Discerning Relevant Information in Discourses Using TFA

Geert-Jan M. Kruijff & Jan Schaake
Department of Computer Science and Department of Philosophy and Social Sciences
University of Twente
P.O. Box 217, 7500 AE Enschede, The Netherlands
email: {kruijff,schaake}@cs.utwente.nl

Abstract

Discourses, whether written or spoken, are intended to convey information. Obviously, it is important to the processing of discourses that one is able to recognize the information that is relevant. The need for a criterion for relevance of information arises out of the idea of developing a tool assisting in the extraction of definitions from philosophical discourses (PAPER/HCRAES-projects).
A way to analyse a discourse with regard to the information expressed in it, is to observe the Topic-Focus Articulation. A topic of (part of) a discourse can be conceived of as already available information, to which more information is added by means of one or more foci. Several topics and foci of a discourse are organized in certain structures, characterized by a thematical progression ("story-line"). The theories about TFA and thematic progression have been developed by the Prague School of Linguistics.
In order to discern the relevant information in a discourse, we try to establish the thematic progression(s) in a discourse. It will turn out that it is important, not only how topics and foci relate to each other with regard to the thematic progression (sequentially, parallelly, etc.), but also how the topics and foci are related rhetorically (e.g. by negation). In this paper we shall come to defining the way in which the information structure of a discourse can be recognized, and what relevant information means in this context.

1 Introduction

Last year a project has been initiated at the University of Twente, as a co-operation between the departments of Computer Science, Public Policy and

Public Administration, and Philosophy and Social Sciences. The objective of this so-called PAPER project is to build a hypertext-system in which, generally speaking, definitions of terms are represented. The terms and their definitions are developed by the American philosopher Ch.S. Peirce (1839–1914). They have been obtained from Peirce's original, unpublished manuscripts as well as published works. By means of PAPER (= Peirce Anthology Project – Electronic Representation) passages considered relevant become electronically available, increasing the possibility to find to-the-point-descriptions of terms, as well as facilitating an easy use of these descriptions in the user's own work.

Since the selections defining the terms were collected by the developers themselves, it became an issue to make the manner of selections as objective as possible, thus, not biased by personal interests. A method has been used (in the beginning only loosely) to develop a formalization of the manner in which information can be communicated by a text, how relevant information may be discerned, and, consequently, how communicated information may be dealt with in a logical fashion. The result of the formalization has been that the procedure of selection has become less subjective. Furthermore, the formalization may serve as the foundation for (a part of) developer-oriented software supporting the developer in creating PAPER-alike systems. This project is called HCRAES: Hyper Card Representation of Anthologies, Entities and other Sources.

Part of the software may be a tool automatically selecting relevant portions in large texts full of examples and side remarks. The selection of relevant passages of text requires, first of all, an analysis of the thematic structure of the text. It is important to discern the different topics dealt with in the text, as well as the segments of text in which these topics are elaborated upon. The latter help in the determination of relevant selections. In order to do so, computational models have to be developed of existing theories with respect to structures of large texts such as Thematic Progression (Daneš 1979) and Rhetorical Structure Theory (Mann and Thompson 1987).

In order to develop and to test these models it has been regarded necessary to choose a domain of smaller texts where discerning relevant information is also needed. This alternative domain we found in the SCHISMA project, that is present within the same PARLEVINK research group where PAPER is located. The SCHISMA project is devoted to the development of a theatre information and booking system. One of the problems to be met in analysing dialogues is to discern what exactly is or are the point(s) made in a turn of the client. As we will see below, in one turn a client may make just one relevant remark, the rest being noise or background information that is not relevant to the system. It may also be the case that two or more relevant points are made in just one turn. These points have to be discerned as being both relevant. In section 3 examples of the occurrence of relevant information in a turn will be given. First, in section 2, an overview will be given of the treatment of relevant information in current theories on dialogue analysis. Subsequently, in section 3, Thematic Progression and Rhetorical Structure Theory will be applied to dialogues taken from the SCHISMA corpus and, in section 4, relevant information will be related to what will be called *generic tasks*; tasks that perform a small function centred around the goal of acquiring a specific piece of information (cf. Chandrasekaran (1986)). Conclusions will be drawn in the final section.

