcentres(some flowers)
topicshift

interpretation-(some flowers were nice)

Parse as s: the end

stack_((((lex jack) (gender m) (pn s3))
         ((lex jill) (gender f) (pn s3)))
        (((lex jack) (gender m) (pn s3))
         ((lex jill) (gender f) (pn s3)))
        (((lex some flowers) (number pl) (pn s3))))

another?
y

Parse as s: jack loved jill
nps-(jack jill)
pronouns-nil
centre-(jack jill)
fcentres(jack jill)
topicshift

interpretation-(jack loved jill)

Parse as s: his friends thought they were happy together
nps-(his friends)
pronouns-(his they)
centre-(jack)
fcentres(jill and jack jack his friends)

interpretation-(jack friends thought jill were happy together)
```

APPENDIX FIVE RULES FOR FOCUS ANALYSIS

Pronoun Interpretation Rules.

Potential local foci are items referred to by NPs, pronouns, and elliptical expressions, excluding those appearing solely in PPs.

When a demonstrative pronoun or stressed pronoun is encountered try:

1) The local focus.
2) The potential foci of the last utterance.
3) Deixis (especially if other clues like gestures suggest it).

When a pronoun is found (including demonstrative pronouns) try:

4) The local focus.
5) The potential foci of the last utterance.
6) Those of the current utterance.
7) The global focus.
8) Previous foci.

9) When a plural pronoun is encountered (and in some cases singular pronouns as well), try generalising from the local focus, keeping the generalised referent in local focus.

10) When an elliptical possessive or quantifier, or a phrase containing the vague pronoun "one(s)", is found, try the following as prototypes for the referent:

a) The global focus.
b) The local focus.
c) A previous focus.

and put the prototype into global focus.

The terms "local focus" and "potential local focus" replace "centre" and "f-centre". "Utterance" is used to allow for the possibilty of elliptical phrases, etc. and I will assume that any two parts of a sentence separated by an interjection from the other speaker (such as an acknowledgement) constitute separate utterances (there is some evidence on the basis of this that an analysis based on clauses, not sentences would be better). "Previous foci" includes both global and local.

Focussing Rules.

11) When a demonstrative or stressed pronoun is used anaphorically, focus (local or global) may have recently shifted to an item just mentioned. Otherwise a contrast may be implied.

12) When a connective like "now" is used, a new focus is signalled

13) When a connective like "so", "then", or "but" is used a shift to a sub-topic or a return to a previous topic may be signalled.

To find the local focus, try:

14) Any item referred to deictically, or by a demonstrative or stressed pronoun.
15) Any item referred to via rule 4.
16) Any other item referred to anaphorically.
17) Any other potential local focus.

In addition, we add these rules for finding global foci:

18) When an item not recently mentioned or in local focus is referred to anaphorically, it may be the global focus.

19) When a PP contains an anaphor that refers back to a previous focus, that becomes the global focus.

20) When a definite NP refers to an "implied" item, the implier is probably one of the foci, and becomes the global focus.

APPENDIX SIX AN EXAMPLE OF FOCUS ANALYSIS

The following is an extract from "Gardening Time", an STV television programme. The two participants, A and B, are talking about chrysanthemums.

1) A: 'and then if you look inside the leaf there you can see perhaps a hundred greenfly even in that little ... ' [points to leaf]

2) B: 'Where does the stickiness come from?'

3) A: 'That's the honeydew that they excrete.'

4) B: 'I see.'

5) A: 'And it just collects over the plant.'

6) B: 'Yes.'

7) A: 'And you get fungus growing on that as well. And some of the trees you see that - a black fungus growing on it.'

8) B: 'Do you really? Very nasty indeed.'

9) A: 'which you notice. And also ladybirds are very useful. There's a little ladybird there look.'

10) B: 'Now what does she do? She eats them, does she?'

11) A: 'Yeah, she eats them. Some people think that they've got to kill everything, but ...'

12) B: 'Ants milk them ...'

13) A: 'Yes, ants milk ...'

14) B: 'And ladybirds eat them.'

15) A: 'Yes. Yes, it is strange, isn't it? But the less that you kill ladybirds and their larvae the better, because they'll certainly eat their way through a few. So that's quite helpful.'

