

Coherence in Spoken Discourse*

Heike Tappe and Frank Schilder

Computer Science Department

Hamburg University

Vogt-Kölln-Str. 30

D-22527 Hamburg

Germany

{tappe, schilder}@informatik.uni-hamburg.de

Abstract

This paper explores the possibilities and limits of a discourse grammar applied to spontaneous speech. Most discourse grammars (e.g. SDRT, Asher, 1993; RST, Mann & Thompson, 1988) tend to be descriptive theories of written discourse which presuppose a coherent structure. This structure is the outcome of a goal directed planning process on the part of the producer. In order to obtain a better understanding of the planning process we analyse spoken discourse elicited in an experimental setting. Subjects describe the pixel-per-pixel development of sketch-maps on a computer screen. This forces the speakers to conceptualise the perceived state of affairs, plan their discourse, and produce a description of the drawing at the same time. Thus we find evidence for the planning process in the recorded data and can show that the discourse structures are less globally coherent than those underlying written text. In our paper we discuss to what extent a flexible discourse grammar based on a Tree Description Grammar (TDG) (Schilder, 1997) can handle such data.

1 Introduction

We investigate in this paper to what extent a discourse grammar is capable of analysing spontaneous speech that is obviously not as well structured as written text. The example text discussed contains questions and remarks which do not seem to be part of the discourse. Nevertheless, we believe that the entire spoken discourse is to be represented by one discourse structure. Evidence for this assumption comes from the observation that anaphoric references are made between questions which apparently

comment on the planning process and the actual description of the sketch-map.

Following Schilder (1997) a discourse grammar based on a Tree Description Grammar (TDG) (Kallmeyer, 1996) is used for the analysis of an example text. TDG is employed to encode the dynamics of the discourse structure. Other discourse theories like Segmented Discourse Representation Theory (SDRT) (Asher, 1993) or Rhetorical Structure Theory (RST) (Mann & Thompson, 1988) offer only a descriptive explanation.

The remaining part of the paper is organised as follows. Section 2 contains a description of the experimental setting in which the example discourse was obtained (Habel and Tappe, forthcoming). Section 3 provides an outline of the example before a short introduction to the discourse grammar is given in section 4. Section 5 offers the formalisation of the example discourse and section 6 concludes and describes areas for ongoing research.

2 Method and material

2.1 Method

Subjects were presented with sketch-maps. These were previously drawn by students who had been asked to sketch the route between the Computer Science Department and the main campus of their university. Since the two landmarks are approximately 6 km apart, all of the sketch-maps included some means of transport. The drawings were made on a drawing tablet and subsequently stored on a computer hard-disc. In the verbalisation-phase replays of the drawings were used as stimulus material. A new group of subjects was presented with one of the drawings. They had to carefully watch what happened and simultaneously describe what they were seeing, while the graphical objects became visible on the previously empty screen in the same chronological order they were produced. The verbalisers were familiar with the route between the two Uni-

* This work is partly funded by the German Science Foundation (DFG), research project 'Conceptualization Processes in Language Production: an Empirically Founded Model on the Basis of Event Description' (Funding Number: HA 1237/10-1).

versity buildings, yet they did not know what material they were going to be confronted with.

2.2 Material

For the present analysis we chose a fragment of one of the online-verbalisations, consisting of the first seven utterances describing the sketch-map segment that is illustrated in figure 1.



Figure 1: The sketch-map

The graphical objects in this sketch represent the following objects: the Computer Science department and the streets leading from the building. This part of the sketch-map is described by a 32 year old, right-handed computer scientist.

3 Analysis of the text fragment

The text fragment contains a variety of features that are characteristic of spoken rather than of written discourse. In this section we will look at each of the utterances in greater detail and show how the discourse coherence is maintained by the speaker. He starts talking as soon as he sees the rectangle being drawn on the screen. The first utterance (U1) can be characterised as a statement about the speaker's current mental state:

U1: Ja, ich weiß ja schon worum es geht, (*Yes, I already know what this is all about,*)

The speaker hereby expresses a self-belief the content of which can be circumscribed as follows: I (the speaker) know which states-of-affairs I am about to see on the computer screen. This utterance serves as a kind of background for what follows. With his statement, the speaker commits himself to prove that he really knows what is going on. With the subsequent utterance (U2) he demonstrates that he has at least some intuition about the stimulus material: He assigns the rectangle the name of the depicted real world object.

U2: also das wird das Informatikgebäude ... mit der Beschreibung daneben. (*well, this is going to be the build-*

ing of the Computer Science department ... with the annotation next to it)

Accordingly, he fulfills part of the felicity conditions that accompany assertions about the possession of knowledge, i.e. he elaborates on the content of his belief-state. The *elaboration*-relation between (U1) and (U2) is triggered by the discourse marker *also*. With the next utterance the speaker adds further information to his states-of-affairs-description.

