Mixed Initiative in Dialogue: An Investigation into Discourse Segmentation

Marilyn Walker
University of Pennsylvania*
Computer Science Dept.
Philadelphia, PA 19104
lyn@linc.cis.upenn.edu

Steve Whittaker
Hewlett Packard Laboratories
Bristol, England BS12 6QZ
& HP Stanford Science Center
sjw@hplb.hpl.hp.com

Abstract

Conversation between two people is usually of MIXED-INITIATIVE, with CONTROL over the conversation being transferred from one person to another. We apply a set of rules for the transfer of control to 4 sets of dialogues consisting of a total of 1862 turns. The application of the control rules lets us derive domain-independent discourse structures. The derived structures indicate that initiative plays a role in the structuring of discourse. In order to explore the relationship of control and initiative to discourse processes like centering, we analyze the distribution of four different classes of anaphora for two data sets. This distribution indicates that some control segments are hierarchically related to others. The analysis suggests that discourse participants often mutually agree to a change of topic. We also compared initiative in Task Oriented and Advice Giving dialogues and found that both allocation of control and the manner in which control is transferred is radically different for the two dialogue types. These differences can be explained in terms of collaborative planning principles.

1 Introduction

Conversation between two people has a number of characteristics that have yet to be modeled adequately in human-computer dialogue. Conversation is BIDIRECTIONAL; there is a two way flow of information between participants. Information is exchanged by MIXED-INITIATIVE. Each participant will, on occasion, take the conversational lead. Conversational partners not only respond to what others say, but feel free to volunteer information that is not requested and sometimes ask questions of their own[Nic76]. As INITIATIVE passes back and forth between the discourse participants, we say that CONTROL over the conversation gets transferred from one discourse participant to another.

Why should we, as computational linguists, be interested in factors that contribute to the interactivity of a discourse? There are both theoretical and practical motivations. First, we wish to extend formal accounts of single utterances produced by single speakers to explain multi-participant, multi-utterance discourses[Pol86, CP86]. Previous studies of the discourse structure of multi-participant dialogues have often factored out the role of MIXED-INITIATIVE, by allocating control to one participant[Gro77, Coh84], or by assuming a passive listener[McK85, Coh87]. Since conversation is a collaborative process[CWG86, SSJ74], models of conversation can provide the basis for extending planning theories[GS90, CLNO90]. When the situation requires the negotiation of a collaborative plan, these theories must account for the interacting beliefs and intentions of multiple participants.

¿From a practical perspective, there is ample evidence that limited mixed-initiative has contributed to lack of system usability. Many researchers have noted that the absence of mixed-initiative gives rise to two problems with expert systems: They don't allow users to participate in the reasoning process, or to ask the questions they want answered[PHW82, Kid85, FL89]. In addition, question answering systems often fail to take account of the system's role as a conversational partner.

*This research was partially funded by ARO grants DAAG29-84-K-0061 and DAAL03-89-C0031PRI, DARPA grant N00014-85-K0018, and NSF grant MCS-82-19196 at the University of Pennsylvania, and by Hewlett Packard, U.K.

For example, fragmentary utterances may be interpreted with respect to the previous user input, but what users say is often in reaction to the system's previous response[CP82, Sid83].

In this paper we focus on interactive discourse. We model mixed-initiative using an utterance type classification and a set of rules for transfer of control between discourse participants that were proposed by Whittaker and Stenton[WS88]. We evaluate the generality of this analysis by applying the control rules to 4 sets of dialogues, including both advisory dialogues (ADs) and task-oriented dialogues (TODs). We analysed both financial and support ADs. The financial ADs are from the radio talk show "Harry Gross: Speaking of Your Money"[1]. The support ADs resulted from a client phoning an expert to help them diagnose and repair various software faults[2]. The TODs are about the construction of a plastic water pump in both telephone and keyboard modality[3].