16) B: 'Well, let's just hope that she won't fly away home.'

17) A: 'Oh look, there's ...'

18) B: 'What's that yellow thing there?' [points to eggs]

19) A: 'Yes, I hadn't seen those. Those are some butterfly eggs.'

20) B: 'Oh, I see.'

21) A: 'Yes, those are from the cabbage white.'

22) B: 'Uh-huh.'

23) A: 'They'll get killed before they hatch out.'

24) B: 'Yes.'

25) A: 'Otherwise they would cause a bit of a problem.'

26) B: 'Uh-huh.'

27) A: 'And there's a little blackfly' [points to blackfly]
'We've got everything on this one, haven't we?'

28) B: 'Yes, we have, haven't we? Yes.'

29) A: 'Shan't have to take this home. Um - there's a little blackfly ...'

30) B: 'You're going to leave it here with us, are you?'

31) A: 'Yes, I wouldn't be very popular if I did that, would I? But - um - blackfly and greenfly are the two commonest problems I think that you get.'

32) B: 'Yes. Now what's the cure then?'

33) A: 'Any of the insecticides that are recommended at your local garden centre.'

34) B: 'Yes.'

35) A: 'But there's a good one here - Malathion - that's fairly safe.'
[holds up a bottle of Malathion]

36) B: 'Yes.'

37) A: 'I don't think you want anything too powerful.'

38) B: 'No, I like the Pyrethrum ones for that reason.'

39) A: 'Yes, Pyrethrum. Yes, Derris as well, that's er - and we use nicosoap - nicotine.'

40) B: 'Oh, do you? Yes. Yes.'

41) A: 'We find that's quite good.'

42) B: 'and Quosia? Do you use that?'

43) A: 'No, we haven't got round to ...'

44) B: 'No.'

45) A: 'I'll have to try that one.'

46) B: 'Yes, Quosia's rather a good one.'

47) A: 'But it's also important to change them round.'

48) B: 'Hmm.'

49) A: 'I mean it's no good saying "Oh I always use Malathion. As soon as you see any ..."'

50) B: 'Because they get used to it.'

51) A: 'Yes they do. And eventually you get some that even thrive on it.'

52) B: 'Oh, I say.'

53) A: 'It's like the Super-rats and things like that. But that's a good standard one.' [Holds up the bottle again.]

54) B: 'Now how - let's just lift this up onto the table so that we can see what a very large and magnificent plant it is, apart from the greenfly of course.

[A puts the chrysanthemum on the table]

'Now I want you to tell us, John, when you first took the cutting for this, how old it is, and what it's going to be.'

55) A: 'Yes well this is one of the large ones - what we used to call the Japs or the Giant Exhibition.'

56) B: 'Yes.'

57) A: 'So we will end up with one bloom, as large as we can get, on the top.'

58) B: 'Uh-huh. A sort of big mop-head type of thing.'

59) A: 'Yes, they do call them mop-heads.'

60) B: 'Yes.'

61) A: 'So because it's rather a specialised thing, we take the cutting fairly early.'

62) B: 'What, January?'

63) A: 'Yes, or even before we go away for the christmas break.'

64) B: 'Oh really?'

65) A: 'The first few go in then.'

66) B: 'Yes.'

67) A: 'There's always some exhibitors that we sell to that want them.'

68) B: 'And these are the tiny cuttings that you get off the stools of the old plants.' [points]

69) A: 'That's right, yes. Yes, they're growing up, and you need to take one that's growing quite nicely, hasn't gone hard or ...'

70) B: 'And take the biggest one, I suppose.'

71) A: 'Yes - well - yes, the best looking one.'

72) B: 'Yes.'

73) A: 'It's very important - if you don't start off with the right thing, who knows what you're going to grow into?'

74) B: 'Well you don't want to start off with the runt of the litter, do you?'

75) A: 'No, no. No, we soon get rid of those.'

76) B: 'So this then is as old as from January.'

77) A: 'Yes. Yes, it starts very slowly ...'

78) B: 'And when will it be blooming?'

79) A: 'This will be blooming in the first week of November - last weekend of October.'

80) B: 'Yes.'

