Discourse-Oriented Anaphora Resolution in Natural Language Understanding: A Review

Graeme Hirst

Department of Computer Science
Brown University
Providence, Rhode Island 02912

Recent research in anaphora resolution has emphasized the effects of discourse structure and cohesion in determining what concepts are available as possible referents, and how discourse cohesion can aid reference resolution. Five approaches, all within this paradigm and yet all distinctly different, are presented, and their strengths and weaknesses evaluated.

1. Introduction

To resolve various forms of definite reference — anaphora in particular — early natural language understanding systems (reviewed in Hirst 1981) typically used a simple kind of history list of concepts previously mentioned in the input, with heuristics for selecting from this list. The history list was usually just a shift register containing the noun phrases from the last sentence or two, and the heuristics would take into account (among other things) selectional restrictions and syntactic constraints on pronominalization. SHRDLU (Winograd 1972) exemplifies this approach. Although able to resolve some types of reference, these systems were not able to handle reference in general, primarily because they did not take into account the effects of discourse structure on reference and pronominalization. This failure motivated work in computational discourse understanding that attempted to exploit discourse structure, especially the relationship between reference and discourse theme, to resolve definite reference.

The present paper[1] is a review of recent work in this area. Five principal approaches are surveyed:

1. Concept activatedness (Kantor) — an examination of the factors affecting the pronominalizability of a concept;

2. Task-oriented dialogues (Grosz) — using a priori knowledge of discourse structure to resolve references;

3. Frames as focus (Sidner) — using discourse cues to choose a frame from a knowledge structure to act as focus;

4. Logical formalism (Webber) — choosing a predicate calculus-like representation to handle problems such as quantification in reference resolution;

5. Discourse cohesion (Hobbs, Lockman, and others) — building a focus and resolving reference by discovering the cohesive ties in a text.

Some preliminary definitions: By **focus** we mean the set containing exactly those concepts available for anaphoric or other definite reference at a point in a text, a set which may conveniently be divided into parts for nominal concepts, temporal concepts, verbal concepts and so forth.[2] The focus is closely related to, but not necessarily identical to, the **theme** of a discourse — what the discourse is about — and since the latter is also sometimes termed *focus*, there is some terminological confusion. (See Section 2.6 and Chapter 4 of Hirst 1981 for further discussion of the distinction between theme and focus.)

Strictly speaking, we mean by the **referent** of an anaphor or reference the real-world entity that it specifies, while by **antecedent** we mean the textual item through which the reference is made. In (1-1):

---

[1] This paper is condensed from a chapter of a longer review of research concerning anaphora and its computational resolution (Hirst 1981).

[2] In this paper we will be concerned mostly with focus for nominal concepts. Temporal, locative and verbal focus are discussed in Hirst (1981).

(1-1) The Queen splutters a little when <u>she</u> speaks.[3]

the antecedent of *she* is the text *The Queen* and the referent is the person who is queen. Generally, however, the two words can be (and are) used interchangeably without confusion.

2. Concept activatedness

Robert Kantor (1977) has investigated the problem of why some pronouns in discourse are more comprehensible than others, even when there is no ambiguity or anomaly. In Kantor's terms, a hard-to-understand pronoun is an example of **inconsiderate** discourse, and speakers (or, more usually, writers) who produce such pronouns lack **secondary [linguistic] competence**. In our terms, an inconsiderate pronoun is one that is not properly in focus.

I will first summarize Kantor's work, and then discuss what we can learn about focus from it.

2.1 Kantor's thesis

Kantor's main exhibit is the following text:

(2-1) A good share of the amazing revival of commerce must be credited to the ease and security of communications within the empire. The Imperial fleet kept the Mediterranean Sea cleared of pirates. In each province, the Roman emperor repaired or constructed a number of skillfully designed roads. <u>They</u> were built for the army but served the merchant class as well. Over them, messengers of the Imperial service, equipped with relays of horses, could average fifty miles a day.

He claims that the *they* in the penultimate sentence is hard to comprehend, and that most informants need to reread the previous text to find its referent. Yet the sentence is neither semantically anomalous nor ambiguous — *the roads* is the only plural NP available as a referent, and it occurs immediately before the pronoun with only a full-stop intervening. To explain this paradox is the task Kantor set himself.

Kantor's explanation is based on discourse topic and the listener's expectations. In (2-1), the discourse topic of the first three sentences is *ease and security of communication in the Roman empire*. In the fourth sentence, there is an improper shift to the roads as the topic: improper, because it is unexpected, and there is no discourse cue to signal it. Had the demonstrative *these roads* been used, the shift would have been okay.

(Note that a definite NP such as *the roads* is not enough.) Alternatively, the writer could have clarified the text by combining the last three sentences with semicolons, indicating that the last two main clauses were to be construed as relating only to the preceding one rather than to the discourse as a whole.

Kantor identifies a continuum of factors affecting the comprehension of pronouns. At one end is **unrestricted expectation** and at the other **negative expectation.** What this says in effect is that a pronoun is easy to understand if its referent is expected, and difficult if it is unexpected. This is not as vacuous as it at first sounds; Kantor provides an analysis of some subtle factors which affect expectation.

The most expected pronominalizations are those whose referent is the discourse topic, or something associated with it (though note the qualifications to this below). Consider:

(2-2) The final years of Henry's reign, as recorded by the admiring Hall, were given over to sport and gaiety, though there was little of the licentiousness that characterized the French court. The athletic contests were serious but very popular. Masques, jousts and spectacles followed one another in endless pageantry. <u>He</u> brought to Greenwich a tremendously vital court life, a central importance in the country's affairs, and above all, a great naval connection.[4]

In the last sentence, *he* is quite comprehensible, despite the distance back to its referent, because the discourse topic in all the sentences is *Henry's reign.* An example of the converse — an unexpected pronoun which is difficult despite recency — can be seen in (2-1) above. Between these two extremes are other cases involving references to aspects of the local topic, changes in topic, syntactic parallelism, and, in topicless instances, recency (though the effect of recency decays very fast). I will not describe these here; the interested reader is referred to Section 2.6.5 of Kantor's dissertation (1977).

Kantor then defines the notion of the **activatedness** of a concept. This provides a continuum of concept givenness, which contrasts with the simple binary given-new distinction usually accepted in linguistics (for example, Chafe 1970). Kantor also distinguishes activatedness from the similar "communicative dynamism" of the Prague school (Firbas 1964). Activated-

---

[3] Underlining is used in this and subsequent examples to indicate the anaphor(s) of interest. It does not indicate stress.

[4] From: Hamilton, Olive and Hamilton, Nigel. *Royal Greenwich.* Greenwich: The Greenwich Bookshop, 1969. Quoted by Halliday and Hasan (1976:14), quoted by Kantor (1977).

ness is defined in terms of the comprehensibility phenomena described above: the more activated a concept is, the easier it is to understand an anaphoric reference to it. Thus activatedness depends upon discourse topic, context, and so forth.

2.2 The implications of Kantor's work

What are the ramifications of Kantor's thesis for focus? Clearly, the notions of activatedness and focus are very similar, though the latter has not generally been thought of as a continuum. It follows that the factors Kantor finds relevant for activatedness and comprehensibility of pronouns are also important for those of us who would maintain focus in computer-based natural language understanding (NLU) systems; we will have to discover discourse topic and topic shifts, generate pronominalization expectations, and so forth.