The application of the control rules to these dialogues lets us derive domain-independent discourse segments with each segment being controlled by one or other discourse participant. We propose that control segments correspond to different subgoals in the evolving discourse plan. In addition, we argue that various linguistic devices are necessary for conversational participants to coordinate their contributions to the dialogue and agree on their mutual beliefs with respect to a evolving plan, for example, to agree that a particular subgoal has been achieved. A final phenomenon concerns shifts of control and the devices used to achieve this. Control shifts occur because it is unusual for a single participant to be responsible for coordinating the achievement of the whole discourse plan. When a different participant assumes control of a discourse subgoal then a control shift occurs and the participants must have mechanisms for achieving this. The control framework distinguishes instances in which a control shift is negotiated by the participants and instances where one participant seizes control.

This paper has two objectives:

- To explore the phenomenon of control in relation to ATTENTIONAL STATE [GS86, GJW86, Sid79][4]. We predict shifts of attentional state when shifts in control are negotiated and agreed by all participants, but not when control is seized by one participant without the acceptance of the others. This should be reflected in different distribution of anaphora in the two cases.
- To test predictions about the distribution of control in different types of dialogues. Because the TOD's embody the master-slave assumption[GS90], and control is allocated to the expert, our expectation is that control should be located exclusively with one participant in the TODs in contrast with the ADs.

2 Rules for the Allocation and Transfer of Control

We use the framework for the allocation and transfer of control of Whittaker and Stenton[WS88]. The analysis is based on a classification of utterances into 4 types[5]. These are:

- **UTTERANCE TYPES**
  - ASSERTIONS: Declarative utterances used to state facts. *Yes* and *No* in response to a question were classified as assertions on the basis that they are supplying information.
  - COMMANDS: Utterances intended to instigate action. Generally imperative form, but could be indirect such as *My suggestion would be that you do .....*
  - QUESTIONS: Utterances which are intended to elicit information, including indirect forms such as *I was wondering whether I should ....*
  - PROMPTS: Utterances which did not express propositional content, such as *Yeah, Okay, Uh-huh ....*

[1] 10 randomly selected dialogues (474 turns) from a corpus that was collected and transcribed by Martha Pollack and Julia Hirschberg[HL87, PHW82].

[2] 4 dialogues (450 turns) from tapes made at one of Hewlett-Packard's customer response centers. See [WS88].

[3] 5 keyboard (224 turns) and 5 telephone dialogues (714 turns), which were collected in an experiment by Phil Cohen to explore the relationship between modality, interactivity and use of referring expressions[Coh84].

[4] The theory of centering, which is part of attentional state, depends on discourse participants' recognizing the beginning and end of a discourse segment[BFP87, Wal89].

[5] The relationship between utterance level meaning and discourse intentions rests on a theory of joint commitment or shared plans[GS90, CLNO90, LCN90]

Note that prompts are in direct contrast to the other options that a participant has available at any point in the discourse. By indicating that the speaker does not want the floor, prompts function on a number of levels, including the expression of understanding or agreement[Sch82].

The rules for the allocation of control are based on the utterance type classification and allow a dialogue to be divided into segments that correspond to which speaker is the controller of the segment.

- **CONTROL RULES**

UTTERANCE	CONTROLLER (ICP)
ASSERTION	SPEAKER, unless response to a Question
COMMAND	SPEAKER
QUESTION	SPEAKER, unless response to Question or Command
PROMPT	HEARER

The definition of controller can be seen to correspond to the intuitions behind the term INITIATING CONVERSATIONAL PARTICIPANT (ICP), who is defined as the initiator of a given discourse segment[GS86]. The OTHER CONVERSATIONAL PARTICIPANT(S), OCP, may speak some utterances in a segment, but the DISCOURSE SEGMENT PURPOSE, must be the purpose of the ICP. The control rules place a segment boundary whenever the roles of the participants (ICP or OCP) change. For example:

