

Referential Movement in Descriptive and Narrative Discourse

Christiane von Stutterheim

Institut für Deutsch als Fremdsprachenphilologie
University of Heidelberg

and Wolfgang Klein

Max-Planck-Institut für Psycholinguistik, Nijmegen, Netherlands
and Institut für Deutsch als Fremdsprachenphilologie
University of Heidelberg¹

1 Introduction

A witness, when asked "What did the robber look like?" may choose to answer this question simply with a single utterance or even a part of a single utterance, for example with

- (1) He looked like Sidney Greenstreet.
- (2) Like Sidney Greenstreet.

What he is asked for by the question, is a description—a set of predicates, visual properties in this case, of a person he has observed. Both answers fulfill this task—they give a description which, though only a summary, is satisfactory for anybody who knows what Sidney Greenstreet looks like. More than this, the first utterance also repeats what the task is about—to describe the robber—although it does not do more to satisfy it.

Alternatively, the speaker may choose to be a bit more explicit with respect to the predicates he or she is asked to specify, and may answer the question by means of a whole series of utterances—i.e., with a text:

- (3) Well, I saw him for only a few seconds. He was a young man, about twenty or so. He was wearing dark trousers, a dark jacket and a brown hat. He was very slim. He wore a mask, so I couldn't see his face at first. But then, when he slung the plastic bag over his shoulder, the mask slipped a bit. He put it back immediately, but for a brief moment I could see his face. He had a thin, black moustache.

There are two important things to notice about this fuller way of answering the question—that is, distributing the information asked for over a series of interconnected utterances. First, not all utterances are directly "to the point", which is, here, to specify some visual properties of the robber. The first utterance, for example, qualifies the reliability of the description. Clearly, this may be relevant and important, but it does not directly serve to answer the question. Later, part of the robber's action is narrated. Again, this is *related* to the description—it gives some background—but is not part of the description itself. We shall call those utterances in the text which directly contribute to answering the question, its *main structure*, and those which give additional—and often important—material, its *side structures*².

What counts as main structures, and what as side structures, clearly depends on the specific question which the text as a whole is produced to answer. In the above example, the speaker is asked to specify a set of predicates which constitutes a description of the robber; everything else—for example narrative sequences—is part of the side structure. Had the question been "What happened?", the categorization as main or side structure would have been the reverse of the above. In that case, the speaker's task would have been to specify a sequence of events, which in its entirety constitutes a robbery (to the extent that he witnessed it); and this "plot line" would then comprise the main structure of the text, its "foreground". Any other material—for example the description of the robber—would then be part of the side structure, or "background", of the text. These familiar concepts, "foreground" and "background", are usually distinguished not in this way, but rather with reference to whether or not they serve to advance a "narrative sequence" (see, for example, Hopper 1979, Reinhart 1985). It should be clear, however, that the background/foreground partitioning of narrative texts is only a special case of a much more general phenomenon, found also in descriptive and in argumentative texts, to which a definition in terms of "narrative sequence" does not apply. Hence, we shall avoid it here. It should be noted, however, that this definition is not inconsistent with the approach taken here; it is merely a special case of it.³

The second feature of (3) is, that in the course of distributing the answer across several utterances, certain meaning components within each utterance are maintained from the preceding utterance (or utterances), while other parts of its meaning are freshly introduced. Thus, when, as in this example, there is one person referred to, the robber, he is introduced in the question itself and

then merely maintained in all subsequent utterances of the main structure. Had there been more than one robber, the related question, "What did the robbers look like?" would have forced the speaker into a somewhat more complicated pattern of referent introduction and maintenance (cf. Marslen-Wilson, Levy and Tyler 1982, Hickmann 1980). Such differences would arise despite the fact that the nature of the question is the same: What is requested of the speaker in both cases, is that he specify visual properties, rather than, for example, events.

Persons are not the only meaning elements which must be introduced or maintained across utterances. Time and place, for example, are kept constant throughout the text, more precisely: throughout its main structure, although place is never explicitly referred to, and time only vaguely by the tense of the verb. This is quite different in other kinds of text. Narratives, for example, typically imply a change of time from one main structure utterance to the next, and often changing places as well. Descriptions of entities larger than a human being may include various places, which then must be explicitly referred to in the appropriate utterances, and descriptions of processes ("How to eat a lobster.") require a temporal sequencing. To summarize, a coherent text, produced as an answer to a question, involves a *referential movement* within various semantic domains—we shall say here *domains of reference*—such as persons, place, time, and others. This referential movement is reflected in the use of specific linguistic devices such as, for example, the use of anaphoric elements rather than indefinite noun phrases, and in the choice of a particular word order, intonation contour, etc.

We can now state the main thesis of this paper: The structure of a text is constrained on both global and local levels by the nature of the question which the text in its entirety is produced to answer. These constraints include:

1. The partitioning of the text into main structure and side structures;
2. the "filling" of various possible domains of reference within each main structure utterance;
3. the assignment of specific meaning components to the topic or to the focus (defined below) of each main structure utterance;
4. the referential movement within the domains from one main structure utterance to the next.