2 Discerning Information in Dialogues

With respect to the way information will be discerned in dialogues it is possible to trace a development from a frame-like approach towards an intentional or plan-based one. The former approach originates with Schaake and Nauta (1994) who introduced the *frame* concept as follows:

> "A frame is a data-structure for representing a stereotyped situation like being in a certain kind of living room or going to a child's birthday party. Attached to each frame are several kinds of information. Some of this information is about how to use the frame. Some is about what one can expect to happen next. Some is about what to do if these expectations are not confirmed."

In natural language processing this frame concept was used immediately in order to represent the information conveyed by sentences or a discourse. A frame was considered as a thing containing a number of slots. For instance, a the frame "birthday party" has to contain a slot for the person whose birthday it is, for the place the party takes place, for the guests, for the cakes and drinks, etc. Some of these slots have default values (the place the party takes place normally will be the home of the person whose birthday it is), all other information has to be gathered from the discourse. The same principle is still used in applications to extract information from, for instance, railway dialogues: a frame is created consisting of the slots departure, destination, reduction, single/return, first/second class, and date. Some slots have default values (departure, second class, today) while at least the destination has to be presented in the dialogue. More complicated frames or frame structures are treated by Minsky (1975) introducing so-called *scripts*. A script, being itself a frame too, often consists of a sequence of frames to be treated in a regular way. Most famous is the so-called "restaurant script" consisting in a sequence of action frames: entering the restaurant, finding a table, receiving the menu, ordering something to drink, etc. Applying this script concept to a theatre information and booking office, we get a sequence of requesting some general information about the program this season, selecting a particular performance, booking seats for it, asking what the price will be and where to park the car. According to this approach the information conveyed in the dialogue is expected already by the participants as soon as the know what kind of dialogue they are involved in (which is the case quite often).

The second, intentional, approach, has been highly inspired by Searle's *Speech Acts* theory (Searle 1969) and can be found with, for instance, Allen and Perrault (1980) and Lambert and Carberry (1991). In this approach utterances are related to a set of presupposed or recognized intentional states like beliefs, knowledge, desires, plans and goals. This so-called *user's* or *domain knowledge* is represented as mental states for both participants, that is to say that both will get a presupposed initial mental state consisting of beliefs, goals, etc., a state that will be updated during the dialogue with knowledge about the intentions and knowledge of the other participant to be recognized in his of her expressions. The recognition of the other's intentions can lead to some co-operative behaviour of an agent. Information is thus considered as the whole content of an utterance or speech act communicating part of the mental state of one participant to that of the other one. Only that information can be used, however, that fits a plan or action model that is also present in the mental state of one of the participants. So, for instance, out of an extended utterance only those portions will bear usable information that fit in an action model present in the hearer's mind. In the unmodified plan theory only the plans present in the hearer's mind determine what part of the information communicated has to be used and what not. The situation in which the utterance has been made, which is determining in the above frame or script theory, doesn't play any role. Neither does any emphasis by the speaker. By this, it is obvious that an intentional model doesn't fit the SCHISMA requirements according to which tasks have to be executed in a certain order and certain information states may cause the execution of a sub task.

In our view it is important to start the analysis of the speaker's utterances with an analysis of its structure in terms of coherence and relevance: what is the relationship between the different parts of one utterance referentially but also functionally? Having analysed the structure of utterances, their meaning has to be related to the current information state, thus, information state changes caused by the utterance have to be performed. Finally, the behaviour of the system has to be determined meeting the changes in the information state and its the general goals and tasks. This flow of analyses fits the theoretical considerations outlined by Schank and Abelson (1977). Moreover, distinct from the above mentioned models, our approach takes the way information is contained in the utterances themselves seriously.