81) A: 'We try and time ours. We have an open weekend - last weekend of October - and we try and get them for then.'

82) B: 'Yes. Now what was the procedure? You've not kept this in the greenhouse all the time.'

83) A: 'No, this is standing outside at the moment.'

84) B: 'Yes.'

85) A: 'It starts off in the greenhouse obviously at that time of year, grows very slowly - keep them as cool as possible.'

86) B: 'Yes. Put them out when? End of May?'

87) A: 'Yes. Yes, once the frost has gone. You put them in the cold frame before then.'

88) B: 'Yes. Yes, marvellous.'

89) A: 'So - and then it's stopped. They've pinched the top out.'

90) B: 'Yes.'

91) A: 'Stopping worries some people, but it shouldn't really.'

92) B: 'No.'

93) A: 'It's quite a simple thing. If you leave the chrysanthemum it will go on and grow normally anyway.'

94) B: 'and become a spray.'

95) A: 'Yes. Yes.'

96) B: 'What they call a spray chrysanthemum.

97) A: 'But if you want them for a particular date, you often .

98) B: 'Tell us then about these little chaps here.' [points to three small plants off to one side]

99) A: 'Yes.'

100) B: 'Which one do you want to talk about first?'

101) A: 'well, the one on the right.'

102) B: 'That one.' [points]

103) A: 'This one.' [points and nods in agreement
'With this hot weather we've been having

104) B: 'Is this called a Charm chrysanthemum?'

105) A: 'Yes, that's a Charm. They're all three Charms. And I think one of the problems people have had is that it's got hot and dry .

106) B: 'Yes.'

107) A: 'And some of these go to bud too early.'

108) B: 'Uh-huh.'

109) A: 'So you get these early flowers which - they don't look very good.'

110) B: 'Yes, some of them there, like that for instance has gone off completely, hasn't it?' [points to flower]

111) A: 'Yes. So this was flowering - it started flowering about the middle of July, which is far too early.'

112) B: 'When do you want them to be in full bloom?'

113) A: 'Again, really in October, but they slow down at the end, so if they're coming in in September ...'

114) B: 'I've forgotten to ask you, what do you fertilise your chrysanthemums with?'

115) A: 'Um - we just use a general fertiliser. We sell our own, and we use that obviously, but ...'

116) B: 'My Grandfather used to collect a bit of sheep manure and press the lumps in with his thumb. I can see him doing it now. And he used to think that that was the most wonderful fertiliser for them.'

117) A: 'I should think it must have worked fairly well.'

118) B: 'Yes. Why sheep, I don't know.'

119) A: 'Yes, I haven't tried that one either. I've obviously got something to learn when I get back, haven't I?'

120) B: 'Well he insisted on having sheep.'

121) A: 'Yes, but all these things work. Some people feel that the more feed you put in, the bigger and better plant you get.'

122) B: 'Yes.'

123) A: 'But this doesn't - this isn't true. It's far too easy to overfeed, and then you get a lush plant that's susceptible to disease.'

124) B: 'And it's very wrong, I understand, to feed an ailing plant. They say it's like giving a baby who's not very well caviar or something like that, and therefore not doing very well. Now what about the dahlias there?'

125) A: 'Um - the dahlias - yes. These are again a few specimens that I - normally dahlias would grow in the open ground.'

126) B: 'Yes, of course.'

127) A: 'And these have been dug up recently, and they're suffering a bit. But ...'

128) B: 'You've just brought them to demonstrate to us.'

129) A: 'Yes, yes. They - they - these have also had not quite enough light. When you're growing dahlias it's important to have a good open area - get plenty of light in.'

130) B: 'Yes.'

131) A: 'They like plenty of feed ...'

132) B: 'Yes'

133) A: '...dahlias do, and they don't want to dry out at all.'

134) B: 'No.'

135) A: 'It's very important with dahlias that they don't dry out too much. But if you take this one here, you can see that there are a central bud and then a bud either side.'

136) B: 'Would you take the side ones off?'

137) A: 'Yes, I would just take those off. Some people worry about taking things off ...'

138) B: 'Yes, well, that's grand. Thank you very much.'