In the following section, we will explain these notions in more detail. Section 3 deals with the specific constraints on the main structure of descriptive versus narrative discourse. Finally, section 4 presents supportive data taken from picture story retellings which combine descriptive and narrative passages.

2 "Quaestio" and Constraints on Text Structure

2.1 Topic and Focus

An utterance such as

- (4) Ruben rang at five.

can answer different questions (though with different intonation contours, as discussed below), for example

- (5) Who rang at five?
 (6) What did Ruben do at five?
 (7) When did Ruben ring?
 (8) What happened?

In each of these cases, it settles an open alternative raised by the question—it specifies one out of a set of candidates at issue. In (5), the alternative to be settled (alternative here being the choice between two or more possibilities) contains those persons who could have called at five, and selected and specified was Ruben. In (6), the possibilities included all those actions which Ruben could have performed at five, ringing being the one specified. In (7), the choice was between the (contextually plausible) time spans during which Ruben could have rung; and five o'clock was the one specified. In (8), the alternative included all (contextually plausible) events that could have occurred on some contextually given occasion, and the event chosen by the speaker was Ruben's ringing at five. Such an alternative at issue will be called the *topic* of the utterance, and the element from that alternative which is chosen and specified, the *focus* of the utterance.

Three points are important here. First, topic and focus, as conceptualized here, are components of the *meaning* of an utterance—for example a person, an action, a time span—not the *means* by which this meaning is expressed—for example a constituent or a series of constituents. If there is need, we will speak of topic expression and focus expression, in contrast to topic and focus. For example, in (5), the focus is person Ruben, and the focus expression is the name "Ruben". The same focus could also be expressed under certain contextual conditions by the word "he" or by some other noun phrase referring to the same person.

Second, we must distinguish sharply between the expression of a topic or focus, on the one hand, and the marking that this entity *is* topic or focus on the

other. In (5), the focus is the person Ruben, and the focus—the person Ruben—is expressed by the word "Ruben". The fact that this entity is focus (and not topic, or part of the topic, as in (6) is marked by intonation. Intonation is not the only device available to make clear what the focus (or the topic) is; word order or specific particles also serve as devices for this, at least in some languages. Very often, however, it is not explicitly marked at all, or is ambiguous.

Third, the distinction between topic and focus must not be confused with that between "given" or "maintained" information on the one hand and "new" or "introduced" information on the other, although these dichotomies may often coincide. Consider, for example, a sequence of utterances such as the following:

- (9) There go Peter and Mary. Which of them do you like more, him or her?
Him (I like more).

The alternative raised is between Peter and Mary, and the focus of the answer is Peter. But obviously, Peter has already been introduced, just as everything else in the answer; the rest of the answer is a repetition of the expression of the topic.

In all examples considered so far, the utterance under discussion was an answer to an explicit question. Obviously, this need not be the case. But we may always imagine an implicit question—the *quaestio*—which the utterance is meant to answer. This *quaestio* can be made explicit, as in the examples above. It may also result from the general context; for example if we hear a loud noise from outside the house and someone says "Ronnie is trying to park the car", then this utterance answers a *quaestio* which could have been made explicit by questions such as "What's that?" or "What happened?"

The *quaestio* answered by an utterance may also be derived from the "higher" *quaestio* of a whole text, to which the utterance belongs and which the text in its entirety is intended to answer. This is exactly the case which we discussed in section 1. Take a question like "What happened to you last week?" which elicits a narrative text. It asks for the specification of some complex event, which the speaker may subdivide into a series of sub-events, each of them happening during some time interval t_i within the time span of the total event. Thus, the *quaestio* of the whole text can be broken down into a temporal sequence (or temporally-ordered series) of *quaestiones* answered by all of those utterances which specify one of the subevents, roughly

- Q₁: What happened to you at t_1 ?
Q₂: What happened to you at t_2 ?
Q₃: What happened to you at t_n ?

This gives us in a straightforward way the "backbone" of the narrative, its main structure. It may be interrupted at any point by utterances which do not answer the overriding *quaestio*, "What happened to you at t_i ?", but rather

such questiones as "How did you feel?" , "What's the moral?", "Why did you do that?", "What did the room look like?", etc., which all contribute to side structures of different types. A particular side structure may also extend across several utterances. For example, a narrative sequence can be interrupted by a descriptive sequence, or vice versa. This is precisely the case in the example to be discussed in section 4. Consider now the quaestio "What happened to you at t_3 ?" It specifies a time span, which is after the time span of the previously mentioned event. It also specifies a person, the addressee in this case ("to you"), and moreover it indicates that what is asked for is an event, not a state, for example, as would be the case with the question "What was the situation at t_1 ?" Hence, it narrows down the focus of the answer to include only those events which could have happened to the addressee during the intervening time since the previously reported event. The utterance which answers the question, and settles the alternative raised, can repeat the topic, either fully or in part; it can also elaborate on it. But it must also contain a part which expresses the focus, that is, which specifies an event that meets the conditions mentioned above. This has many consequences for the structure of the utterance. If, for instance, t_3 is expressed by an adverbial (ADV), the protagonist by a noun phrase (NP), and the event in the narrower sense by a verb (V), and if the language in question has a rule "topic expression before focus expression", then a word order such as NP V ADV is impossible: V must be last. This may conflict with purely syntactic constraints on word order, such as "V in second position", and different languages have found different ways to solve competing requirements of this type: intonation, special particles, cleft constructions, passives, etc. or by simply not requiring consistent marking of what is focus and what is topic in these cases (cf. Li and Thompson 1976, Givon 1985).