Abdication Example
E: "And they are, in your gen you'll find that they've relocated into the labelled common area"
(ASSERT - E control)
C: "That's right." (PROMPT - E control)
E: "Yeah" (PROMPT - E abdicates control)
—— CONTROL SHIFT TO C ——
C: "I've got two in there. There are two of them." (ASSERT - C control)
E: "Right" (PROMPT - C control)
C: "And there's another one which is % RESA"
(ASSERT - C control)
E: "OK um" (PROMPT - C control)
C: "VS" (ASSERT - C control)
E: "Right" (PROMPT - C control)
C: "Mm" (PROMPT - C abdicates control)
—— CONTROL SHIFT TO E ——
E: "Right and you haven't got - I assume you haven't got local labelled common with those labels"
(QUESTION - E control)

Whittaker and Stenton also performed a post-hoc analysis of the segment boundaries that are defined by the control rules. The boundaries fell into one of three types:

- **CONTROL SHIFT TYPES**
  - ABDICATION: *Okay, go on.*
  - REPETITION/SUMMARY: *That would be my recommendation and that will ensure that you get a logically integral set of files.*
  - INTERRUPTION: *It is something new though um.*

ABDICATIONS[6] correspond to those cases where the controller produces a prompt as the last utterance of the segment. The class REPETITION/SUMMARY corresponds to the controller producing a **redundant** utterance. The utterance is either an exact repetition of previous propositional content, or a summary that realizes a proposition, P, which could have been inferred from what came before. Thus orderly control shifts occur when the controller explicitly indicates that s/he wishes to relinquish control. What unifies ABDICATIONS and REPETITION/SUMMARIES is that the controller supplies no new propositional content. The remaining class, INTERRUPTIONS, characterize shifts occurring when the noncontroller displays initiative by seizing control. This class is more general than other definitions of Interruptions. It properly contains cross-speaker interruptions that involve topic shift, similar to the true-interruptions of Grosz and Sidner[GS86], as well as clarification subdialogues[Sid83, LA90].

This classification suggests that the transfer of control is often a collaborative phenomenon. Since a noncontroller(OCP), has the option of seizing control at any juncture in discourse, it would seem that controllers(ICPs), are in control because the noncontroller allows it. These observations address problems raised by Grosz and Sidner, namely how ICPs signal and OCPs recognize segment boundaries. The claim is that shifts of control often do not occur until the controller indicates the end of a discourse segment by abdicating or producing a repetition/summary.

3 Control Segmentation and Anaphora

To determine the relationship between the derived control segments and ATTENTIONAL STATE we

[6] Our abdication category was called prompt by [WS88].

looked at the distribution of anaphora with respect to the control segments in the ADs. All data were analysed statistically by χ^2 and all differences cited are significant at the 0.05 level. We looked at all anaphors (excluding first and second person), and grouped them into 4 classes.

- **Classes of Anaphors**
  - 3RD PERSON: *it, they, them, their, she, he, her, him, his*
  - ONE/SOME, *one of them, one of those, a new one, that one, the other one, some*
  - DEICTIC: Noun phrases, e.g. *this, that, this NP, that NP, those NP, these NP*
  - EVENT: Verb Phrases, Sentences, Segments, e.g. *this, that, it*

The class DEICTIC refers to deictic references to material introduced by noun phrases, whereas the class EVENT refers to material introduced clausally.

3.1 Hierarchical Relationships

The first phenomenon we noted was that the anaphora distribution indicated that some segments are hierarchically related to others[7]. This was especially apparent in cases where one discourse participant interrupted briefly, then immediately passed control back to the other.

Interrupt/Abdicate 1

A: ... the only way I could do that was to take a to take a one third down and to take back a mortgage (ASSERTION)
————INTERRUPT SHIFT TO B——-
2. B: When you talk about one third put a number on it (QUESTION)
3. A: uh 15 thou (ASSERTION, but response)
4. B: go ahead (PROMPT)
————ABDICATE SHIFT BACK TO A——
5. A: and then I'm a mortgage back for 36

The following example illustrates the same point.

Interrupt/Abdicate 2

1. A: The maximum amount ... will be $400 on THEIR tax return. (ASSERTION)
———-INTERRUPT SHIFT TO B——-
2. B: 400 for the whole year? (QUESTION)
3. A: yeah it'll be 20% (ASSERTION, but response)
4. B: um hm (PROMPT)
———-ABDICATE SHIFT BACK TO A——-
5. A: now if indeed THEY pay the $2000 to your wife....