In other words, if we could dynamically compute (and maintain) the activatedness of each concept floating around, we would have a measure for the ordering of the focus set by preferability as referent; the referent for any given anaphor would be the most highly activated element which passes basic tests for number, gender and semantic reasonableness. And to find the activatedness of the concepts, we follow Kantor's pointers (which he himself concedes are very tenuous and difficult) to extract and identify the relevant factors from the text.

It may be objected that by applying Kantor's insights all we have done is produce a mere notational variant of our original problem. This is partly true. One should not gainsay the power of a good notation, however, and what we can buy here even with mere notational variance is the power of Kantor's investigations. And there is more. Previously, it has been suggested that items either are in focus or they aren't, and that at each separate anaphor we need to compute a preference ranking of the focus elements for that anaphor. What Kantor tells us is that such a ranking exists independently of the actual use of anaphors in the text, and that we can find the ranking by looking at things like discourse topic.

Some miscellaneous comments on Kantor's work:

1. It can be seen as a generalization albeit a weakening of Grosz's (1977a, 1977b, 1978) findings on focus in task-oriented dialogues (where each sub-task becomes the new discourse topic, opening up a new set of possible referents), which are discussed below in Section 3. (Kantor and Grosz were apparently unaware of each other's work; neither cites the other.)

2. It provides an explanation for focus problems that have previously baffled us. For example, in Hirst (1977a) I contemplated the problem of the ill-formedness of this text:

(2-3) *John left the window and drank the wine on the table. <u>It</u> was brown and round.

I had previously thought this to be due to a syntactic factor — that cross-sentence pronominal reference to an NP in a relative clause or adjectival phrase qualifying an NP was not possible. However, it can also be explained as a grossly inconsiderate pronoun which does not refer to the topic properly — *the table* occurs only as a descriptor for the wine, and not as a concept "in its own right". This would be a major restriction on possible reference to sub-aspects of topics.

3. Like too many other researchers, Kantor makes many claims about comprehensibility and the degree of well-formedness of sentences which others (as he concedes) may not agree with. He uses only himself (and his friends, sometimes) as an informant, and then only at an intuitive level.[5] Claims as strong and subtle as Kantor's cry out for empirical testing.[6]

3. Focus of attention in task-oriented dialogues

Barbara Grosz (1977a, 1977b, 1978) studied the maintenance of the focus of attention in task-oriented dialogues and its effect on the resolution of definite reference, as part of SRI's speech understanding system project (Walker 1978). By a **task-oriented** dialogue is meant one which has some single major well-defined task as its goal. For example, Grosz collected and studied dialogues in which an expert guides an apprentice in the assembly of an air compressor. She found that the structure of such dialogues parallels the structure of the task. That is, just as the major task is divided into several well-defined sub-tasks, and these perhaps into sub-sub-tasks and so on, the dialogue is likewise divided into sub-dialogues, sub-sub-dialogues, etc,[7] each corresponding to a task component, much as a well-structured Algol program is composed of blocks within blocks within blocks. As the dialogue progresses, each sub-dialogue in turn is performed in a strict depth-first order corresponding to the order of sub-task performance in the task goal (though note that some sub-tasks may not be ordered with respect to

---

[5] For a discussion of the problem of idiosyncratic well-formedness judgments, and a suggested solution, see Sections 4.2 and 7.3 of Hirst (1981).

[6] Kantor tells me that he hopes to test some of his assertions by observing the eye movements of readers of considerate and inconsiderate texts, to find out if inconsiderate texts actually make readers physically search back for a referent.

[7] Below I will use the prefix *sub-* generically to include *sub-sub-sub-* . . . to an indefinite level.

others). As we will see, this dialogue structure can be exploited in reference resolution.

Grosz's aim was to find ways of determining and representing the **focus of attention** of a discourse — that is, roughly speaking, its global theme and the things associated therewith — as a means for constraining the knowledge an NLU system needs to bring to bear in understanding discourse. In other words, the focus of attention is that knowledge which is relevant at a given point in a text for comprehension of the text.[8] Grosz claims that antecedents for definite reference can be found in the focus of attention. That is, the focus of attention is a superset of focus in our sense, the set of referable concepts (in this case definite reference, not just anaphoric reference). Moreover, no element in the focus of attention is excluded from being a candidate antecedent for a definite NP. Grosz thereby implies that all items in the focus of attention can be referred to, and that hence the two senses of the word *focus* are actually identical.

3.1 Representing and searching focus

In Grosz's representation, which uses a partitioned semantic net formalism (Hendrix 1975, 1978), an **explicit focus** corresponds to a sub-dialogue, and includes, for each concept in it, type information about that concept and any situation in which that concept participates. For each item in the explicit focus, there is an associated **implicit focus**, which includes subparts of objects in explicit focus, subevents of events in explicit focus, and participants in those subevents. The implicit focus attempts to account for reference to items that have a close semantic distance to items in focus, or which have a close enough relationship to items in focus to be able to be referred to. The implicit focus is also used in detecting focus shifts (discussed below).

Then, at any given point in a text, antecedents of definite non-pronominal NPs can be found by searching through the explicit and implicit focus for a match for the reference. After checking the other non-pronominal NPs in the same sentence to see if the reference is intrasentential, the **currently active** explicit focus (the focus corresponding to the present sub-dialogue) is searched, and then if that search is not successful, the other currently open focus spaces (that is, those corresponding to sub-dialogues that the present sub-dialogue is contained in) are searched in order, back up to the top of the tree. As part of the search the implicit focus associated with each explicit focus is checked, as are subset relations, so that if a novel, say,

is in focus, it could be referred to as *the book*. If there is still no success after this, one then checks whether the NP refers to a single unique concept (such as the sun), contains new information (such as *the red coat*, when a coat is in focus, but not yet known to be red), or refers to an item in implicit focus.

A similar search method could be used for pronouns. However, since pronouns carry much less information than other definite NPs, more inference is required by the reference matching process to disambiguate many syntactically ambiguous pronouns, and it would be necessary to search focus exhaustively, comparing the reasonableness of candidate referents, rather than stopping at the first plausible one. In addition, other constraints on pronoun reference, such as local (rather than global) theme, and default referent, would also need to be taken into account; Grosz's mechanisms do not do this. However, Grosz does show how a partitioned network structure can be used to resolve certain types of ellipsis by means of syntactic and semantic pattern matching against the immediately preceding utterance, which may itself have been expanded from an elliptical expression. She leaves open for future research most of the problems in relating pronouns to focus.

3.2 Maintaining focus

Given this approach, one is then faced with the problem of deciding what the focus is at a given point in the discourse. For highly constrained task-oriented dialogues such as those Grosz considered, the question of an initial focus does not arise; it is, by definition, the overall task in question. The other component of the problem, handling changes and shifts in the focus, is attacked by Grosz in a top-down manner using the task structure as a guide.

A shift in focus can be indicated explicitly by an utterance, such as:

(3-1) Well, the reciprocating afterburner nozzle
        speed control is assembled. Next, it must
        be fitted above the preburner swivel hose
        cover guard cooling fin mounting rack.