What has been said here about main structure utterances within narrative texts applies analogously to other text types, for example descriptive texts. In that case, the focus is not an event which happened to a protagonist, but rather, for example, specification of some spatial location, of some visual property, as in (3) above (cf. for a discussion of different text types: van Dijk 1977, Werlich 1979, Rehbein 1984).

2.2 Domains of Reference within an Utterance

Before turning to the constraints which the quaestio imposes on the "filling" of various domains of reference within an utterance, it is necessary to have a brief look at the interplay of linguistic meaning proper and contextual information within an utterance.

A speaker who, at some occasion, utters

- (10) He swallowed a frog

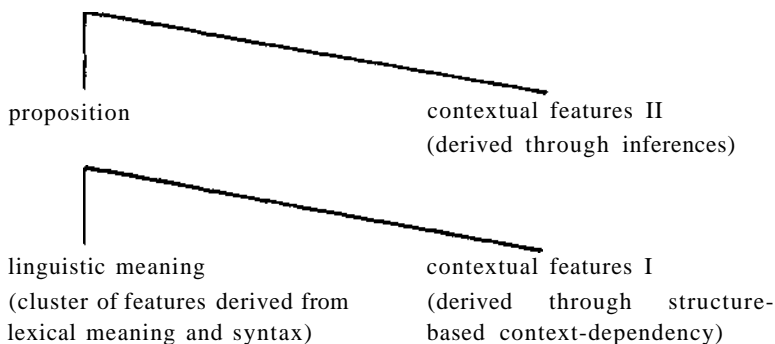
expresses a certain meaning which results from the lexical meanings of the individual words and the way, in which they are put together. (This, admittedly, makes a long story short.) The hearer may combine the meaning given by the *linguistic expressions* with other information available to him or her, e.g. from previous utterances or from situational perception, in general: from contextual information. It is useful to distinguish two ways in which contextual information is applied to complete the utterance above and beyond what is made explicit by linguistic means.

First, there is contextual information which is directly linked to context-dependent verbal means in the utterance, such as deixis, anaphora, ellipsis. The interpretation of an utterance such as *Me, too* is based on the knowledge of the meaning of these two English words and the rules of ellipsis in English, on the one hand, and on having access to the necessary contextual information, on the other (roughly, the hearer must assume the speaker was able to identify who is speaking, and had heard the immediately preceding utterance). In these cases, we may speak of *structure-based context-dependency*. By integrating the purely linguistic meaning with the information derivable from structure-based context-dependency, the listener can achieve a first interpretation, which we will call the *proposition* of the utterance. In addition, the listener may infer, with varying degrees of certainty, other features of the actual incident described in (10), for example that "he" is now less hungry than before. This inference is not directly linked to structural means but more globally related to the proposition. Therefore, *inference* in this sense is less accessible to systematic linguistic analysis than structure-based context-dependency; but it is not less important for communication. We shall say that the proposition in combination with various types of inferences leads the listener in a given case to an *utterance interpretation*.

So, when talking about the meaning of an utterance such as (10) for a particular listener in a particular situation we must distinguish (at least) three progressively inclusive levels of interpretation, each of which is in turn suggestive of a complex cluster of various temporal, spatial, personal, and other features:

1. the *linguistic meaning*, based on the lexical context of the words (or morphemes) and the rules of syntax (including, for example, case marking)
2. the *proposition*, where structure-based information is added
3. the *utterance interpretation*, which enriches the proposition by means of inferrable information of various types. Obviously, the utterance interpretation may vary from listener to listener, depending on his particular state of knowledge [relationship to the speaker, past communicative history, etc.].

(11) utterance interpretation



Two additional comments are in order here. First, the same distinction also applies to constituents smaller than full utterances for example noun phrases: they also have an intrinsic linguistic meaning, a second level of interpretation where structure-based context information is added (I above), and a third inferentially enriched "full" interpretation (II above).

In what follows, we are mainly concerned with meaning entities on the second level, and we shall call them "referents", regardless of what kind they are: persons, places, times, actions, or even modalities. This is simply a terminological convention. We could also say "meaning entities based on linguistic meaning proper and on structure-based contextual features".

Second, all of these meaning levels are abstract representations, very different from the "real" action, in this case the swallowing of a frog by someone, or whatever the matter spoken about may be. Suppose the speaker of (10) has really observed such an incident at some occasion and renders it correctly by (10). Then (10), by its linguistic meaning, gives only a very selective picture of this incident, and regardless of what may be added to this picture by contextual information—it must still remain a partial picture. And obviously, there needn't even be a counterpart in reality, at all: (10) could relate a fictitious event. The three meanings on different levels arise in the mind of anyone who knows the language, the specific context and the world in general.