[7] Similar phenomena has been noted by many researchers in discourse including[Gro77, Hob79, Sid79, PH90].

The control segments as defined would treat both of these cases as composed of 3 different segments. But this ignores the fact that utterances (1) and (5) have closely related propositional content in the first example, and that the plural pronoun straddles the central subsegment with the same referents being picked out by *they* and *their* in the second example. Thus we allowed for hierarchical segments by treating the interruptions of 2-4 as subsegments, and utterances 1 and 5 as related parts of the parent segments. All interruptions were treated as embeddings in this way. However the relationship of the segment after the interruption to the segment before must be determined on independent grounds such as topic or intentional structure.

3.2 Distribution

Once we extended the control framework to allow for the embedding of interrupts, we coded every anaphor with respect to whether its antecedent lay outside or within the current segment. These are labelled X (cross segment boundary antecedent) NX (no cross segment boundary), in Figure 1. In addition we break these down as to which type of control shift occurred at the previous segment boundary.

	3rd Pers		One		Deictic		Event	
	X	NX	X	NX	X	NX	X	NX
Abdication	1	105	0	10	13	27	7	18
Summary	3	33	0	4	3	5	2	5
Interrupt	7	27	0	0	8	9	2	11
TOTAL	11	165	0	14	24	41	11	34

Figure 1: Distribution of Anaphora in Finance ADs

We also looked at the distribution of anaphora in the Support ADs and found similar results.

For both dialogues, the distribution of anaphors varies according to which type of control shift occurred at the previous segment boundary. When we look at the different types of anaphora, we find that third person and one anaphors cross bound-

	3rd Pers		One		Deictic		Event	
	X	NX	X	NX	X	NX	X	NX
Abdication	4	46	0	3	4	12	4	8
Summary	4	26	1	4	10	6	9	24
Interrupt	8	40	0	4	5	5	5	10
TOTAL	16	112	1	11	19	23	18	42

Figure 2: Distribution of Anaphora in Support ADs

aries extremely rarely, but the event anaphors and the deictic pronouns demonstrate a different pattern. What does this mean?

The fact that anaphora is more likely to cross segment boundaries following interruptions than for summaries or abdications is consistent with the control principles. With both summaries and abdications the speaker gives an explicit signal that s/he wishes to relinquish control. In contrast, interruptions are the unprompted attempts of the listener to seize control, often having to do with some 'problem' with the controller's utterance. Therefore, interruptions are much more likely to be within topic.

But why should deixis and event anaphors behave differently from the other anaphors? Deixis serves to pick out objects that cannot be selected by the use of standard anaphora, i.e. we should expect the referents for deixis to be outside immediate focus and hence more likely to be outside the current segment[Web86]. The picture is more complex for event anaphora, which seems to serve a number of different functions in the dialogue. It is used to talk about the past events that lead up to the current situation, *I did THAT in order to move the place.* It is also used to refer to sets of propositions of the preceding discourse, *Now THAT'S a little background* (cf [Web88]). The most prevalent use, however, was to refer to future events or actions, *THAT would be the move that I would make - but you have to do IT the same day.*

SUMMARY EXAMPLE

A: As far as you are concerned THAT could cost you more .... what's your tax bracket? (QUESTION)

B: Well I'm on pension Harry and my wife hasn't worked at all and ..(ASSERT/RESP)

A: No reason at all why you can't do THAT. (ASSERTION)

———SUMMARY SHIFT to B——

B: See my comment was if we should throw even the $2000 into an IRA or something for her. (ASSERTION)

———REPETITION SHIFT to A——

A: You could do THAT too. (ASSERTION)

Since the task in the ADs is to develop a plan, speakers use event anaphora as concise references to the plans they have just negotiated and to discuss the status and quality of plans that have been suggested. Thus the frequent cross-speaker references to future events and actions correspond to phases of plan negotiation[PHW82]. More importantly these references are closely related to the control structure. The example above illustrates the clustering of event anaphora at segment boundaries. One discourse participant uses an anaphor to summarize a plan, but when the other participant evaluates this plan there may be a control shift and any reference to the plan will necessarily cross a control boundary. The distribution of event anaphora bears this out, since 23/25 references to future actions are within 2 utterances of a segment boundary (See the example above). More significantly every instance of event anaphora crossing a segment boundary occurs when the speaker is talking about future events or actions.