In this case, the reciprocating afterburner nozzle speed control assembly sub-task and its corresponding sub-dialogue and focus are closed, and new ones are opened for the reciprocating afterburner nozzle speed control fitting, dominated by the same open sub-tasks/sub-dialogues/focuses in their respective trees that dominated the old ones. If however the new sub-task were a sub-task of the old one, then the old one would not be closed, but the new one added to the hierarchy below it as the new active focus space. The newly created focus space initially contains only those

---

[8] In her later work (Grosz 1978), Grosz emphasizes focusing as an active process carried out by dialogue participants.

items referred to in the utterance, and those objects associated with the current sub-task. (Being **able** to bring in the associated objects at this time is, of course, the crucial point on which the whole system relies.) As subsequent non-shift-causing utterances come in, their new information is added to the active focus space.

Usually, of course, speakers are not as helpful as in (3-1), and it is necessary to look for various clues to shifts in focus. For Grosz, the clues are definite NPs. If a definite NP from an utterance cannot be matched in focus, then this is a clue that the focus has shifted, and it is necessary to search for the new focus. If the antecedent of a definite NP is in the current implicit focus, this is a clue that a sub-task associated with this item is being opened. If the task structure is being followed, then the new focus will reflect the opening or closing of a sub-task.

Shifting cannot be done until a whole utterance is considered, because clues may conflict, or the meaning of the utterance may contraindicate the posited shift. In particular, recall that the task structure is only a guide, and does not define the dialogue structure absolutely. For example, the focus may shift to a problem associated with the current sub-task with a question like this:

(3-2) Should I use the box-end ratchet wrench
        to do that?

This does not imply a shift to the next sub-task requiring a box-end ratchet wrench (assuming that the current task doesn't require one) (cf Grosz 1977b:105).

We can see here that the problem of the circularity of language comprehension looms dangerously: to determine the focus one must resolve the references, and to resolve the references, one must know the focus. In Grosz's work, the strong constraints of the structure of task-oriented dialogues provide a toehold. Whether generalization to the case of discourse with other structures, or with no particular structure, is possible is unclear, as it may not be possible to determine so nicely what the knowledge associated with any new focus is. (See however my remarks in Section 2.2 above on the relationship between Grosz's work and that of Kantor, and Section 6 on approaches which attempt to exploit local discourse structure.)

In addition, Grosz's mechanisms are limited in their ability to resolve anaphora that require inference or are intersentential (or both). The assumption that global focus of attention equals all and only possible referents (except where the focus shifts), while perhaps not unreasonable in task-oriented domains, is probably untrue in general. For example, it is unclear that such mechanisms could handle the effects of local as opposed to global theme that exclude *the table* from

the focus for almost all speakers in (2-3). Similarly, could the level of world knowledge and inference required to resolve the different referents of *she* in (3-3) and (3-4) be integrated into the partitioned semantic net formalism?

(3-3) When Nadia visited Sue for dinner,
        she ate sukiyaki au gratin.

(3-4) When Nadia visited Sue for dinner,
        she served sukiyaki au gratin.

Could entities evoked by, but not explicit in, a text of only moderate structure be identified and instantiated in focus? Grosz did not address these issues (nor did she need to for her immediate goals), but they would need to be resolved in any attempt to generalize her approach. (Some other related problems, including those of focus shifting, are discussed in Grosz 1978.)

Grosz's contribution was to demonstrate the role of discourse structure in the identification of theme, relevant world knowledge and the resolution of reference. We now turn to another system which aspires to similar goals, but in a more general context.

4. Focus in the PAL system

The PAL personal assistant program (Bullwinkle 1977a) is a system designed to accept natural language requests for scheduling activities. A typical request (from Bullwinkle 1977b:44) is:

(4-1) I want to schedule a meeting with Ira. It
        should be at 3 pm tomorrow. We can
        meet in Bruce's office.

The section of PAL that deals with discourse pragmatics and reference was developed by Candace Sidner [Bullwinkle] (Bullwinkle 1977b; Sidner 1978a). Like Grosz's system, PAL attempts to find a focus of attention in its knowledge structures to use as a focus for reference resolution. Sidner sees the focus as equivalent to the discourse topic; in fact in Bullwinkle (1977b) the word *topic* is used instead of *focus*.

There are three major differences from Grosz's system:

1. PAL does not rely heavily on discourse structures.
2. Knowledge is represented in frames.
3. Focus selection and shifting are handled at a more superficial level.

I will discuss each difference in turn.

4.1 PAL's approach to discourse

Because a request to PAL need not have the rigid structure of one of Grosz's task-oriented dialogues, PAL does not use discourse structure to the same extent, instead relying on more general local cues. However, as we shall see below, in focus selection and

shifting, Sidner was forced to use ad hoc rules based on observations of typical requests to PAL.

4.2 The frame as focus

The representation of knowledge in PAL is based on frames, and its implementation uses the FRL frame representation language (actually a dialect of LISP) developed by Roberts and Goldstein (1977a, 1977b).

In PAL, the frame corresponds to Grosz's focus space. Following Rosenberg's (1976, 1977) work on discourse structure and frames, the antecedent for a definite NP is first assumed to be either the frame itself, or one of its slots. So, for example, in (4-2):

(4-2) I want to have a meeting with Ross[1]. It should be at three pm. The location will be the department lounge. Please tell Ross[2].

it refers to the MEETING frame (**not** to the text *a meeting*) which provides the context for the whole discourse; *the location* refers to the LOCATION slot that the MEETING frame presumably has (thus the CLOSELY ASSOCIATED WITH relation (Hirst 1981) is handled), and *Ross[2]* to the contents[9] of the CO-MEETER slot, previously given as Ross.

If the antecedent cannot be found in the frame, it is assumed to be either outside the discourse or inferred. In (4-2), PAL would search its database to find referents for *Ross[1]* and *the department lounge*. Personal names are resolved with a special module that knows about the semantics of names (Bullwinkle 1977b:48). PAL carries out database searches for references like *the department lounge* apparently by searching a hierarchy of frames, looking at the frames in the slots of the current focus, and then in the slots of these frames, and so on (Sidner 1978a:211), though it is not apparent why this should usefully constrain the search in the above example.[10]

---

[9] Sidner only speaks of reference to slots (1978a:211), without saying whether she means the slot itself or its contents; it seems reasonable to assume, as I have done here, that she actually means both.

[10] In fact there is no need in this particular example for a referent at all. The personal assistant need only treat *the department lounge* as a piece of text, presumably meaningful to both the speaker and Ross, denoting the meeting location. A human might do this when passing on a message he or she didn't understand:

(i) Ross asked me to tell you to meet him in the arboretum, whatever the heck that is.

On the other hand, an explicit antecedent **would** be needed if PAL had been asked, say, to deliver coffee to the meeting in the department lounge. Knowing when to be satisfied with ignorance is a difficult problem which Sidner does not consider, preferring the safe course of always requiring an antecedent.

4.3 Focus selection

In PAL, the initial focus is the first NP following the main verb of the first sentence of the discourse — usually, the object of the sentence — or, if there is no such NP, then the subject of that sentence. This is a short-cut method, which seems to be sufficient for requests to PAL, but which Sidner readily admits is inadequate for the general case (Sidner 1978a:209). I will briefly review some of the problems.

Charniak (1978) has shown that the frame-selection problem (which is here identical to the initial focus selection problem, since the focus is just the frame representing the theme of the discourse) is in fact extremely difficult, and is not in the most general case amenable to solution by either strictly top-down or bottom-up methods. Sidner's assumption that the relevant frame is given by an explicitly mentioned NP is also a source of trouble, even in the examples she quotes, such as these two (Sidner 1978b:92):

(4-3) I was driving along the freeway the other day. Suddenly the engine began to make a funny noise.