U3: und die Straßen die jetzt angefangen werden zu malen... (*and the streets that are now started to be drawn...*)

Therefore, we can categorise the relation between (U2) and (U3) as a *narration*-relation. This relation does not add a new perspective or a new theme to the ongoing discourse, but rather supports its continuation. On contrast, (U4) establishes a break in the ongoing discourse. The discourse marker *eigentlich* signals that the speaker has build up an expectation about the continuation of the drawing event on the basis of his belief state.

U4: Eigentlich würde ich erwarten, (*Actually I would expect*)

The content of the belief state is — as mentioned before — that the speaker believes to know what will be drawn. Yet, this belief state ends here, because even though the speaker rightly interprets the developing double lines to represent streets (cf. U3 above) his further expectation is not met. The content of this expectation is expressed in (U5):

U5: daß irgendwo die Bushaltestelle noch eingezeichnet wird, da im... (*that the bus stop was drawn into it somewhere, there in the...*)

Obviously the speaker expects that the drawing will contain a symbol representing a bus-stop near to the building. This is not the case. Therefore the rhetorical relation between (U4) and (U1) is that of a *termination*. We see that rhetorical relations do not necessarily hold between adjacent utterances only, but that an utterance may open a subtree that can be closed off by an utterance that is verbalised a couple of utterances later. (U5) breaks off with a prepositional phrase that lacks the location argument (...*da im...*). The speaker is quite obviously insecure about the name of the street that contains the bus-stop. (U6) reveals his insecurity.

U6: (a) wie heißt das Ding, heißt das Gazellenkamp? (b) Ja, ne? ... (*what is it called, is it called Gazellenkamp? Yes, isn't it? ...*)

The structure in U6 is very typical for spoken dis-

course. It is not in a strict sense part of the ongoing discourse, but the verbalisation of vocabulary search and planning processes. We hold that the interrogative intonation functions as a signal, allowing the integration of a substructure that is not connected to the previous discourse via a prototypical rhetorical relation. The substructure itself can be interpreted as a meta-comment about the ongoing mental processes. This substructure is closed off by (U7) which begins with *aber* ('but').

U7: Aber ... keine Bushaltestelle (*But ... no bus stop*)

This discourse marker allows the speaker to return to the branching node of the discourse structure where the digression was introduced.

4 Discourse grammar

4.1 Tree descriptions

A definition of TDG is given by Kallmeyer (1996) who introduces *tree descriptions* consisting of constraints for finite labelled trees. A dominance relation (\triangleleft^*) between node labels indicates that these two labels can be equated or have a path of arbitrary length inserted between them. The second relation between nodes is the parent relation (\triangleleft) which is irreflexive, asymmetric and intransitive.

The tree's root node **D** labelled k_1 in figure 2, for example, dominates another node labelled k_2 . According to the definition of \triangleleft^* these two nodes may be equal or an arbitrary number of other nodes may be in between them. An adjoining operation

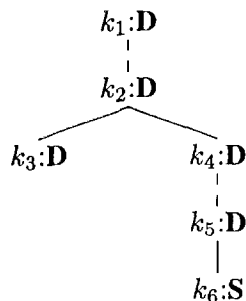


Figure 2: A labelled tree description

is easily defined because of this property. Further *tree descriptions* can be inserted between such nodes. The descriptions which are, formally speaking, negation-free formulae of constraints on the nodes, are conjoined. The nodes where the adjunction takes place are set to equal.¹

¹Figure 3 shows an example.

4.2 A flexible discourse grammar

According to Schilder (1997), feature value structures are added to the tree logic in order to enrich it with rhetorical relations and further discourse information. One non-terminal symbol is used for the **D**(iscourse) segments, whereas the terminals are the **S**(entences).

Two features are added to the tree description to encode the semantic content of the sentence and the 'topic' information expressed in a discourse. Firstly, **S** gets associated with the meaning of a sentence via a feature **CONT**(ENT) containing all discourse referents and the conditions imposed on them.²

Secondly, a feature **PROMI**(NENT) is added that is used to define the notion of openness within a discourse. This feature reflects the fact that one situation described by an utterance (e.g. situation e_1 described by U1) is subordinated by another one when combined via a rhetorical relation. It furthermore exhibits the restriction of the further utterances to the *right frontier* of the discourse tree (cf. (Webber, 1991)).