We also looked at the TODs for instances of anaphora being used to describe a future act in the way that we observed in the ADs. However, over the 938 turns in the TODs, there were only 18 instances of event anaphora, because in the main there were few occasions when it was necessary to talk about the plan. The financial ADs had 45 event anaphors in 474 utterances.

4 Control and Collaborative Plans

To explore the relationship of control to planning, we compare the TODs with both types of ADs (financial and support). We would expect these dialogues to differ in terms of initiative. In the ADs, the objective is to develop a collaborative plan through a series of conversational exchanges. Both discourse participants believe that the expert has knowledge about the domain, but only has partial information about the situation. They also believe that the advisee must contribute both the problem description and also constraints as to how the problem can be solved. This information must be exchanged, so that the mutual beliefs necessary to develop the collaborative plan are established in the conversation[Jos82]. The situation is different

in the TODs. Both participants here believe at the outset that the expert has sufficient information about the situation and complete and correct knowledge about how to execute the Task. Since the apprentice has no need to assert information to change the expert's beliefs or to ask questions to verify the expert's beliefs or to issue commands, we should not expect the apprentice to have control. S/he is merely present to execute the actions indicated by the knowledgeable participant.

The differences in the beliefs and knowledge states of the participants can be interpreted in the terms of the collaborative planning principles of Whittaker and Stenton[WS88]. We generalize the principles of INFORMATION QUALITY and PLAN QUALITY, which predict when an interrupt should occur.

- INFORMATION QUALITY: The listener must believe that the information that the speaker has provided is true, unambiguous and relevant to the mutual goal. This corresponds to the two rules: (A1) TRUTH: If the listener believes a fact P and believes that fact to be relevant and either believes that the speaker believes not P or that the speaker does not know P then interrupt; (A2)AMBIGUITY: If the listener believes that the speaker's assertion is relevant but ambiguous then interrupt.
- PLAN QUALITY: The listener must believe that the action proposed by the speaker is a part of an adequate plan to achieve the mutual goal and the action must also be comprehensible to the listener. The two rules to express this are: (B1)EFFECTIVENESS: If the listener believes P and either believes that P presents an obstacle to the proposed plan or believes that P is part of the proposed plan that has already been satisfied, then interrupt; (B2) AMBIGUITY: If the listener believes that an assertion about the proposed plan is ambiguous, then interrupt.

These principles indirectly provide a means to ensure mutual belief. Since a participant must interrupt if any condition for an interrupt holds, then lack of interruption signals that there is no discrepancy in mutual beliefs. If there is such a discrepancy, the interruption is a necessary contribution to a collaborative plan, not a distraction from the joint activity.

We compare ADs to TODs with respect to how often control is exchanged by calculating the average number of turns between control shifts[8]. We also investigate whether control is shared equally between participants and what percentage of control shifts are represented by abdications, interrupts, and summaries for each dialogue type. See Figure 3.

	Finance	Support	Task-Phone	Task-Key
Turns/Seg	7.49	8.03	15.68	11.27
Exp-Contr	60%	51%	91%	91%
Abdication	38%	38%	45%	28%
Summary	23%	27%	7%	6%
Interrupt	38%	36%	48%	67%

Turns/Seg: Average number of turns between control shifts
Exp-Contr: % total turns controlled by expert
Abdication: % control shifts that are Abdications
Summaries: % control shifts that are Reps/Summaries
Interrupt: % control shifts that are Interrupts

Figure 3: Differences in Control for Dialogue Types

Three things are striking about this data. As we predicted, the distribution of control between expert and client is completely different in the ADs and the TODs. The expert has control for around 90% of utterances in the TODs whereas control is shared almost equally in the ADs. Secondly, contrary to our expectations, we did find some instances of shifts in the TODs. Thirdly, the distribution of interruptions and summaries differs across dialogue types. How can the collaborative planning principles highlight the differences we observe?