Any proposition, then, is a web of "referents" (in the sense explained above) of various kinds—spatial, personal, modal etc. Ever since Aristotle, there have been innumerable proposals to categorize various domains of referents in this sense. We shall assume here, without further discussion, that there are five of them:

(12) 1. T: time spans

2. L: places

3. P: persons and objects involved (if such an entity plays a prominent role in the text, we will call it a "protagonist")
4. A: states, events, properties, in brief "predicates"; clearly this is a somewhat heterogeneous class which includes the (contextually enriched) meaning of expressions such as "to sleep", "to explode", "to be green", "to be on the table", etc.
5. M: modalities, such as the "possible", "real", "necessary", etc.

The problems with this Gordian categorization are obvious; but they should not concern us at present.

Any utterance selects referents from these domains and integrates them into a whole, the proposition. This is not done arbitrarily; rather it follows certain principles. Most typically, an element from P is combined with an element from A to form the "inner core" of the proposition; this inner core is then located in a situational frame—by reference to a time span and a place—and finally, the resulting "periphery" provides it with a modality. Obviously, this elementary picture can be complicated in various ways. Not all domains of reference must be represented; it does not make much sense to associate a proposition such as the one "Two plus two makes four" with a place. On the other hand, any particular domain of reference may be represented several times; one referent from r_1 may be used to identify a protagonist (the guy to my left), while another referent from r_1 may be used to situate the whole event. In any case, the result is a more or less complex structure of referents from these five domains.

The crucial point in the present context is, that the ways in which the proposition is "filled" with representations from the various domains of reference is constrained by the quaestio of the text. The above-mentioned quaestio, "What happened to you at t_1 ?" requires that the referent given in P be the protagonist; the referent given in A, an event; and the referent given in T, a definite time span in the past, etc. In actual fact, the constraints are much more complicated; we shall discuss constraints for narrative and descriptive texts in section 3 below.

2.3 Referential Movement

Referential movement is the way in which information from the five domains of reference shifts from one utterance to the next. Suppose, for simplicity's sake, that in two subsequent utterances S_i and S_j , the proposition consists of just one referent from each of the five domains; these referents are labelled here M_i , T_i , L_i , P_i , A_i , and M_j , T_j , L_j , P_j , A_j . This gives the following picture of possible referential movements.

$$(13) \quad \begin{array}{cccccc} S_i & M_i & T_i & L_i & P_i & A_i \\ | & | & | & | & | & | \\ S_j & M_j & T_j & L_j & P_j & A_j \end{array}$$

Consider, for example, the transition from T_i to T_j , the transition from the time span referred to in S_i to the time span referred to in S_j . It may be that the time span remains the same, although the *expression* used to refer to it varies. In this case, we speak of *maintenance*. It may also be that T_j is different, but it is to be temporally brought into relation with T_i . This is a case of *shift*, as typically expressed by such expressions as "then", "the next day", "some time later". Both types of referential movement contrast to cases where there is no connection to a previously introduced time span. In this case, we speak of *introduction*, as in the expression "On a summer morning of the year 1987". This simple typology of referential movement can be refined in various ways.

The quaestio of the text imposes constraints on which kind of referential movement is possible across main structure utterance. For example, the quaestio of a narrative requires that T_j be after T_i ; this is sometimes called the "principle of chronological order" or "principle of natural order" in narratives. Similar constraints hold for other domains of reference. Again, these constraints are quite different for descriptions of other types. There is no "principle of chronological order" governing the description of various visual properties in response to a question such as "What did the robber look like?" The structure of this type of text is governed by other constraints, however, as we shall see in the next section.

3 Main Structure Constraints in Narrative and Descriptive Texts

The quaestio, this is the general idea, imposes three types of constraints on main structure utterances: It constrains the "referential filling", i.e., how the various domains of reference can be represented in the utterance; it constrains the "referential movement", i.e., what can be, or has to be maintained or newly introduced and finally; it constrains what has to become topic element and what focus element of an utterance. Some examples of these constraints have been given already. In this section, we shall discuss them more systematically for narrative and for descriptive texts. Narrative texts are those which are meant to answer a question such as "What happened to p at t ", where p is some protagonist (or protagonists)—for example, the addressee—and t some time span, during which the entire narrated incident occurred—for example "last weekend" or "when you first met this woman". Obviously, there is some variation, depending on the precise nature of p and t . And the answer to the quaestio may also vary in several respects; in particular, it may include elaborate side structures, e.g. descriptions, evaluations, arguments, etc. But here, we shall be concerned only with the main structure. Descriptive texts, in contrast, are much more variable, depending on the entity to be described. They may answer questions such

as "What did the robber look like?", "How does an electronic flash attachment work?", and similar ones.⁴ In these cases, the entire "mental representation", which the description actually only partly renders, is differently organized: it differs with regard to the objects, which make it up, as well as with regard to the temporal, spatial, and functional relations which exist among these objects. For present purposes, we shall simply distinguish between two types of descriptions—descriptions of static arrangements and descriptions of processes, i.e. dynamic descriptions.