For the discourse structure two types of tree descriptions have to be distinguished. One tree structure allows attachment on two levels of the right frontier of the tree. This tree is called *subordinated tree* and the structure is schematically indicated in figure 2. The other one is a *subordinating* structure that is triggered by discourse relations such as *narration* or *result*. Further attachment is only possible at the last uttered sentence.³

5 Formalisation

The discourse structure obtained for the first three sentences of the example text is reflected in figure 3. At first an *elaboration* relation is established between (U1) and (U2). The imposed discourse structure (i.e. a *subordinated* tree as in figure 2) allows attachment at two levels. Note furthermore that the *elaboration* relation holds between the mental state of the producer (i.e. *I already know what this is all about*) and the description of what is happening on the screen.⁴

(U3) is connected with (U2) via *narration*. The adjunction operation in figure 3 shows how the

²We presume that this content is represented by a discourse representation structure as standard DRT would predict (Kamp and Reyle, 1993).

³See the right tree in figure 3.

⁴These rhetorical relations are underlined in the figure to highlight their different status.

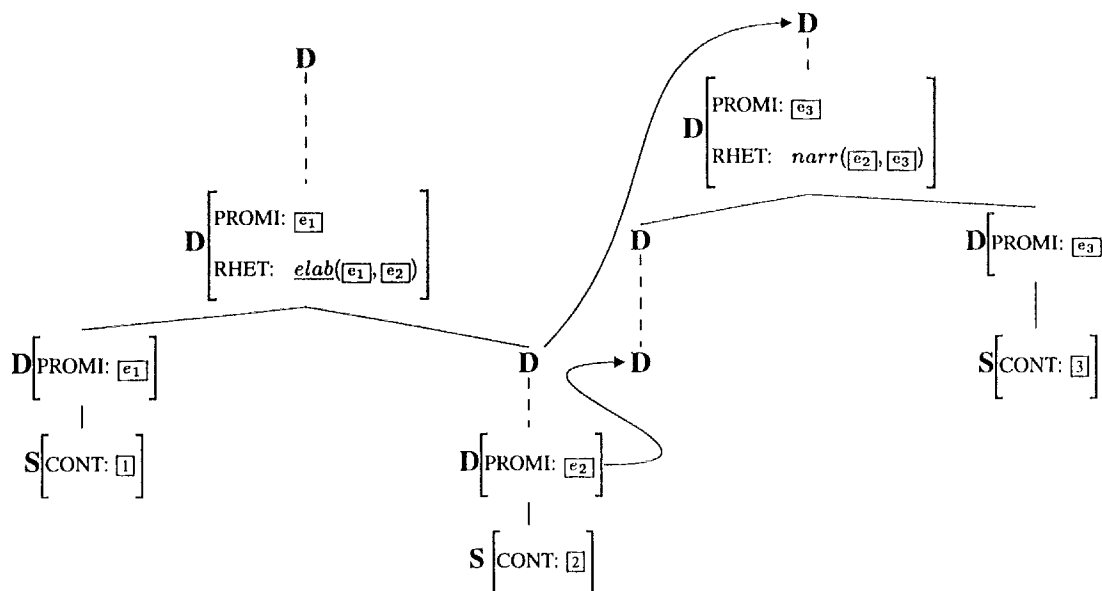


Figure 3: Two discourse segments combined

newly generated sentence is incorporated in the current discourse structure.

Although the production took place under a certain amount of pressure, the right frontier principle was never violated. The speaker never went back or made anaphoric references to discourse referents being behind this *frontier*.

Having demonstrated how the production of the discourse structure can be formally described for the first three utterances, we now want to focus on a particularly interesting problem exhibited by the sequence (U4) to (U6). This sequence contains rhetorical questions, which describe the ongoing planning process of the speaker.⁵

The sequence starts with an expectation (i.e. (U4)) the subject utters. Again the proposition expressed is related to the mental state of the speaker. Interestingly enough, he has to return to the top level of the discourse tree and continue from there. Consequently, the discourse segment containing (U2) and (U3) is 'cut off' and not available for further attachment.

Embedded within the expectation is an utterance describing the ongoing planning and searching process. The verbalised questions reflect the request to the mental lexicon and the mental map the subject has got of this area.

The discourse grammar consequently has to be

⁵Note that such a sequence would never be found in a written text.

extended in order to maintain a coherent discourse structure for the modelling of the *producer*. Thus rhetorical relations describing planning processes are introduced. With these, the discourse grammar becomes capable of representing a coherent discourse structure for the spoken language despite the fact that the entire discourse segment does not seem as coherent as written text.

Figure 4 contains the discourse structure after the search for the street name has come to an end. One rhetorical relation introduced is *p(lan)_comment* which describes the ongoing planning process. It also involves a search for the correct word in the lexicon. The rhetorical *quest(ion)* is asked whether the correct word has been chosen and this question answered by the subject. The summarising *yes, isn't it* (i.e. (U6b)) ends the search process and closes the discourse structure at the right frontier.