There seem to be two reasons why shifts occur in the TODs. First, many interruptions in the TODs result from the apprentice seizing control just to indicate that there is a temporary problem and that plan execution should be delayed.

TASK INTERRUPT 1, A is the Instructor
A: It's hard to get on (ASSERTION)
———INTERRUPT SHIFT TO B——-
B: Not there yet - ouch yep it's there. (ASSERTION)
A: Okay (PROMPT)
B: Yeah (PROMPT)
———ABDICATE SHIFT TO A——
A: All right. Now there's a little blue cap ..

Second, control was exchanged when the execution of the task started to go awry.

[8] We excluded turns in dialogue openings and closings.

TASK INTERRUPT 2, A is the Instructor
A: And then the elbow goes over that ... the big end of the elbow. (COMMAND)
———INTERRUPT SHIFT TO B———
B: You said that it didn't fit tight, but it doesn't fit tight at all, okay ... (ASSERTION)
A: Okay (PROMPT)
B: Let me try THIS - oops - again(ASSERTION)

The problem with the physical situation indicates to the apprentice that the relevant beliefs are no longer shared. The Instructor is not in possession of critical information such as the current state of the apprentice's pump. This necessitates an information exchange to resynchronize mutual beliefs, so that the rest of the plan may be successfully executed. However, since control is explicitly allocated to the instructor in TODs, there is no reason for that participant to believe that the other has any contribution to make. Thus there are fewer attempts by the instructor to coordinate activity, such as by using summaries to synchronize mutual beliefs. Therefore, if the apprentice needs to make a contribution, s/he must do so via interruption, explaining why there are many more interruptions in these dialogues.[9] In addition, the majority of Interruptions (73%) are initiated by apprentices, in contrast to the ADs in which only 29% are produced by the Clients.

Summaries are more frequent in ADs. In the ADs both participants believe that a plan cannot be constructed without contributions from both of them. Abdications and summaries are devices which allow these contributions to be coordinated and participants use these devices to explicitly set up opportunities for one another to make a contribution, and to ensure mutual beliefs. The increased frequency of summaries in the ADs may result from the fact that the participants start with discrepant mutual beliefs about the situation and that establishing and maintaining mutual beliefs is a key part of the ADs.

5 Discussion

It has often been stated that discourse is an inherently collaborative process and that this is manifested in certain phenomena, e.g. the use of anaphora and cue words [GS86, HL87, Coh87] by which the speaker makes aspects of the discourse structure explicit. We found shifts of attentional state when shifts in control are negotiated and agreed by all participants, but not when control is seized by one participant without the acceptance of the others. This was reflected in different distribution of anaphora in the two cases. Furthermore we found that not all types of anaphora behaved in the same way. Event anaphora clustered at segment boundaries when it was used to refer to preceding segments and was more likely to cross segment boundaries because of its function in talking about the proposed plan. We also found that control was distributed and exchanged differently in the ADs and TODs. These results provide support for the control rules.

In our analysis we argued for hierarchical organization of the control segments on the basis of specific examples of interruptions. We also believe that there are other levels of structure in discourse that are not captured by the control rules, e.g. control shifts do not always correspond with task boundaries. There can be topic shifts without change of initiation, change of control without a topic shift[WS88]. The relationship of cue words, intonational contour[PH90] and the use of modal subordination[Rob86] to the segments derived from the control rules is a topic for future research.

A more controversial question concerns rhetorical relations and the extent to which these are detected and used by listeners[GS86]. Hobbs has applied COHERENCE RELATIONS to face-to-face conversation in which mixed-initiative is displayed by participants[HA85, Hob79]. One category of rhetorical relation he describes is that of ELABORATION, in which a speaker repeats the propositional content of a previous utterance. Hobbs has some difficulties determining the function of this repetition, but we maintain that the function follows from the more general principles of the control rules: speakers signal that they wish to shift control by supplying no new propositional content. Abdications, repetitions and summaries all add no new information and function to signal to the listener that the speaker has nothing further to say right now. The listener certainly must recognize this fact.