We shall differentiate between *absolute* and *relative* constraints. Absolute constraints are categorical and pertain to only one domain of reference, for example: "the narrative main structure utterance must contain reference to a definite time span situated temporally before the time of speaking". Relative constraints are conditional and most often concern relationships involving several domains of reference simultaneously, for example: "If referential domain X is filled by a referent of type a, then referential domain Y has to be filled by a referent of type b". Some of these relative constraints embody conceptual or logical necessities, such as: "If the person referred to is maintained and a new spatial referent is introduced, then a new time span must also be introduced". In what follows, we shall mainly consider absolute constraints (3.1); relative constraints will be briefly discussed in section 3.2.

3.1 Absolute Constraints

Modality

Let us begin with the referential domain "modality", and keep in mind that "referent", as used here, is only a meaning component based on linguistic meaning proper and structure-based contextual features. Modality, put in very general terms qualifies the relation between proposition and "reality". Actually, the notion covers two distinct, though related, concepts. One concerns the validity of the proposition with respect to possible worlds; this leads to categories such as logical necessity (true in all possible worlds), possibility (true in at least one possible world), real (in a distinguished, the "real" world), fictitious (true in a world which is not the "real" one), and maybe others. The other concept of modality includes the speaker's point of view on the validity of the proposition, leading to notions such as "likelihood", "certainty", "desirability", and others. This level of detail is sufficient for present purposes (for more extended discussion see Lyons 1976, Palmer 1987).

Main structure utterances of narratives require one type of modality, and that is the modality relating to validity in the "real" world. This is true for both "real" real-world narratives, which are typically deictically anchored, and for fictitious stories, in which the story-world is treated as if it were real.⁵ This constraint on referential filling implies that in this domain the only possible

type of referential movement between main structure utterances is "maintenance". Any "shift" leads to side structures, such as comments, evaluations, explanations. Finally, modality belongs to the topic component. Utterances such as "This was almost certainly the case", where the marking of modality is in focus, are side structures.

In a descriptive text, be it of the static or of the dynamic type, the constraints on modal reference look very similar, except for the first constraint: it need not be the real world, nor need it be presented as the real world, as indicated, for example, by the generic use of present tense.

Temporal Domain

In the temporal domain, to which we now turn, there are much clearer differences between the two types of text. The quaestio of a narrative demands that the temporal referent be a specific time interval (contained in the time of the total event talked about). This time interval need not to be as precisely identified as, for example, "Last Sunday, from three to four o'clock", although this is often the case. Less precise expressions such as "once", or "about three years ago, in late autumn" are often sufficient. The fact that the temporal referent is a distinct time span, with a left and a right boundary (even when not precisely identified) allows it to be used as an anchoring point from which to move on to the event presented in the subsequent utterance. Therefore unbounded states, habituals, generics and also negated propositions are generally excluded from the main structure of a narrative (with some interesting exceptions). Referential movement can't be of the "maintenance"-type. Typically, it is a shift: the time span referred to is a new one, but it is related to the time span of the event reported in the preceding utterance in that it takes place afterwards. This is referred to as "the principle of chronological order", which is often considered to be the defining criterion for "foreground utterances" in narratives.⁶

Next we turn to the topic-focus condition. In narratives, temporal reference belongs to the topic. Utterances like "This happened yesterday" are side structures, since the temporal reference is in focus.

Compare this now to descriptive texts. Here, we get a very different picture, which may vary according to the type of description. In the description of objects, specific temporal reference is normally excluded, temporal location on the time axis would lead to side structures. Temporal reference remains constant throughout the text,⁷ and it is part of the topic. Dynamic descriptions, on the other hand, typically include a shift (still as part of the topic), but in contrast to narratives, they do not refer to specific time spans. A typical quaestio would be "What do you usually do when you get up in the morning?", which leads to a description of temporally ordered routines.

Spatial Domain

Turning to the referential domain of space now, the only constraint on the "filling" in narratives is that main structure utterances cannot have a generic spatial reference. In principle, the event reported must have a definite spatial referent, although this referent need not be precisely identified in the utterance in question. There are no constraints on referential movement: the place can be maintained, shifted or introduced. It can be part of the topic, but also part of the focus; it can't constitute the focus, however; so, it is possible to have "Then, he jumped into the river", with "in the river" being the new place introduced as part of the focal event "jump into the river"; an utterance such as "This happened on the river" would be a side structure, however.

In descriptive texts, there are no general constraints. Depending on the *quaestio*, either a specific or an unspecific spatial referent may be required, referential movement can be of any type, and the place can be part of either topic of focus; it may even constitute the focus: "the book is on the table" is clearly a spatial description.

In many types of descriptions, spatial reference plays an important role for structuring the underlying conceptual representation. In texts produced to answer a question such as "What do you see?" (while looking at a picture or some physical object), the spatial domain of reference will play the same role as the temporal domain does in narratives. It provides the thread for the sequential organization of the information, reflected in a shift in spatial reference from one utterance to the next (cf. Linde and Labov 1975, Weissenborn 1984, Ehrich and Koster 1983).