ANALYSIS

The above text is here analysed using the rules of appendix 5. Numbers on the left refer to lines in the text - those labelling utterances that play no focussing role (e.g. acknowledgements) have been ommitted. Numbers on the right refer to the rules used for interpretation and focussing, so that "4,15" in the second column means that a pronoun has been used to refer to the local focus using rules 4 for interpretation and rule 15 for focussing, while "10" in the fourth column means that, by rule 10, a phrase like "a few" has just put some item into global focus. Letters refer to additional notes below.

	LOCAL FOCI	RULES	GLOBAL FOCI	RULES
1)	GREENFLY	14,13	CHRYSANTHEMUM?	A
2)	HONEYDEW	14	?	
3)	HONEYDEW	1,11	GREENFLY	8,18
5)	HONEYDEW	4,15	GREENFLY?	B
7)	FUNGUS	17	HONEYDEW	1,11
9)	LADYBIRDS	17	?	
	LADYBIRD	17	?	
10)	LADYBIRD	4,15	GREENFLY	8,18
11)	LADYBIRD	4,15	GREENFLY	7,18
12)	GREENFLY	7,16	?	
13)	GREENFLY	4,15	?	
14)	GREENFLY	4,15	?	
15)	LADYBIRDS & LARVAE	17,13	GREENFLY	10
16)	LADYBIRD	8,16	?	
18)	EGGS	3,14	?	
19)	EGGS	1,14	?	
21)	EGGS	1,14	?	
23)	EGGS	1,15	?	
25)	EGGS	1,15	?	
27)	BLACKFLY	14	?	
	"EVERYTHING"	17	CHRYSANTHEMUM	19, C
29)	CHRYSANTHEMUM	3,14	?	
	BLACKFLY	14	?	
30)	CHRYSANTHEMUM	4,15	?	
31)	BLACKFLY & GREENFLY	17,13	CHRYSANTHEMUM(S)	20
32)	CURE	17,12	PESTS	20,D,E
33)	INSECTICIDES	17	PESTS?	20
35)	MALATHION	4,14,13	INSECTICIDES	10
37)	"ANYTHING"	17	INSECTICIDES	10, C
38)	PYRETHRUM	17	INSECTICIDES	10
39)	PYRETHRUM	17	INSECTICIDES?	10
	DERRIS	17,18	INSECTICIDES?	
	NICOSOAP	17	INSECTICIDES?	
41)	NICOSOAP	18	INSECTICIDES?	
42)	QUOSIA	17,18	INSECTICIDES?	
45)	QUOSIA	14	INSECTICIDES	10
46)	QUOSIA	17	INSECTICIDES	10
47)	INSECTICIDES	7,13,16	?	

49)	MALATHION?	17	INSECTICIDES?	
50)	PESTS	8,16	INSECTICIDE	19, E
51)	PESTS	4,15	INSECTICIDE	7,19
	SOME PESTS	17	INSECTICIDE	7,19
53)	MALATHION	3,14,13	INSECTICIDE	10
54)	CHRYSANTHEMUM	12;3,14,4,15	?	
55)	CHRYSANTHEMUM	1,14	CHRYSANTHEMUMS	10
57)	BLOOM	17;13	CHRYSANTHEMUM	20
58)	BLOOM	17	CHRYSANTHEMUM?	
59)	BLOOM/CHRYSANTH.	9	?	F
61)	CHRYSANTHEMUM	13,8,16	?	
62)	CUTTING	17	CHRYSANTHEMUM	20
63)	CUTTING	17	CHRYSANTHEMUM	20
65)	FIRST CUTTINGS	17	CUTTING	10
67)	FIRST CUTTINGS	4,15	CUTTING?	B
68)	CUTTINGS	3,14	CHRYSANTHEMUMS?	
69)	CUTTINGS	4,15	CHRYSANTHEMUMS?	
	CUTTING	17	CUTTINGS	10
70)	CUTTING	17	CUTTINGS	10
71)	CUTTING	17	CUTTINGS	10
73)	"RIGHT THING"?		CUTTINGS	10,C
74)	"RUNT"?		CUTTINGS?	B,C
75)	"RUNTS"	9	CUTTINGS?	B
76)	CHRYSANTHEMUM	13,3,14	?	
77)	CHRYSANTHEMUM	4,15	?	
78)	CHRYSANTHEMUM	4,15	?	
79)	CHRYSANTHEMUM	1,14	?	