Summaries appear to have an additional function of synchronization, by allowing both participants to agree on what propositions are mutually believed at that point in the discussion. Thus this work highlights aspects of collaboration in discourse, but

[9] The higher percentage of Interruptions in the keyboard TODs in comparison with the telephone TODs parallels Oviatt and Cohen's analysis, showing that participants exploit the wider bandwidth of the interactive spoken channel to break tasks down into subtasks[Coh84, OC89].

should be formally integrated with research on collaborative planning[GS90, LCN90], particularly with respect to the relation between control shifts and the coordination of plans.

6 Acknowledgements

We would like to thank Aravind Joshi for his support, comments and criticisms. Discussions of joint action with Phil Cohen and the members of CSLI's DIA working group have influenced the first author. We are also indebted to Susan Brennan, Herb Clark, Julia Hirschberg, Jerry Hobbs, Libby Levison, Kathy McKeown, Ellen Prince, Penni Sibun, Candy Sidner, Martha Pollack, Phil Stenton, and Bonnie Webber for their insightful comments and criticisms on drafts of this paper.

References

[BFP87] Susan E. Brennan, Marilyn Walker Friedman, and Carl J. Pollard. A centering approach to pronouns. In *Proc. 25th Annual Meeting of the ACL*, pages 155–162, 1987.

[CLNO90] Phillip R. Cohen, Hector J. Levesque, Jose H. T. Nunes, and Sharon L. Oviatt. Task oriented dialogue as a consequence of joint activity, 1990. Unpublished Manuscript.

[Coh84] Phillip R. Cohen. The pragmatics of referring and the modality of communication. *Computational Linguistics*, 10:97–146, 1984.

[Coh87] Robin Cohen. Analyzing the structure of argumentative discourse. *Computational Linguistics*, 13:11–24, 1987.

[CP82] Phillip R. Cohen, C. Raymond Perrault, and James F. Allen 1982. Beyond question answering. In Wendy Lehnert and Martin Ringle, editors, *Strategies for Natural Language Processing*, pages 245–274. Lawrence Erlbaum Ass. Inc, Hillsdale, N.J., 1982.

[CP86] Philip R. Cohen and C. Raymond Perrault. Elements of a plan-based theory of speech acts. In Bonnie Lynn Webber Barbara J. Grosz, Karen Sparck Jones, editor, *Readings in Natural Language Processing*, pages 423–440. Morgan Kauffman, Los Altos, Ca., 1986.

[CWG86] Herbert H. Clark and Deanna Wilkes-Gibbs. Referring as a collaborative process. *Cognition*, 22:1–39, 1986.

[FL89] David M. Frohlich and Paul Luff. Conversational resources for situated action. In *Proc. Annual Meeting of the Computer Human Interaction of the ACM*, 1989.

[GJW86] Barbara J. Grosz, Aravind K. Joshi, and Scott Weinstein. Towards a computational theory of discourse interpretation. Unpublished Manuscript, 1986.

[Gro77] Barbara J. Grosz. The representation and use of focus in dialogue understanding. Technical Report 151, SRI International, 333 Ravenswood Ave, Menlo Park, Ca. 94025, 1977.

[GS86] Barbara J. Grosz and Candace L. Sidner. Attentions, intentions and the structure of discourse. *Computational Linguistics*, 12:pp. 175–204, 1986.

[GS90] Barbara J. Grosz and Candace L. Sidner. Plans for discourse. In *Cohen, Morgan and Pollack, eds. Intentions in Communication, MIT Press, Cambridge, MA.*, 1990.

[HA85] Jerry R. Hobbs and Michael H. Agar. The coherence of incoherent discourse. Technical Report CSLI-85-38, Center for the Study of Language and Information, Ventura Hall, Stanford University, Stanford, CA 94305, 1985.