(4-4) I went to a new restaurant with Sam. The waitress was nasty. The food was great.

(Underlining indicates what Sidner claims is the focus.) In (4-3), Sidner posits a chain of inferences to get from *the engine* to the focus, the FREEWAY frame. This is more complex than is necessary; if the frame/focus were DRIVING (with its LOCATION slot containing the FREEWAY frame), then the path from the frame to *the engine* is shorter and the whole arrangement seems more natural. Thus we see that focus need not be based on an NP at all.

In (4-4), our problem is what to do with Sam, who could be referenced in a subsequent sentence. It is necessary to integrate Sam into the RESTAURANT frame/focus, since clearly he should not be considered external to the discourse and sought in the database. While the RESTAURANT frame may indeed contain a COMPANION slot for Sam to sit in, it is clear that the first sentence could have been *I went <anywhere at all> with Sam*, requiring that any frame referring to something occupying a location must have a COMPANION slot. This is clearly undesirable. But the RESTAURANT frame **is** involved in (4-4); otherwise *the waitress* and *the food* would be external to the discourse. A natural solution is that the frame/focus of (4-4) is actually the GOING-SOMEWHERE frame (with Sam in its COMPANION slot), containing the RESTAURANT frame in its PLACE slot, with both frames together taken as the focus. Sidner does not consider mechanisms for a multi-frame focus.

It is, of course, not always true that the frame/focus is explicit. Charniak (1978) points out that (4-5) is somehow sufficient to invoke the MAGICIAN frame:

(4-5) The woman waved as the man on stage
        sawed her in half.

(See also Charniak (1981) for more on frame invocation problems.)

Focus shifting in PAL is restricted: the only shifts permitted are to and from sub-aspects of the present focus (Sidner 1978a:209). Old topics are stacked for possible later return. This is very similar to Grosz's open-focus hierarchy. It is unclear whether there is a predictive aspect to PAL's focus-shift mechanism,[11] but the basic idea seems to be that any new phrase in a sentence is picked as a potential new focus. If in a subsequent sentence an anaphoric reference is a semantically acceptable coreferent for that potential focus, then a shift to that focus is ipso facto indicated (Sidner 1978a:209). Presumably this check is done after a check of focus has failed, but before any database search. A potential focus has a limited life span, and is dropped if not shifted to by the end of the second sentence following the one in which it occurred.

An example (Sidner 1978a:209):

(4-6) I want to schedule a meeting with George, Jim, Steve and Mike. We can meet in my office. It's kind of small, but the meeting won't last long anyway.

(4-7) I want to schedule a meeting with George, Jim, Steve and Mike. We can meet in my office. It won't take more than 20 minutes.

In the second sentence *my office* is identified as a potential focus, and *it*, in the first reading of the third sentence, as an acceptable coreferent to *my office* confirms the shift. In the second reading, *it* couldn't be *my office*, so no shift occurs. The acceptability decision is based on selectional and case-like restrictions.

While perhaps adequate for PAL, this mechanism is, of course, not sufficient for the general case, where a true shift, as opposed to an expansion upon a previ-

ously mentioned point, may occur. This is exemplified by many of the shifts in Grosz's task-oriented dialogues.

Another problem arising from this shift mechanism is that two different focus shifts may be indicated at the same time, but the mechanism has no way to choose between them. For example:

(4-8) Schedule a meeting of the Experimental
        Theology Research Group, and tell Ross
        Andrews about it too. I'd like him to
        hear about the deocommunication work
        that they're doing.

Each of the two underlined NPs in the first sentence would be picked as a potential focus. Since each is pronominally referenced in the second sentence, the mechanism would be confused as to where to shift the focus. (Presumably *Ross Andrews* would be the correct choice here.)

4.4 Conclusions

The shortcomings of Sidner's work are mainly attributable to two causes: her avoidance of relying on the highly constrained discourse structures that Grosz used, and the limited connectivity of frame systems, compared to Grosz's semantic nets.[12] With respect to the former point, perhaps Sidner's main contribution has been to show the difficulties and pitfalls that lie in wait for anyone attempting to generalize Grosz's work, even to the extent that PAL does.

5. Webber's formalism

In the preceding sections of this paper, we saw approaches to anaphor resolution that were mainly top-down in that they relied on a notion of theme and/or focus of attention to guide the selection of focus (although theme determination may have been bottom-up). An alternative approach has been suggested by Bonnie [Nash-]Webber (Nash-Webber and Reiter 1977; Webber 1978a, 1978b), wherein a set of rules is applied to a logical-form representation of the text to derive the set of entities that that text makes available for subsequent reference. Webber's formalism attacks some problems caused by quantification that have not otherwise been considered by workers in NLU.

---

[11] On page 209 of Sidner (1978a) we are told: "Focus shifts cannot be predicted; they are detectable only after they occur". Yet on the following page, Sidner says: "Sentences appearing in mid-discourse are assumed to be about the focus until the coreference module predicts a focus shift . . . . Once an implicit focus relation is established, the module can go onto [sic] predictions of focus shift". My interpretation of these remarks is that one cannot be certain that the next sentence will shift focus, but one can note when a shift might happen, requiring later checking to confirm or disconfirm the shift.

[12] In her thesis (1979) [which was not available to me when this paper was first written], Sidner subsequently proposed the use of an association network instead of frames, and presented more sophisticated focus selection and shifting algorithms. I have emphasized her earlier work here, as it has received much wider circulation.

I can only give the flavor of Webber's formalism here, and I shall have to assume some familiarity with logical forms. Readers who want more details should see her thesis (1978a); readers who find my exposition mystifying should not worry unduly — the fault is probably mine — but should turn to the thesis for illumination.

In Webber's formalism, it is assumed that an input sentence is first converted to a parse tree, and then, by some semantic interpretation process, to an **extended restricted-quantification predicate calculus representation.** It is during this second conversion that anaphor resolution takes place. When the final representation, which we shall simply call a **logical form**, is complete, certain rules are applied to it to generate the set of referable entities and descriptions that the sentence evokes. Webber considers three types of antecedents — those for definite pronouns, those for *one*-anaphora,[13] and those for verb phrase ellipsis. Each type has its own set of rules; we will briefly look at the first. (The others are discussed in Sections 5.4.2 and 5.4.3 of Hirst 1981.)

5.1 Definite pronouns

The antecedents for definite pronouns are **invoking descriptions** (IDs); these are in effect focus elements that are explicit in the text. IDs are derived from the logical form representation of a sentence by a set of rules that attempt to take into account factors, such as NP definiteness or references to sets, that affect what antecedents are evoked by a text. There are six of these ID-rules;[14] which one applies depends on the structural description of the logical form.

Here is one of Webber's examples (1978a:64):

(5-1) Wendy bought a crayon.

This has this representation:

(5-2) $(\exists x{:}\text{Crayon})$. Bought Wendy,x

Now, one of the ID-rules says that any sentence S whose representation is of this form:

(5-3) $(\exists x{:}C)$. Fx

where C is an arbitrary predicate on individuals and Fx an arbitrary open sentence in which x is free, evokes an entity whose representation is of this form:

(5-4) e_j $\iota x{:}$ Cx & Fx & evoke S,x

where e_j is an arbitrary label assigned to the entity and ι is the definite operator. Hence, starting at the left of (5-2), we obtain this representation for the crayon of (5-1):

(5-5) e_1 $\iota x{:}$ Crayon x & Bought Wendy,x & evoke (5-1),x

which may be interpreted as e_1 *is the crayon mentioned in sentence (5-1) that Wendy bought.* Similarly we will obtain a representation of e_2, Wendy, which is then substituted for *Wendy* in (5-5) after some matching process has determined the identity of the two.