Reference to Persons and Objects

The next domain of reference, persons or objects, is a special case insofar as it is most often represented by more than one referent in the proposition. We will begin with the case where there is only the subject, and then consider further referents.

The *quaestio* of a narrative does not impose definite constraints on this domain. There are, however, strong tendencies toward a typical narrative pattern, as numerous studies have shown (cf. Givon 1983). Since narratives report singular events in a spatio-temporal frame, reference is usually made to one specific person, the narrator, or to a group of persons—the protagonists of the story. Generic reference is not allowed.

As for referential movement, maintenance of *p* is a major device for establishing coherence. Often, the events reported are brought into a meaningful relation with reference to their relevance to the actor or experiencer. This means that reference to person is maintained in main structure utterances, a fact which is shown in many studies on "topic" continuity in discourse, the term "topic"

being taken here in a narrower sense: it is a special case of the concept of topic as used here.

Now, having only one referent from this domain in subject position is but the simplest case. If the narrator wants to introduce an additional protagonist, this is typically not done in subject position at first. Rather a new referent tends to be introduced as a participant in the same event as that involving the already-introduced protagonist. In such circumstances, the new referent is part of the focus, most often in form of the grammatical object, as in "Suddenly he saw a little girl". In subsequent utterances, references to this new person may then switch position and grammatical function, and become part of the topic: "she ...". Another possibility is to introduce a new person via a side structure, where again she is part of the focus such as in "Then there was a little girl. She ...". It is also possible to introduce a new referent directly in the subject position of a main structure utterance, replacing the protagonist of the preceding event. But in such cases there are special constraints on the referential structure of the rest of the utterance: it has to be linked to the preceding utterance in the spatial domain or in the predicate domain (discussed below) as, for example, in the sequence "He got on the bus. A fat man followed him". This is a case of a relative constraint.

These examples show that there are different possibilities for introducing a new person or object and that these different possibilities have consequences as to *what* can be presented in the remainder of the utterance and *how* this information will be linguistically encoded. Thus, the choice of the particular anaphoric device in German—*er, sie, es* vs. *der, die, das* vs. zero anaphora, all of them indicating maintenance—is crucially governed by whether the maintenance goes from topic to topic, from focus to topic, or from topic to focus.

Descriptive texts vary considerably in the constraints on this domain of reference, depending on the precise nature of the quaestio. In the *static type*, the referent may be specific or non-specific, it belongs to the topic, and the possible types of referential movement are maintenance or shift, as is clearly illustrated in the case of a person description like (3) above. Introduction is not fully excluded, but where it occurs, there are special relative constraints on the rest of the utterance, as was seen above for narratives. If the text question is of the kind "what do you see ...", as, for example in describing a picture, then the description includes both static entities and activities. Then, the quaestio imposes almost no constraint in the particular domain of reference, as becomes clear from examples like "There is a little boy. He is reading. The sun is shining. A cow is standing next to him".

Comparing narrative and descriptive texts with respect to the domain "reference to person and objects" we can make the following distinction: If the quaestio of a text establishes a person or object as belonging to the text-topic, then the two text-types follow basically the same pattern. But if the quaestio of

the description fixes the p-domain as the focus element, then the two text types differ with respect to referential movement and topic/focus distribution.

This is quite different for the last domain of reference considered here: predicates.

The Predicate Domain

Remember that the term "predicate" is used here to denote meaning entities such as states, events, processes and properties represented in a proposition on the basis of linguistic meaning and particular contextual features. The grammatical predicate is used to *express* a "predicate" in this sense, taken together with the context.

In narratives, the predicate has to be a single event. It has to include new information, and it belongs to the focus. There are some problems here as to what may qualify as a single event. It is common to base the distinction between events, processes, and states on inherent temporal properties of the verbs which denote them. This covers the concept of event, as used here, only in part. In narratives, main structure utterances can also use verbs which express processes or even states, so long as it is clear that the whole predicate is temporally bounded; it must *include* a switch to a qualitatively distinct state. This may result from the "inherent" linguistic meaning of the verb, or, as in the case of process verbs or state verbs—it may be explicitly marked by some adverbial.

This is different from descriptive texts, in which the nature of the constraints is governed more closely by the precise nature of the question. In object descriptions, predicates will be unbounded states or properties, such as the visual properties of (3). They are not maintained, and they belong to the focus. In dynamic descriptions, the predicates are typically unbounded processes, which may be generic or habitual; excluded are only single events; the predicate must not be maintained, and it belongs to the focus.

The discussion has shown that there is a number of absolute constraints; there is also a considerable amount of variation, depending on the precise nature of the quaestio. This is in full accordance with the general theory.

3.2 Relative Constraints

Relative constraints are conditional and may impose constraints on several referential domains simultaneously. Consequently, they may be relatively complicated. In what follows, we shall consider one example only, to illustrate the general principle.