3 The Communication of Information

Surely, it might almost sound like a commonplace that a dialogue conveys, or communicates, information[1]. But what can we say about the exact features of such communication? If we want to our logical theory of information to be of any use, we should elucidate how we arrive at the information we express in

[1]Supposed that the dialogue is meant be purposeful, of course. Otherwise, they are called "parasitic" with respect to communicative dialogues (cf. Habermas).

information states. Such elucidation is the issue of the current section.

The assumption we make about the dialogues to be considered is that they are coherent. Rather than being a set of utterances bearing no relation to each other, a dialogue—by the assumption—should have a 'story line'. For example, the utterances can therein be related by referring to a common topic, or by elaborating a little further upon a topic that was previously introduced. More formally, we shall consider utterances to be constituted of a Topic and Focus pair. The Topic of an utterance stands for *given information*, while the Focus of an utterance stands for *new information*. The theory of the articulation of Topic and Focus (TFA) has been developed by members of the Modern Prague School, notably by Hajicova (cf. Hajicova (1993a, 1994b, 1994a, 1993b)).

Consequently, the 'story line' of a dialogue becomes describable in terms of relations between Topics and Foci. The communication of information thus is describable in terms of how given information is used and new information is provided. The relations between Topics and Foci may be conceived of in two ways, basically: Thematically, and rhetorically. The thematical way concerns basically the coreferential aspect, while the rhetorical way concerns the functional relationship between portions of a discourse. Let us therefore have a closer look at each of these ways, and how they are related to each other.

First, the relations between Topics and Foci can be examined at the level of individual utterances. In that case we shall speak of *thematic* relations, elucidating the *thematic progression*. Thematic progression is a term introduced by Daneš (1979) as a means to analyse the thematic build-up of texts. We shall use it here in the analysis of the manner in which given and new information are bound to each other by utterances in a dialogue. According to Daneš, there are three possibilities in which Topics and Foci are bindable, which are described as the following kinds of progression:

1. **Sequential progression**: The Focus of utterance m, F_m, is constitutive for the Topic of a (the) next utterance n, T_n.
   Diagrammatically:
   $$\begin{array}{llll} T_m & \rightarrow & F_m & \\ & & \| \ \mathbf{seq} & \\ & & T_n & \rightarrow \quad F_n \end{array}$$

2. **Parallel progression**: The Topic of utterance m, T_m, bears much similarity to the Topic of a (the) next utterance n, T_n.
   Diagrammatically:
   $$\begin{array}{llll} T_m & \rightarrow & F_m & \\ & & \| \ \mathbf{par} & \\ & & T_n & \rightarrow \quad F_n \end{array}$$

3. **Hypertheme progression**: The Topic of utterance m, T_m, as well as the Topic of utterance n, T_n, refer to an overall Topic called the *Hypertheme*, T_H. Utterances m and n are said to be related hyperthematically.
   Diagrammatically:
   $$T_H \left\{ \begin{array}{lll} T_m & \rightarrow & F_m \\ T_n & \rightarrow & F_n \end{array} \right.$$

The following sentences are examples of these different kinds of progression:

(1) The brand of GJ's car is Trabant. The Trabant has a two-stroke engine.

(2) Trabis are famous for their funny motor-sound. Trabis are also well-known for the blue clouds to puff.

(3) Being a car for the whole family, the Trabant has several interesting features. One feature is that about every person can repair it. Another feature is that a child's finger-paint can easily enhance the permanent outlook of the car.

It might be tempting to try to determine the kind of thematic progression between utterances by merely looking at the predicates and entities involved. In other words, directly in terms of information states. Especially sentences like (1) and (2) tend to underline such a standpoint. However, consider the following revision of (1), named (1′):

(1′) GJ has a Trabant. The motor is a cute two-stroke engine.