In this next, more complex example (Webber 1978a:73), we see how quantification is handled:

(5-6) Each boy gave each girl a peach.
$(\forall x{:}\text{Boy})$ $(\forall y{:}\text{Girl})$ $(\exists z{:}\text{Peach})$. Gave x,y,z

This matches the following structural description (where Q_j stands for the quantifier $(\forall x_j \in e_j)$, where e_j is an earlier evoked discourse entity, and ! is the left boundary of a clause):

(5-7) $!Q_1 \ldots Q_n$ $(\exists y{:}C)$. Fy

and hence evokes an ID of this form:

(5-8) e_i $\iota y{:}$ maxset$(\lambda(u{:}C)[(\exists x_1 \in e_1)$ $\ldots$
$(\exists x_n \in e_n)$. Fu & evoke S,u]$)$ y

(For any one-place predicate P, maxset$(P)y$ is true if and only if y is the set of all items u such that Pu holds.) Another rule has already given us:

(5-9) e_1 $\iota x{:}$ maxset(Boy) x
"the set of all boys"
e_2 $\iota x{:}$ maxset(Girl) x
"the set of all girls"

and so (5-8) is instantiated as:

(5-10) e_3 $\iota z{:}$ maxset$(\lambda(u{:}\text{Peach})$ $[(\exists x \in e_1)$ $(\exists y \in e_2)$. Gave x,y,u & evoke (5-6),y]$)$ z
"the set of peaches, each one of which is linked to (5-6) by virtue of some member of e_1 giving it to some member of e_2"

Although such rules could (in principle) be used to generate all IDs (explicit focus elements) that a sentence evokes, Webber does not commit herself to such an approach, instead allowing for the possibility of generating IDs only when they are needed, depending on subsequent information such as speaker's perspective. She also suggests the possibility of "vague, temporary" IDs for interim use (1978a:67).

There is a problem here with intrasentential anaphora, since it is assumed that a sentence's anaphors are resolved before ID rules are applied to find what may be the antecedents necessary for that resolution. Webber proposes that known syntactic and selectional

---

[13] *One-anaphors* are those such as *those, one*, and some uses of *it* that refer to a description rather than a specific entity. An example:

(i) Wendy didn't give either boy a green tie-dyed T-shirt, but she gave Sue a red <u>one.</u>

[14] Webber regards her rules only as a preliminary step towards a complete set that considers all relevant factors. She discusses some of the remaining problems, such as negation, in Webber (1978a:81-88).

constraints may help in this conflict, but this is not always sufficient. For example:

(5-11) Mary bought each girl a cotton T-shirt,
       but none of <u>them</u> were the style de rigeur
       in high schools.

The IDs for both the set of girls and the set of T-shirts are needed to resolve *them*, but *them* needs to be resolved before the IDs are generated. In this particular example, the clear solution is to work a clause at a time rather than at a sentence level. However, this is not always an adequate solution, as (5-12) shows:

(5-12) The rebel students annoyed the teachers
       greatly, and by the end of the week none
       of the faculty were willing to go to <u>their</u>
       classes.

In this ambiguous sentence, one possible antecedent for *their*, *the faculty*, occurs in the same clause as the anaphor. Thus neither strictly intraclausal nor strictly interclausal methods are appropriate. Webber is aware of this problem (1978a:48), and believes that it suffices that such information as is available be used to rule out impossible choices; the use of vague temporary IDs then allows the anaphor to be resolved.

5.2 Conclusions

It remains to discuss the strengths and weaknesses of Webber's approach, and she herself (in contradistinction to some other workers) is as quick to point out the latter as the former. The reader is therefore referred to her thesis (1978a) for this. However, I will make some global comments on the important aspects relevant here.

Webber's main contributions, as I see them, are as follows:

1. The anaphor resolution problem is approached from the point of view of determining what an adequate representation would be, rather than trying to fit (to straitjacket?) a resolution mechanism into some pre-existing and perhaps arbitrarily chosen representation; and the criteria of adequacy for the representation are rigorously enumerated.

2. A formalism in which it is possible to compute focus elements as they are needed, rather than having them sitting round in advance (as in Grosz's system), perhaps never to be used, is provided (but compare my further remarks below).

3. Webber brings to NLU anaphora research the formality and rigor of logic, something that has been previously almost unseen.

4. Previously ignored problems of quantification are dealt with.

5. The formalism itself is an important contribution.

The shortcomings, as I see them, are as follows:

1. The formalism relies very much on antecedents being in the text. Entities evoked by, but not explicit in, the text cannot in general be adequately handled (in contrast to Grosz's system).

2. The formalism is not related to discourse structure. So, for example, it contains nothing to discourage the use of *the table* as the antecedent in (2-3). It remains to be seen if discourse pragmatics can be adequately integrated with the formalism or otherwise accounted for in a system using the formalism.

3. Intrasentential and intraclausal anaphora are not adequately dealt with.

4. Webber does not relate her discussions of representational adequacy to currently popular knowledge representations. If frames, for example, are truly inadequate we would like to have some watertight proof of this before abandoning current NLU projects attempting to use frames.

It will be noticed that contribution 2 and shortcoming 1 are actually two sides of the same coin — it is static pre-available knowledge that allows non-textual entities to be easily found — and clearly a synthesis will be necessary here.

6. Discourse-cohesion approaches to anaphora resolution

Another approach to coreference resolution attempts to exploit local discourse cohesion, building a representation of the discourse with which references can be resolved. This approach has been taken by (inter alia) Klappholz and Lockman (1977; Lockman 1978). By using only cues to the discourse structure at the sentence level or lower, one avoids the need to search for referents in pre-determined dialogue models such as those of Grosz's task-oriented dialogues, or rigidly predefined knowledge structures such as scripts (Schank and Abelson 1977) and frames (Minsky 1975), which Klappholz and Lockman, for example, call overweight structures that inflexibly dominate processing of text. Klappholz and Lockman emphasize that the structure through which reference is resolved must be dynamically built up as the text is processed; frames or scripts could assist in this building, but cannot, however, be reliably used for refer-

ence resolution, because deviations by the text from the pre-defined structure will cause errors.

The basis of this approach is that there is a strong interrelationship between coreference and the cohesive ties in a discourse that make it coherent. By determining what the cohesive ties in a discourse are, one can put each new sentence or clause, as it comes in, into the appropriate place in a growing structure that represents the discourse. This structure can then be used as a focus to search for coreference antecedents, since not only do coherently connected sentences tend to refer to the same things, but knowledge of the cohesion relation can provide additional reference resolution restraints. Hobbs (1979) in particular sees the problem of coreference resolution as being automatically solved in the process of discovering the coherence relations in a text. (An example of this will be given in Section 6.2.) Conversely, it is frequently helpful or necessary to resolve coreference relations in order to discover the coherence relations. This is not a vicious circle, claims Hobbs, but a spiral staircase.

In our discussion below, we will cover four issues:

1. deciding on a set of possible coherence relations;

2. detecting them when they occur in a text;

3. using the coherence relations to build a focus structure; and

4. searching for referents in the structure.