Narratives normally require a new predicate, an event, as part of the focus. There are utterances, though, which don't introduce a new predicate but which apparently are direct answers to a question such as "What happened at t_1 ?", and hence should be considered as a part of the main structure. This is possible

only if certain requirements in other domains of reference are met. Consider the following three sentence pairs:

- (14) (a) Bond put ten pounds on noir.
 (b) The lady next to him *did the same*.
- (15) (a) Bond put ten pounds on noir.
 (b) He did the same again.
- (16) (a) Bond put ten pounds on noir.
 (b) He did the same.

In the (b)-utterance, the predicate is maintained, as reflected in the anaphoric expression "to do the same". In the first two cases, this is acceptable—though against the normal constraints—whereas it is not in the third. In the first case, the introduction of a new agent in subject position makes it clear, indeed, that the predicate is a new instance of the same kind of event, and hence that the referent is not really maintained although the expression seems to suggest this. A similar effect could be achieved by introducing a shift in the spatial domain, for example by continuing with "He did the same at another table". In (16)(b), however, there is no indication of any such shift: the event is interpreted as being the same. If this should be avoided in the absence of a change in some other domain, i.e., the "new" event must be explicitly marked as being another instance of the same type of event; and this is the function of "again" in (15)(b).

There are many other relative constraints of this type, most of them quite subtle. But this one may suffice to illustrate the point.

4 Analysis of Selected Phenomena

4.1 Description of Data

In this section we shall have a look at some data to see how the theoretical issues raised above can be pursued empirically. The texts are re-tellings of three different picture stories, without captions, from a comic book (see appendix for two selections). Subjects were three adults and two children (aged 9 and 11), each of them telling the same three stories. They were asked to look at the whole comic strip first, and then to tell the story of the main protagonist. While talking, they were allowed to have the book open. The interlocutor could not see the pictures, however. Thus, the use of deictic pointing was precluded. This technique elicits a very specific type of text, a hybrid of narrative and description. Depending on whether the speaker understands the text question as "what do you see in the pictures?" or as "what happens to person X?" the text will follow a descriptive or a narrative pattern. As the data show, the

speakers all chose the latter as main structure for their stories, but descriptive elements often also came in. They showed up on several different levels:

- Perspective shifts from the protagonist's to the observer's point of view, signalled by expressions such as "man sieht" (one sees), "es scheint" (it seems), introduce a descriptive frame.
- Segmentation into larger episodes corresponds to that imposed by the pictures, or corresponding to the parsing of events in the individual frames of the picture book and the relation between these segments is not established in the temporal domain; the segments are added one to the other like pieces in a puzzle. Typical connectors for this 'additive' relation are "als nächstes" (next), "jetzt" (now) or explicitly "auf dem nächsten Bild" (in the next picture).
- The dominant relation between several entities taken from one picture is established not in the temporal but in the spatial domain.
- The tense form used in all main structure utterances is the present tense, reflecting the fact that the story is being or has been perceived 'second hand' from the pictures, and is not situated in time as a sequence of singular events (cf. the use of present tense in film/story retellings Weinrich 1964, Schiffrin 1986). Past tense forms only come in in side structures in which explanations or evaluations are given.

All of the stories analyzed contain both narrative and descriptive elements, and this is the reason why they are particularly well suited for clarifying the general ideas outlined in the first parts.

We assumed that children follow basic patterns of one particular global structure more rigidly than adults, that they produce fewer side-structure utterances, and that they have difficulties in integrating the different perspectives which are necessary for an understandable account of the picture stories. Therefore, a comparison will help to highlight different degrees of elaborateness with respect to the referential structure of a text. We do not intend to make specific claims about development.

4.2 Data Analysis

For present purposes, we consider two complete stories, one from an adult, AI, one from a child, CI (see appendix). The analysis will focus on a few aspects which have been examined in all texts. Supporting material from the other texts will be provided selectively, as needed.

Text-Structure and Referential Movement

We will first have a look at the global text structures and the distribution of 'main structure' (ms) and 'side structure' (ss) utterances in the two texts.

The stories can be segmented into three parts. The speakers begin with a description of the situation in the first picture. Then a narrative passage follows as the main part of the text. Finally, another description of the changed situation is given.⁸ The referential movement in the different domains should reflect these text structures.

In the following analysis of the referential movement we will focus on the referential domains T (time), L (space) and P (person/object) which allow for the determination of the specific type of cohesion between utterances. The predicate domain, which sometimes contains a number of constituents, will not be analyzed in its different parts, but will be treated as one unit. All propositions are claimed to be valid in the same world. So, M (modality) remains constant throughout the texts (with one exception) and we will not consider it here.

The categories used for the description of the referential movement are the ones introduced in section 2.3, repeated here for convenience:

- 'new' stands for referentially not previously introduced/related/mentioned information ('introduction');
- 'shift' indicates a new reference which presupposes some preceding reference in that domain and cannot be interpreted independently of it. Typically, shifts take place in T and L and are expressed by anaphoric adverbials such as "then" or "next to".
- 'main' (maintenance) stands for the carry over of a specific reference from a preceding utterance. This category encloses total as well as partial maintenance.

The numbers following the categories indicate the utterance to which the given reference is related, if it is not the immediately preceding one. The label (i) is used to mark those cases in which reference is made implicitly.