81)	A's CHRYSANTHS	17	CHRYSANTHEMUM	10
82)	PROCEDURE?	12,17	CHRYSANTHEMUM	20
	CHRYSANTHEMUM	3,14	?	
83)	CHRYSANTHEMUM	1,14	?	
85)	CHRYSANTHEMUM	4,15	?	
	CHRYSANTHEMUMS	9,16	?	
86)	CHRYSANTHEMUMS	4,15	?	
87)	CHRYSANTHEMUMS	4,15	?	
89)	CHRYSANTHEMUM	13;8,16	?	
	TOP	17	CHRYSANTHEMUM	20
91)	STOPPING	5,16	CHRYSANTHEMUM?	20, G
93)	STOPPPING	4,15	CHRYSANTHEMUM?	B
	CHRYSANTHEMUM	6,16	?	
94)	CHRYSANTHEMUM	17	?	
96)	CHRYSANTHEMUM	17	?	
97)	CHRYSANTHEMUMS	9	?	
98)	3 CHARMS	13,14	?	
100)	CHARM	17	3 CHARMS	10
101)	CHARM	14	3 CHARMS	10
102)	CHARM	14	3 CHARMS	10
103)	CHARM	14	3 CHARMS	10
104)	CHARM	1,14	?	
105)	CHARM	1,14	?	
	3 CHARMS	8,16	?	
107)	SOME CHARMS	17	CHARMS	10
109)	FLOWERS	13,14,4,15	?	
110)	FLOWER	3,14,4,15	FLOWERS	10
111)	CHARM	13,3,14,4,15	?	

112)	CHARMS	9	?	
113)	CHARMS	4,15	?	
114)	FERTILISER?		CHRYSANTHEMUMS?	
115)	FERTILISER	17	CHRYSANTHEMUMS?	
	A's FERTILISER	17,18,1,14	FERTILISER	10
116)	GRANDFATHER?	4,15	FERTILISER?	B
	SHEEP MANURE	18,2,14	CHRYSANTHEMUMS	8,19
117)	SHEEP MANURE	4,15	CHRYSANTHEMUMS?	B
118)	SHEEP MANURE	17	CHRYSANTHEMUMS?	B
119)	SHEEP MANURE	1,14	FERTILISER	8,10
120)	SHEEP MANURE?		GRANDFATHER?	8,18
121)	"ALL THESE"	13,17	FERTILISER?	10, C
	?		?	H
123)	?		?	H
124)	?		?	H
	THE DAHLIAS	12,14	?	
125)	THE DAHLIAS	17	?	
	DAHLIAS	17	?	
127)	THE DAHLIAS	3,14,4,15	?	
128)	THE DAHLIAS	4,15	?	
129)	THE DAHLIAS	4,15,1,14	?	
	OPEN AREA	17	DAHLIAS 19	
131)	DAHLIAS	7,16	?	
133)	DAHLIAS	4,15	?	
135)	DAHLIAS	4,15	?	
	DAHLIA	13,3,14	DAHLIAS 10, I	
136)	SIDE BUDS	17	BUDS 10	
137)	SIDE BUDS	1,14	BUDS?	
	"THINGS"?		BUDS? 10	

Further Notes

A) Although the rules do not say so definitely, it would be reasonable to assume that it is the plant that is in global focus here, and that it is the overall topic of the whole of this section.

B) A reasonable general principle is that when the local focus stays the same, the global focus does too, unless the rules say otherwise.

C) This is an example of vague lexis where the interpretation seems somehow constrained by the foci.

D) Although an NP is cited as focus, these "what" questions more accurately focus the "what", that is to say that the answer is what moves into focus next.

E) One interpretation would be that the focus here is on blackfly and greenfly, but the meaning seems to be more general than that, so it seems that some process of generalisation, akin to rule 9, yields "pests". Similarly, in line 50, "it" may mean "Malathion", but more probably means "some insecticide".

F) The sense seems ambiguous here, as mentioned in the preliminary analysis.

G) "stopping" may behave rather like a definite NP like "the top" - it elliptically implies the plant.

H) It is not at all clear what goes on in this section, which almost has the character of a prolonged metastatement. It may be that, if we get a clearer idea of the role of such statemnts, and of subordination, this will become clear.

I) See my suggestions about the role of clauses for an alternative analysis of this section.