6.1 Coherence relations

The first thing required by this approach is a complete and computable set of the coherence relations that may obtain between sentences and/or clauses. Various sets have been suggested by many people, including Eisenstadt (1976), Phillips (1977), Pitkin (1977a, 1977b), Hirst (1977b, 1978), Lockman (1978), Hobbs (1978, 1979) and Reichman (1978).[15] None of these sets fulfill all desiderata; and while Halliday and Hasan (1976) provide an extensive analysis of cohesion, it does not fit within our computational framework of coherence relations, and those, such as Hobbs, Lockman, Eisenstadt and Hirst, who emphasize computability, provide sets insufficient, I believe, to capture all the semantic subtleties of discourse cohesion. Nevertheless, the works cited above undoubtedly serve as a useful starting point for development of this area.

To illustrate what a very preliminary set of cohesion relations could look like, I will briefly present a set abstracted from the various sets of Eisenstadt, Hirst, Hobbs, Lockman and Phillips (but not faithful to any one of these).

The set contains two basic classes of coherence relations: expansion or elaboration on an entity, concept or event in the discourse, and temporal continuation or time flow. Expansion includes relations like EFFECT, CAUSE, SYLLOGISM, ELABORATION, CONTRAST, PARALLEL and EXEMPLIFICATION. In the following examples, "■" is used to indicate the point where the cohesive tie illustrated is acting:

(6-1) [ELABORATION] To gain access to the latch-housing, remove the control panel cover. ■ Undo both screws and rock it gently until it snaps out from the mounting bracket.

(6-2) [CONTRAST] The hoary marmot likes to be scratched behind the ears by its mate, ■ while in the lesser dormouse, nuzzling is the primary behavior promoting pair-bonding.

(6-3) [EFFECT] Ross pulled out the bottom module. ■ The entire structure collapsed.

(6-4) [CAUSE] Ross scratched his head furiously. ■ The new Hoary Marmot™ shampoo that he used had made it itch unbearably.

(6-5) [SYLLOGISM] Nadia goes to the movies with Ross on Fridays. Today's Friday, ■ so I guess she'll be going to the movies.

(6-6) [PARALLEL] Nearly all our best men are dead! Carlyle, Tennyson, Browning, George Eliot! — ■ I'm not feeling very well myself![16]

(6-7) [EXEMPLIFICATION] Many of our staff are keen amateur ornithologists. ■ Nadia has written a book on the Canadian triller, and Daryel once missed a board meeting because he was high up a tree near Gundaroo, watching the hatching of some rare red-crested snipes.

(One may disagree with my classification of some of the relations above; the boundaries between categories are yet ill-defined, and it is to be expected that some people's intuitions will differ from mine.)

---

[15] Reichman's coherence relations operate at paragraph level rather than sentence or clause level.

[16] From: A lament [cartoon caption]. *Punch, or the London charivari*, CIV, 1893, page 210.

Temporal flow relations involve some continuation forwards or backwards over time:

(6-8) VICTORIA — A suntanned Prince Charles arrived here Sunday afternoon, ■ and was greeted with a big kiss by a pretty English au pair girl.[17]

(6-9) SAN JUAN, Puerto Rico — Travel officials tackled a major job here Sunday to find new accommodations for 650 passengers from the burned Italian cruise liner *Angelina Lauro.*
　　　　■ The vessel caught fire Friday while docked at Charlotte Amalie in the Virgin Islands, but most passengers were ashore at the time.[18]

Temporal flow may be treated as a single relation, as Phillips, for example, does, or it may be subdivided, as by Eisenstadt and Hirst, into categories like TIME STEP, FLASHBACK, FLASHFORWARD, TIME EDIT, and so on. Certainly, time flow in a text may be quite contorted, as in (6-10) (from Hirst 1978); "■" indicates a point where the direction of the time flow changes:

(6-10) Slowly, hesitantly, Ross approached Nadia. ■ He had waited for this moment for many days. ■ Now he was going to say the words ■ which he had agonized over ■ and in the very room ■ he had often dreamed about. ■ He gazed lovingly at her soft green eyes.

It is not clear, however, to what extent an analysis of time flow is necessary for anaphor resolution. I suspect that relatively little is necessary — less than is required for other aspects of discourse understanding.

I see relations like those exemplified above as **primitives** from which more complex relations could be built. For example, the relation between the two sentences of (6-3) above clearly involves FORWARD TIME STEP as well as EFFECT. I have hypothesized elsewhere (Hirst 1978) the possibility of constructing a small set of discourse relations (with cardinality about twenty or less) from which more complex relations may be built up by simple combination, and, one hopes, in such a way that the effects of relation $R_1 + R_2$ would be the sum of the individual effects of relations R_1 and R_2. Rules for permitted combinations would be needed; for example, FORWARD TIME STEP could combine with EFFECT, but not with BACKWARD TIME STEP.

---

[17] From: *The Vancouver express*, 2 April 1979, page A1.

[18] From: *The Vancouver express*, 2 April 1979, page A5.

What would the formal definition of a coherence relation be like? Here is Hobbs's (1979:73) definition of ELABORATION: Sentence S_1 is an ELABORATION of sentence S_0 if some proposition P follows from the assertions of both S_0 and S_1, but S_1 contains a property of one of the elements of P that is not in S_0. The example in the next section will clarify this.

6.2 An example of anaphor resolution using a coherence relation

It is appropriate at this stage to give an example of the use of coherence relations in the resolution of anaphors. I will present an outline of one of Hobbs's; for the fine details I have omitted, see Hobbs (1979:78-80). The text is this:

(6-11) John can open Bill's safe. <u>He</u> knows <u>the combination</u>.

We want an NLU system to recognize the cohesion relation operating here, namely ELABORATION, and identify *he* as John and *the combination* as that of Bill's safe. We assume that in the world knowledge that the system has are various axioms and rules of inference dealing with such matters as what combinations of safes are and knowledge about doing things. Then, from the first sentence of (6-11), which we represent as (6-12):

(6-12) can (John, open (Bill's-safe))

(we omit the details of the representation of *Bill's safe*), we can infer:

(6-13) know (John, cause (do (John, *a*),
　　　　　　open (Bill's-safe)))
　　"John knows that he can perform an action *a* that will cause Bill's-safe to be open"

From the second sentence of (6-11), namely:

(6-14) know (*he*, combination (*comb, y*))
　　"someone, *he*, knows the combination *comb* to something, *y*"

we can infer, using knowledge about combinations:

(6-15) know (*he*, cause (dial (*comb,y*), open (*y*)))
　　"*he* knows that by causing the dialing of *comb* on *y*, the state in which *y* is open will be brought about"

Recognizing that (6-13) and (6-15) are nearly identical, and assuming that some coherence relation does hold, we can identify *he* with *John*, *y* with *Bill's-safe*, and the definition of the ELABORATION relation is satisfied. In the process, the required referents were found.

6.3 Lockman's contextual reference resolution algorithm

Given a set of discourse cohesion relations, how may their use in a text be computationally recognized and employed to build a structure that represents the discourse and can be used as a focus for reference resolution? Only Hobbs (1978, 1979) and Lockman (1978; Klappholz and Lockman 1977) seem to have considered these aspects of the problem, though Eisenstadt (1976) discusses some of the requirements in world knowledge and inference that would be required. In this section we look at Lockman's work.