Similar to (1) we would like to regard (1′) as a sequential progression. Yet, If we would consider only predicates and entities, we would not be able to arrive at that preferred interpretation. It is for that reason that we propose to determine the kind of thematic progression obtaining between two utterances as follows. Instead of discerning whether the predicates and entities of a Topic T_m or a Focus F_m are the same as those of a Topic T_n, we want to establish whether F_m or T_m and T_n are *coreferring*. We take coreference to mean that two expressions, E_1 and E_2

- are referring to the same *concept*, or
- are referring to a conceptual structure, where E_1 is referring to a concept C_{E_1} which is the parent of a concept C_{E_2}, to which E_2 is referring.

Hence, the following relations hold[2]:

1.	F_m and T_n are coreferring	$\rightarrow$	sequential progression
2.	T_m and T_n are coreferring	$\rightarrow$	parallel progression
3.	T_H, T_m and T_n are coreferring	$\rightarrow$	hypertheme progression

For our purposes we establish the thematic progression between a number utterances making up a single turn in a dialogue. As we already noted above, utterances can also be related rhetorically, besides thematically. Now, whereas the thematic progression shows us how information is being communicated by individual utterances, the rhetorical structure elucidates how parts of the communicated information functions in relation to other parts of information communicated within the same turn. In other words, the rhetorical structure considers the function of the information communicated by clusters of one or more utterances of a single turn. Such clusters will be called *segments* hereafter.

[2] The presented ideas about thematic progression and coreference result from discussions between Geert-Jan Kruijff and Ivana Korbayová.

When performing an analysis in order to explicate the rhetorical structure, one can make use of for example Mann and Thompson's Rhetorical Structure Theory (RST) as laid down in (Mann and Thompson 1987). Basically, RST enables us to structure a turn into separate segments that are functionally related to each other by means of so-called *rhetorical relations*. Examples of such rhetorical relations are:

(4) Segment S **is evidence for** segment N:
 (N) The engine of my car works really well nowadays.
 (S) It started yesterday within one minute.

(5) Segment S **provides background for** segment N:
 (S) I spend a significant part of the year in Prague.
 (N) Nowadays, I am the proud owner of a Trabant.

(6) Segment S **is a justification for** segment N:
 (S) When parking a little carelessly, I broke one of the rear lights.
 (N) I should buy a new rear light.

A study of a corpus of dialogues we have gathered reveals that within our domain the following rhetorical relations are of importance:

1. **Solutionhood**: S provides the solution for N;
   "Yes, but grandma is a little cripple, so, well, then we'll go with the two of us.'

2. **Background**: S provides background for N;
   "I would like to go to an opera. Is there one on Saturday?"

3. **Conditional**: S is a condition for N;
   "If the first row is right opposite to the stage, then the first row, please."

4. **Elaboration**: S elaborates on N;
   "I would like to go to Wittgenstein, because he was really entertaining last time."

5. **Restatement**: S restates or summarizes N;
   "So I have made a reservation for ..."

6. **Contrast**: Several N's are contrasted;
   "I would like to, but my friend does not. So, then we'd better not go to an opera; Can we go to an other performance?"

7. **Joint**: Several N's are joined;
   "How expensive would that be, and are there still vacant seats?"

In case of rhetorical relations 1 through 3 the S is uttered after N, while in case of the relations 4 through 5 S is uttered before N. Relations 6 and 7 are constituted by multiple nuclei.

The reader might have come to wonder what the S and N stand for. Therefore, let us provide some explanation. N stands for **nucleus**, while S stands for **satellite**. Obviously, both are segments. The distinction between them can be pointed out as follows. A nucleus is defined as a segment that serves as the locus of attention. A satellite is a segment that gains its significance through a nucleus. The concept of nuclearity is important to us: We would still have a coherent dialogue if we would consider the nuclei only. In our understanding, nuclearity is thus an expressive source that directs the response to a turn of a dialogue.