Interestingly enough, the clue given by the discourse marker *but* uttered in (U7) is absolutely essential. The speaker indicates with this marker that he wants to return to the top level of the discourse tree and to add a *contrast* relation to the expectation. The construction of the discourse structure continues therefore at the top level of the tree in figure 4.

6 Conclusion

We have shown that spoken discourse can be formalised by a discourse grammar based on TDG. Even planning processes that surface as rhetor-

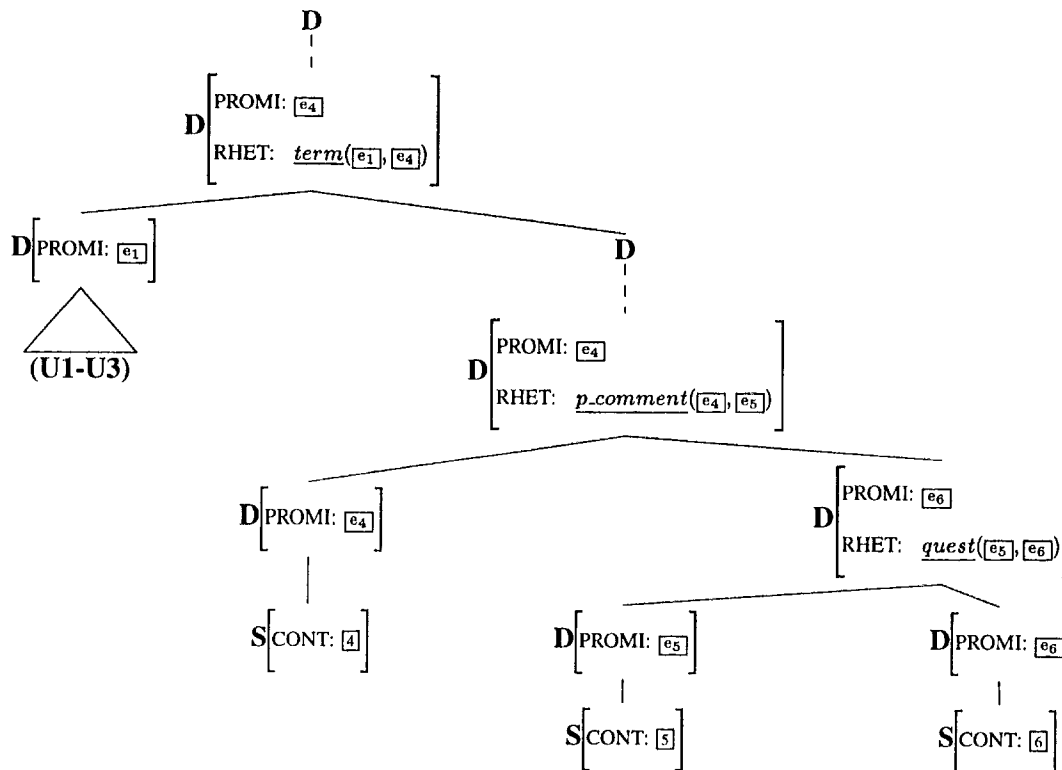


Figure 4: The planning process within the discourse structure

ical questions can be incorporated into the discourse structure generated. New rhetorical relations were introduced that should prove useful for NLP-applications. In ongoing research we focus on the interaction between planning sequences, discourse structure and intentional structure.

References

- Nicholas Asher. 1993. *Reference to abstract Objects in Discourse*, volume 50 of *Studies in Linguistics and Philosophy*. Kluwer Academic Publishers, Dordrecht.
- Christopher Habel and Heike Tappe. forthcoming. Processes of segmentation and linearization in describing events. In Ch. von Stutterheim and R. Meyer-Klabunde, editors, *Processes in language production*.
- Laura Kallmeyer. 1996. Underspecification in Tree Description Grammars. Arbeitspapiere des Sonderforschungsbereichs 340 81, University of Tübingen, Tübingen, December.
- Hans Kamp and Uwe Reyle. 1993. *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language*, volume 42 of *Studies in Linguistics and Philosophy*. Kluwer Academic Publishers, Dordrecht.
- William Mann and Sandra Thompson. 1988. Rhetorical structure theory: Toward a functional theory of text organisation. *Text*, 8(3):243–281.
- Frank Schilder. 1997. *Temporal Relations in English and German Narrative Discourse*. Ph.D. thesis, University of Edinburgh, Centre for Cognitive Science.
- Bonnie L. Webber. 1991. Structure and ostension in the interpretation of discourse deixis. *Language and Cognitive Processes*, 6(2):107–135.