Lockman does not separate the three processes of recognizing cohesion, resolving references and building the representation of the discourse. Rather, as befits such interrelated processes, all three are carried out at the same time. His contextual reference resolution algorithm (CRRA) works as follows:

The structure to be built is a tree, initially null, of which each node is a sentence and each edge a coherence relation. As each new sentence comes in, the CRRA tries to find the right node of the tree to attach it to, starting at the leaf that is the previous sentence and working back up the tree in a specified search order (discussed below) until a connection is indicated. Lockman assumes the existence of a judgment mechanism that generates and tests hypotheses as to how the new sentence may be **feasibly connected** to the node being tested. The first hypothesis whose likelihood exceeds a certain threshold is chosen.

The hypotheses consider both the coherence and the coreference relations that may obtain. Each member of the set of coherence relations is hypothesized, and for each one, all possible coreference relations between the conceptual tokens of the new sentence and tokens in the node under consideration (or nearby it in the tree) are posited. (The search for tokens goes back as far as necessary in the tree until suitable tokens are found for all unfulfilled definite noun phrases.) The hypotheses are considered in parallel; if none are judged sufficiently likely, the next node or set of nodes will be considered for feasible connection to the current sentence.

The search order is as follows: First the **immediate context**, the previous sentence, is tried. If no feasible connection is found, then the immediate ancestor of this node, and all its other descendants, are tried in parallel. If the algorithm is still unsuccessful, the immediate ancestor of the immediate ancestor, and the descendants thereof, are tried, and so on up the tree. If a test of several nodes in parallel yields more than one acceptable node, the one nearest the immediate context is chosen.

If the current sentence is not a simple sentence, it is not broken into clauses dealt with individually, but rather converted to a small sub-tree, reflecting the semantic relationship between the clauses. The conversion is based simply upon a table look-up indexed on the structure of the parse tree of the sentence. One of the nodes is designated by the table look-up as the head node, and the sub-tree is attached to the pre-existing context tree, using the procedure described above, with the connection occurring at this node. Similarly one (or more) of the nodes is designated as the immediate context, the starting point for the next search. (The search will be conducted in parallel if there is more than one immediate context node.)

There are some possible problems with Lockman's approach. The first lies in the fact that the structure built grows without limit, and therefore a search in it could, in theory, run right through an enormous tree. Normally, of course, a feasible connection or desired referent will be found fairly quickly, close to the immediate context. However, should the judgment mechanism fail to spot the correct one, the algorithm may run a little wild, searching large areas of the structure needlessly and expensively, possibly lighting on a wrong referent or wrong node for attachment, with no indication that an error has occurred. In other words, Lockman's CRRA places much greater trust in the judgment mechanism than a system like Grosz's that constrains the referent search area — more trust than perhaps should be put in what will necessarily be the most tentative and unreliable part of the system.

Secondly, I am worried about the syntax-based table look-up for sub-trees for complex sentences. On the one hand, it would be nice if it were correct, simplifying processing. On the other hand, I cannot but feel that it is an over-simplification, and that effects of discourse theme cannot reliably be handled in this way. However, I have no counterexamples to give, and suggest that this question needs more investigation.

The third possible problem, and perhaps the most serious, concerns the order in which the search for a feasible connection takes place. Because the first hypothesis whose likelihood exceeds the threshold is selected, it is possible to miss an even better hypothesis further up the tree. In theory, this could be avoided by doing all tests in parallel, the winning hypothesis being judged on both likelihood and closeness to the immediate context. In practice, given the ever-growing context tree as discussed above, this would not be feasible, and some way to limit the search area would be needed.

The fourth problem lies in the judgment mechanism itself. Lockman frankly admits that the mechanism, incorporated as a black box in his algorithm, must have abilities far beyond those of present state-of-the-art inference and judgment systems. The problem is that it is unwise to predicate too much on the nature of this unbuilt black box, as we do not know yet if its input-output behavior could be as Lockman posits. It may well be that to perform as required, the mechanism will need access to information such as the sentence following the current one (in effect, the ability to delay a decision), or more information about the previous context than the CRRA retains or ever determines; in fact, it may need an entirely different discourse structure representation from the tree being built. In other words, while it is fine in theory to design a reference resolver around a black box, in practice it may be computationally more economical to design the reference resolver around a knowledge of how the black box actually works, exploiting that mechanism, rather than straitjacketing the judgment module into its pre-defined cabinet; thus Lockman's work may be premature.

None of these problems are insurmountable. However it is perhaps a little unfortunate that Lockman's work offers little of immediate use for NLU systems of the present day.

6.4 Conclusions

Clearly, much work remains to be done if the coherence/cohesion paradigm of NLU is to be viable. Almost all aspects need refinement. However, it is an intuitively appealing paradigm, and it will be interesting to see if it can be developed into functioning NLU systems.

7. Epilogue

Each approach examined offers a different insight into some aspect or aspects of the use of discourse structure to resolve anaphora. So far there has been no attempt to integrate these insights into a single cohesive system or model; indeed this will be an extremely difficult task. It should, however, be a most fruitful one, and is the logical next step in computational anaphora resolution.

Acknowledgements

References

Bullwinkle, Candace Lee. *see also* Sidner, Candace Lee.

Bullwinkle, Candace Lee (1977a). The semantic component of PAL: The personal assistant language understanding program. Working paper 141, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, March 1977.

Bullwinkle, Candace Lee (1977b). Levels of complexity in discourse for anaphora disambiguation and speech act interpretation. [1] *Proceedings of the fifth international joint conference on artificial intelligence.* Cambridge, Massachusetts, August 1977, 43-49. [2] an earlier version was published as: Memo 413, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, May 1977.

Chafe, Wallace L (1970). *Meaning and the structure of language.* The University of Chicago Press, 1970.

Charniak, Eugene (1978). With spoon in hand this must be the eating frame. in: Waltz 1978, 187-193.

Charniak, Eugene (1981). Context recognition in language comprehension. in: Wendy Grace Lehnert and Martin Ringle (editors). *Knowledge representation for language processing systems,* Hillsdale, NJ: Lawrence Erlbaum Associates, 1981 [forthcoming].

Eisenstadt, Marc (1976). Processing newspaper stories: some thoughts on fighting and stylistics. *Proceedings of the AISB summer conference,* Society for the Study of Artificial Intelligence and the Simulation of Behaviour, July 1976, 104-117.

Firbas, Jan (1964). On defining the theme in functional sentence analysis. *Travaux linguistiques de Prague, 1,* 1964, 267-280.

Grosz, Barbara Jean (1977a). The representation and use of focus in a system for understanding dialogs. [1] *Proceedings of the fifth international joint conference on artificial intelligence.* Cambridge, Massachusetts, August 1977, 67-76. [2] Technical note 150, Artificial Intelligence Center, SRI International, June 1977.

Grosz, Barbara Jean (1977b). The representation and use of focus in dialogue understanding. [1] Unpublished PhD thesis, Department of Computer Science, University of California, Berkeley, June 1977. [2] published, slightly revised, as: Technical note 151, SRI International, Artificial Intelligence Center, July 1977. [3] a newer revised version appears in: Walker 1978, section 4.

Grosz, Barbara Jean (1978). Focusing in dialog. [1] in: Waltz 1978, 96-103. [2] Technical note 166, Artificial Intelligence Center, SRI International, July 1978.