Thus, revisiting the thematic and rhetorical structure of a turn in a dialogue, we observe the following. The established thematic progression elucidates the actual flow of communicated information. Therein, we can observe which utterances convey what information. The rhetorical structure clarifies how information expressed by nuclei and satellites are functionally related to each other. Clearly, the question that might be raised subsequently is How does the segmentation of a turn into nuclei and satellites arise from the thematic progression?

To answer the question, we should realize that we are actually dealing with three smaller problems:

1. The segmentation-problem: How does a thematic progression segment a turn?

2. The problem of recognizing rhetorical relations: Which rhetorical relations are actually involved?

3. The problem of recognizing nuclei and satellites.

The answer to the first problem is as follows. A thematic progression divides a turn into discernible segments according to the flow of information. Intuitively, one might say that every time a new flow of information is commenced, a new segment is introduced. As we shall see in the example provided below, this means in general that when a parallel progression or hypertheme progression is invoked, a new segment starts. Regarding the second problem, Mann and Thompson describe how rhetorical relations can be recognized by means of conditions (or constraints) that should hold for the textual structure. We conjecture that, in terms of our approach, rhetorical relations can be recognized by taking the thematic progression and the formed conceptual structure into account. Rephrased, rhetorical relations are conditioned by the thematic progression and the conceptual structure involved. Once the rhetorical relation has been recognized, the third problem is also solved (as Mann and Thompson state), which follows inter alia from the canonical order of each rhetorical relation.

At the end of this section, we provide an example analysis of a turn into thematic progression and ensuing rhetorical structure. As will become obvious from the example, recognizing the thematic progression as well as the rhetorical structure enables us to observe which parts of a turn are to be considered as relevant. The issue of discerning relevance will be elaborated upon in the next section.

(7) For Wittgenstein tonight it is, yes. For four persons is fine. But the other one doesn't know. And because it is his birthday we would like to have our picture taken. Can you ask that too? Oh yes, and my husband would like to join us for dinner if that would be possible. No foreign stuff. So that is for three. Are you also in charge of the food?

Assuming that we have decent means to analyse the dialogue linguistically, let us commence with discerning the thematic progression. The schema displays sequential progressions (**seq**) and parallel progressions (**par**).

$$
\begin{array}{llll}
T_1 & [\text{It}] & \rightarrow & F_1 \ \ [\text{Wittgenstein tonight}] \\
 & & & \|\mathbf{seq} \\
 & & & T_2 \ \ \langle\text{ellipsis}\rangle \ \rightarrow \ F_3 \ [\text{four persons}] \\
 & & & \\
T_3 & \rightarrow & \langle\ \rangle & F_3 \ \ [\text{the other one}] \\
 & & & \|\mathbf{seq} \\
 & & & T_4 \ \ \ldots[\text{his birthday}]\ldots \rightarrow \ldots \ \ F_4 \ [\text{picture}] \\
 & & & \qquad\qquad\qquad\qquad\qquad\qquad\quad \|\mathbf{seq} \\
 & & & \qquad\qquad\qquad\qquad\qquad\qquad\quad T_5 \ \text{Question} \\
 & & & \\
T_6 & [\text{husband}] & \rightarrow & F_6 \ \ [\text{to join for dinner}] \\
 & & & \|\mathbf{seq} \\
 & & & T_7 \ \ \langle\text{ellipsis}\rangle \ \rightarrow \ F_7 \ [\text{foreign stuff}] \\
 & & & \|\mathbf{par} \\
 & & & T_8 \ \ \langle\text{dinner}\rangle \ \rightarrow \ F_8 \ [\text{three } \langle\text{persons}\rangle] \\
 & & & \|\mathbf{par} \\
 & & & T_9 \ \text{Question}
\end{array}
$$

T_3 and T_6 refer hyperthematically to F_3, being "(members of) the group that is going to the performance", but we shall not consider such in the case at hand. More interesting to observe is that the thematic progression quite naturally segmentates the turn of the dialogue, as we conjectured. Let us call the three segments S_{T_1}, S_{T_3} and S_{T_6}, the subscript denoting the Topic that initiates the segment.