[HL87] Julia Hirschberg and Diane Litman. Now lets talk about now: Identifying cue phrases intonationally. In *Proc. 25th Annual Meeting of the ACL*, pages 163–171, Stanford University, Stanford, Ca., 1987.

[Hob79] Jerry R. Hobbs. Coherence and coreference. *Cognitive Science*, 3:67–90, 1979.

[Jos82] Aravind K. Joshi. Mutual beliefs in question-answer systems. In *Neil V.*

Smith eds. Mutual Knowledge, Academic Press, New York, New York, pages 181–199, 1982.

[Kid85] Alison Kidd. The consultative role of an expert system. In P. Johnson and S. Cook, editors, *People and Computers: Designing the Interface*. Cambridge University Press, Cambridge, U.K., 1985.

[LA90] Diane Litman and James Allen. Recognizing and relating discourse intentions and task-oriented plans. In *Cohen, Morgan and Pollack, eds. Intentions in Communication, MIT Press, Cambridge, MA.*, 1990.

[LCN90] Hector J. Levesque, Phillip R. Cohen, and Jose H. T. Nunes. On acting together. In *AAAI90*, 1990.

[McK85] Kathleen R. McKeown. Discourse strategies for generating natural language text. *Artificial Intelligence*, 27(1):1–42, September 1985.

[Nic76] R.S. Nickerson. On converational interaction with computers. In SiegFried Treu, editor, *User-Oriented Design of Interactive Graphics Systems*, pages 101–65. Elsevier Science, 1976.

[OC89] Sharon L. Oviatt and Philip R. Cohen. The effects of interaction on spoken discourse. In *Proc. 27th Annual Meeting of the ACL*, pages 126–134, 1989.

[PH90] Janet Pierrehumbert and Julia Hirschberg. The meaning of intonational contours in the interpretation of discourse. In *Cohen, Morgan and Pollack, eds. Intentions in Communication, MIT Press, Cambridge, MA.*, 1990.

[PHW82] Martha Pollack, Julia Hirschberg, and Bonnie Webber. User participation in the reasoning process of expert systems. In *Proc. National Conference on Artificial Intelligence*, 1982.

[Pol86] Martha Pollack. Inferring domain plans in question answering. Technical Report 403, SRI International - Artificial Intelligence Center, 1986.

[Rob86] Craige Roberts. *Modal Subordination and Anaphora*. PhD thesis, Linguistics Dept, University of Massachusetts, Amherst, 1986.

[Sch82] Emanuel A. Schegloff. Discourse as an interactional achievement: Some uses of 'uh huh' and other things that come between sentences. In D. Tannen, editor, *Analyzing Discourse: Text and Talk*, pages 71–93. Georgetown University Press, 1982.

[Sid79] Candace L. Sidner. Toward a computational theory of definite anaphora comprehension in english. Technical Report AI-TR-537, MIT, 1979.

[Sid83] Candace Sidner. What the speaker means: the recognition of speakers plans in discourse. *International Journal of Computers and Mathematics*, 9:71–82, 1983.

[SSJ74] Harvey Sacks, Emmanuel Schegloff, and Gail Jefferson. A simplest systematics for the organization of turn-taking in conversation. *Language*, 50:pp. 325–345, 1974.

[Wal89] Marilyn A. Walker. Evaluating discourse processing algorithms. In *Proc. 27th Annual Meeting of the ACL*, pages 251–261, 1989.

[Web86] Bonnie Lynn Webber. Two steps closer to event reference. Technical Report MS-CIS-86-74, Linc Lab 42, Department of Computer and Information Science, University of Pennsylvania, 1986.

[Web88] Bonnie Lynn Webber. Discourse deixis: Reference to discourse segments. In *Proc. 26th Annual Meeting of the ACL, Association of Computational Linguistics*, pages 113–123, 1988.

[WS88] Steve Whittaker and Phil Stenton. Cues and control in expert client dialogues. In *Proc. 26th Annual Meeting of the ACL, Association of Computational Linguistics*, pages 123–130, 1988.