Halliday, Michael Alexander Kirkwood and Hasan, Ruqaiya (1976). *Cohesion in English.* (= Longman English Language Series 9). London: Longman, 1976.

Hendrix, Gary Grant (1975). Expanding the utility of semantic networks through partitioning. *Advance papers of the fourth international joint conference on artificial intelligence.* Tblisi, Union of Soviet Socialist Republics, September 1975, 115-121.

Hendrix, Gary Grant (1978). The representation of semantic knowledge. in: Walker 1978, 121-181.

Hirst, Graeme John (1977a). Focus in reference resolution in natural language understanding. Paper presented at the Language and Speech Conference, Melbourne, November 1977.

Hirst, Graeme John (1977b). Cohesive discourse transitions and reference resolution: The cinema metaphor and beyond into the transfinite. Unpublished manuscript, 20 December 1977.

Hirst, Graeme John (1978). A set of primitives for discourse transitions. Unpublishable manuscript, 1 February 1978.

Hirst, Graeme John (1981). *Anaphora in natural language under-standing: a survey.* [1] (Lecture notes in computer science), NY: Springer-Verlag, 1981. [2] Technical report 79-2, Department of Computer Science, University of British Columbia, Vancouver, B.C., Canada, May 1979.

Hobbs, Jerry Robert (1978). Why is discourse coherent? [1] Technical note 176, Artificial Intelligence Center, SRI International, 30 November 1978. [2] in: F Neubauer (editor). *Coherence in natural language texts,* 1979.

Hobbs, Jerry Robert (1979). Coherence and coreference. [1] *Cognitive Science, 3*(1), January-March 1979, 67-90. [2] Technical note 168, Artificial Intelligence Center, SRI International, 4 August 1978.

Kantor, Robert Neal (1977). The management and comprehension of discourse connection by pronouns in English. PhD thesis, Department of Linguistics, Ohio State University, 1977.

Klappholz, A David and Lockman, Abe David (1977). The use of dynamically extracted context for anaphoric reference resolution. Unpublished MS, Department of Electrical Engineering and Computer Science, Columbia University, New York, February 1977.

Lockman, Abe David (1978). Contextual reference resolution [1] PhD dissertation, Faculty of Pure Science, Columbia University, May 1978. [2] Technical report DCS-TR-70, Department of Computer Science, Rutgers University, 1978.

Minsky, Marvin Lee (1975). A framework for representing knowledge. [1] in: Patrick Henry Winston (editor). *The psychology of computer vision.* McGraw-Hill, 1975, 211-280. [2] Memo 306, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, June 1974. [3] a condensed version appears in: Schank and Nash-Webber 1975, 118-130. [4] version 3 also appears as: Frame-system theory. in: Philip Nicholas Johnson-Laird and Peter Cathcart Wason (editors). *Thinking: Readings in cognitive science.* Cambridge University Press, 1977, 355-376.

Nash-Webber, Bonnie Lynn. *See also* Webber, Bonnie Lynn.

Nash-Webber, Bonnie Lynn (1976). Semantic interpretation revisited. Report 3335 (AI report 48), Bolt Beranek and Newman Inc, Cambridge, Massachusetts, July 1976.

Nash-Webber, Bonnie Lynn and Reiter, Raymond (1977). Anaphora and logical form: on formal meaning representations for natural language. [1] *Proceedings of the fifth international joint conference on artificial intelligence.* Cambridge, Massachusetts, August 1977, 121-131. [2] Technical report CSR-36, Center for the Study of Reading, University of Illinois at Urbana-Champaign, 1977.

Phillips, Brian (1977). Discourse connectives. Technical report KSL-11, Knowledge Systems Laboratory, Department of Information Engineering, University of Illinois at Chicago Circle, March 1977.

Pitkin, Willis L Jr (1977a). Hierarchies and the discourse hierarchy. *College English, 38*(7), March 1977, 648-659.

Pitkin, Willis L Jr (1977b). X/Y: Some basic strategies of discourse. *College English, 38*(7), March 1977, 660-672.

Reichman, Rachel (1978). Conversational coherency. *Cognitive science, 2*(4), October-December 1978, 283-327.

Roberts, R Bruce and Goldstein, Ira P (1977a). The FRL primer. Memo 408, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, July 1977.

Roberts, R Bruce and Goldstein, Ira P (1977b). The FRL manual. Memo 409, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, September 1977.

Rosenberg, Steven T (1976). Discourse structure. Working paper 130, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 17 August 1976.

Rosenberg, Steven T (1977). Frame-based text processing. Memo 431, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, November 1977.

Schank, Roger Carl and Abelson, Robert P (1977). *Scripts, plans, goals and understanding: An enquiry into human knowledge structures.* Hillsdale, New Jersey: Lawrence Erlbaum Associates, 1977. 1975, 309-328.

Schank, Roger Carl and Nash-Webber, Bonnie Lynn (1975). *Theoretical issues in natural language processing: An inter-disciplinary workshop.* Cambridge, Massachusetts: Association for Computational Linguistics, June 1975.

Sidner, Candace Lee. *see also* Bullwinkle, Candace Lee.

Sidner, Candace Lee (1978a). A progress report on the discourse and reference components of PAL. [1] in: *Proceedings of the second national conference,* Canadian Society for Computational Studies of Intelligence/Societe canadienne des etudes d'intelligence par ordinateur. Toronto. July 1978, 206-213. [2] Memo 468, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, 1978.

Sidner, Candace Lee (1978b). The use of focus as a tool for the disambiguation of definite noun phrases. in: Waltz 1978, 86-95.

Sidner, Candace Lee (1979). Towards a computational theory of definite anaphora comprehension in English discourse. [1] PhD thesis, Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, 16 May 1979. [2] revised as: Technical report 537, Artificial Intelligence Laboratory, Massachusetts Institute of Technology, June 1979.

Walker, Donald E (editor) (1978). *Understanding spoken language.* (The computer science library, Artificial intelligence series 5), New York: North-Holland, 1978.

Waltz, David L (editor) (1978). *TINLAP-2: Theoretical issues in natural language processing 2.* [1] University of Illinois at Urbana-Champaign, 25-27 July 1978. [2] reprinted in: *American journal of computational linguistics,* fiche 78-80, 1978.

Webber, Bonnie Lynn. *see also* Nash-Webber, Bonnie Lynn.

Webber, Bonnie Lynn (1978a). *A formal approach to discourse anaphora.* [1] (Outstanding dissertations in linguistics), NY: Garland Publishing, 1978. [2] PhD thesis, Department of Applied Mathematics, Harvard University, 1978. [3] Report 3761, Bolt Beranek and Newman Inc, May 1978.

Webber, Bonnie Lynn (1978b). Description formation and discourse model synthesis. in: Waltz 1978, 42-50.

Winograd, Terry (1972). *Understanding natural language.* [1] New York: Academic Press, 1972. [2] Edinburgh University Press, 1972. [3] also published in: *Cognitive psychology, 3*(1), 1972, 1-191.

Woods, William A; Kaplan, Ronald M and Nash-Webber, Bonnie Lynn (1972). The Lunar Science Natural Language Information System: Final report. Report 2378, Bolt Beranek and Newman Inc, Cambridge, Massachusetts, June 1972.

Graeme Hirst is a doctoral candidate in the Department of Computer Science at Brown University. He received the M.Sc. degree in Engineering Physics from the Australian National University in 1980.