Subsequently, the segments can be said—quite uncontroversially, hopefully—to be rhetorically related as follows:

$$
\begin{array}{lllll}
S_{T_1} & \longleftarrow & [\mathbf{elaboration}] & \longrightarrow & S_{T_3} \\
S_{T_6} & \longleftarrow & [\mathbf{elaboration}] & \longrightarrow & S_{T_3}
\end{array}
$$

Using the canonical order noted earlier, we can consequently determine the nuclei and satellites and construct the following hierarchical organization:

S_{T_1} = Nucleus
\
[**elaboration**]
\
S_{T_3} = Satellite w.r.t. S_{T_1} / Nucleus w.r.t. S_{T_6}
\
[**elaboration**]
\
S_{T_6} = Nucleus

Apparently, it suffices to maintain only the nucleus S_{T_1} and still have a coherent and justly purposeful dialogue. As we stated already, the concept of nuclearity is important to us. It directs the response to the turn of the dialogue, which in this case could for example be that there is no performance by Wittgenstein tonight at all.

4 Relevancy and Generic Tasks

The current section will explain the fashion in which we discern *relevant information* in a dialogue, thereby building forth upon the previous section. First and foremost we should then clarify what we understand *by relevance.*

When we state that a particular piece of information is relevant, we mean that it is relevant from a certain point of view. We do not want to take all the information that is provided into consideration. Rather, we are looking for information that fits our purposes. And what are these purposes? Recall the discussion above, where the concept of generic tasks was introduced. Generic tasks were presented as units to carry out simple tasks, units which could be combined into an overall structure that would remain flexible due to the functional individuality of the simple tasks. These generic tasks are our 'purposes'.

More specifically, when carrying out a generic task, we look among the nuclei found in the rhetorical structure for one that presents us with the information that we need for performing the task at hand. In other words, such a nucleus presents us with relevant information. For example, if when carrying out the task IDENTIFY_PERFORMANCE, the following information is of importance to uniquely identify a performance:

- the name of the entertainer, the performing group, or the performance itself;
- the day (and if more performances on one day, also the time).

Obviously, the nucleus ST1 is highly relevant to this task. For it provides us with both ENTERTAINER_NAME as well as PERFORMANCE_DAY. Interesting to note is that once we have such information, a proper response can be generated by the dialogue manager. For example, the system could respond that there is no performance by the entertainer on the mentioned day, or ask (in case of several performances on the same day), whether one would like to go in the afternoon or in the evening.

Furthermore, things also work the other way around. As we noted earlier, a nucleus directs response. Therefore, a nucleus should also be regarded as a possibility to initiate the execution of a particular generic task. Such requires the following assumptions, though. First of all, a linguistic analysis should provide us with the concepts that are related to words or word-groups. Observe that this assumption has been made already above. Second, from each generic task it should be known which concepts are involved in the performance of that task. Thus, what kinds of information it gathers. It basically boils down to the following then. Namely, if we know the concepts involved, we should be able to identify the generic task that should be initiated to respond properly to the user.

It is realistic to assume that, based on all the information the user provides, several generic tasks might be invoked. Such tasks should then be placed in an order that would appear natural to the user. We must note, though, that it will not be the case that different generic tasks will be invoked based on identical information. Each generic task is functionally independent and has a simple goal, and as such works with information that is not relevant to other generic tasks.

Recapitulating, we perceive of relevance in terms of information that is needed for the performance of tasks that are functionally independent and have simple goals: The so-called generic tasks. Based on the thematic progression and the rhetorical structure, we look for information in the nuclei that we have identified. If the information found is needed for a task that is currently being carried out, or if it can be used to initiate a new task, then we consider the information to be *relevant information.*

Clearly, our system thereby no longer organizes its responses strictly to prefixed scripts nor strictly to a recognition of the user's intentions. Due to our use of generic tasks and integrated with our understanding of relevant information, our system carries out its tasks corresponding the way the user provides it with information. Thus, the system is able to respond more flexibly as well as more natural to the user.

5 Conclusions

In this paper we stated that the information we are basically interested in is relevant information, and we provided the means by which one can arrive at relevant information. For that purpose, we discussed the Praguian concepts of Topic and Focus Articulation (TFA) and thematic progression, the structure in which Topics and Foci get organized. Subsequently, we examined rhetorical structures in the light of Rhetorical Structure Theory, and showed how the rhetorical structure of a turn builds forth upon the turn's thematic progression. We identified genuine nuclei in a rhetorical structure to be potentially providers of relevant information. That is, information that a currently running generic task would need or that could initiate a generic task. We closed our discussion by noting how such leads to a system that is capable of responding to a user in a flexible and natural way.

A couple of concluding remarks could be made. First of all, in the discussion we do not treat of thematic progressions spanning over more than one turn. Currently, thematic progressions and thus rhetorical structures are bound to single turns of a dialogue. We intend to lift this restriction after examining how we can completely integrate our logical theory of information with the views presented here. Second, we would like to elaborate on how the mechanisms described here would fit into a dialogue manager that parses dialogues on the level of generic tasks.

Regarding the segmentation of discourses and its relation to the dynamics of the communication of information, a topic for further research could be to compare our point of view to that of Firbas' Communicative Dynamism as described in (Firbas 1992).

References

Allen, J. F., and Perrault, C. R. (1980). Analyzing intention in utterances. *Artificial Intelligence*, 15.

Chandrasekaran, B. (1986). Generic tasks in knowledge-based reasoning: High-level building blocks for expert system design. *IEEE Expert*.

Daneš, F. (1979). Functional sentence perspective and the organization of text. In Daneš, F., editor, *Papers on Functional Sentence Perspective*. Academia, Praha.

Firbas, J. (1992). *Functional sentence perspective in written and spoken communication*. Studies in English Language. Cambridge: Cambridge University Press.

Hajicova, E. (1993a). From the topic/focus articulation of the sentence to discourse patterns. Vilem Mathesius Courses in Linguistics and Semiotics, Praha.

Hajicova, E. (1993b). Issues of sentence structure and discourse patterns. Technical report, Charles University, Prague.

Hajicova, E. (1994a). Topic/focus and related research. In Luelsdorff, P. A., editor, *Prague School of Structural and Functional Linguistics*, vol. 41 of *Linguistic & Literary Studies in Eastern Europe*. Amsterdam: John Benjamins.

Hajicova, E. (1994b). Topic/focus articulation and its semantic relevance. Vilem Mathesius Courses in Linguistics and Semiotics, Praha.

Lambert, L., and Carberry, S. (1991). A tripartite plan-based model of dialogues. In *Proceedings of the 29th Annual Meeting of the Association for Computational Linguistics.*

Mann, W. C., and Thompson, S. A. (1987). Rhetorical structure theory: A theory of text organization. Technical report, Information Sciences Institute, Marina del Rey (CAL). Reprint.

Minsky, M. L. (1975). A framework for representing knowledge. In Wilston, P., editor, *The Psychology of Computer Vision.* New York: McGraw-Hill.

Schaake, J., and Nauta, D. (1994). Een perspectief op informatieverwerking vanuit een pragmaticistisch taalbegrip. unpublished manuscript, Enschede.

Schank, R. C., and Abelson, R. (1977). *Scripts, Plans, Goals and Understanding.* Hillsdale (NJ): Lawrence Erlbaum Associates.

Searle, J. R. (1969). *Speech Acts, An Essay in the Philosophy of Language.* Cambridge: Cambridge University Press.