In addition, the tables 1 and 2 below contain information about the function of an utterance within the overall text structure: 's' stands for side-structure, 'm' for main structure, 'd' indicates those utterances which are part of a descriptive passage.

The flow of information in the three domains R_p , R_t , R_l marks off different segments in the two stories. Both speakers start with a passage in which temporal reference is maintained, corresponding to temporally unbounded states in the predicate domain. New information is introduced in the p-domain, and the spatial domain is used for structuring the selection of successive information. The question eliciting this particular text structure could be paraphrased

Table 1: Referential movement in AI

	R _p	R _t	R ₁	R _a
1	d	new	new	new (state)
2	d	main (i)	main	new (state)
3	d	new	main	new (state)
4	d	main (2)	main	new (state)
5	d	new	main	new (state)
6	s	new	new	new (state)
7	m	new	main	new (event)
8	m	main (4)	shift	new (state)
9	m	main (i)	shift	new (event)
10	m	main (i)	shift	new (event)
11	s	main (6)	main	new (state)
12	m	main (10)	shift (10)	new (event)
13	s	main (3)	main	main (3) (state)
14	s	main (5)	main	main (5) (state)
15	m	main (12)	shift (12)	new (event)
16	m	main (i)	main (i)	main
17	s	main (14)	main	main (14) (state)
18	s	main (13)	main	main (13) (state)
19	m	main (15)	shift (15)	new (event)
20	m	main (i)	shift	new (event)
21	s	—	—	—
22	m	main (i)	shift	new (event)
23	s	new	new	new (event)
24	m	main (22)	shift (22)	new (event)
25	m	new	shift	new (event)
26	s	—	—	—
27	s	new	shift	new (state)
28	m	main (24)	shift (25)	new (state)
29	s	new	new	new (event)
30	m	main (28)	shift (28)	new (event)
31	s	—	—	—
32	m	main	shift	new (event)
33	s	new	main	new (state)
34	m	main (32)	shift (32)	new (event)
35	m	main	shift	new (event)
36	m	main	shift	new (state)
37	s	—	—	—
38	s	main	shift	new (event)
39	s	main	shift	main (state)
40	s	main (i)	main (i)	main (state)
41	s	new	main (38)	new (state)
42	s	main (40)	shift (38)	new (event)

	R _p	R _t	R ₁	R _a
1	d	new	new	new (state)
2	d	new	main	new (state)
3	d	new	main	new (i) (state)
4	m	new	new	new (event)
5	m	main (1)	shift	new (event)
6	m	main	shift	new (event)
7	m	main (i)	shift	new (event)
8	m	main	shift	new (event)
9	m	main	shift	new (event)
10	m	main (i)	shift	new (event)
11	m	main	implicit main	new (event)
12	m	main	of region;	new (event)
13	m	main (i)	implicit shift	new (event)
14	m	main (i)	of reference	new (event)
15	m	main (i)	shift	new (event)
16	m	main	shift	new (event)
17	m	new	shift	new (event)
18	s	main (16)	shift	new (process)
19	m	main (16)	shift (17)	new (event)
20	m	main (i)	shift	new (event)
21	m	main	shift	new (event)
22	m	main	shift	new (event)
23	m	main (i)	shift	new (event)
24	s	new	main	new (state)
25	m	main (23)	shift (23)	new (event)
26	m	main (i)	shift	new (event)
27	m	main	shift	new (event)
28	d	new	main	new (state)
29	d	new	main (i) shift	new (i) (state)
30	d	new	main (i) shift (i)	new (i) (state)
31	d	new	main shift	new (state)
32	d	new	main (i) main	new (i) (state)
33	m	main (27)	shift (27) shift	new (event)

Table 2: Referential movement in CI

as "what do you see in the picture?", excluding temporal reference as a structuring device, suggesting either an object-oriented organization or a structure which follows the spatial organization of the picture.

The characteristics of the first text segment with respect to the referential movement correspond to what has been said for descriptive texts. The referential fillers show properties typical for this text type: the time reference is unspecified, and the predicates are all unbounded states.

At the end of text CI the same picture emerges again. Between 28 and 34 R_t is maintained, R_p introduces new information and the spatial domain is used for interrelating the pieces of information.

Between the two descriptive segments the retellings are organized as narratives. In all texts the descriptive introduction is interrupted by a new time reference to some singular point in time, which is, however, not identifiable for the listener: ("auf einmal", "plötzlich" (suddenly) in CI/4 and AI/6). From this the child moves on, strictly following the referential pattern of a narrative. In the temporal domain 'shift' appears throughout the whole text, meeting the topic condition for this text type. The reference in R_p is maintained except for two cases (17 and 24) with a dummy subject "it" and one object reference in 18. In L the region established in the descriptive passage is maintained until the last utterance, but within that region the protagonist moves around. His activities imply changes of state, but in none of the ms utterances is spatial reference used as a criterion for selecting successive information. The predicate domain always contains new information, and, in contrast to the descriptive passage, the verbs refer to activities or processes.

The child's text exhibits a very homogeneous structure, with all the utterances except two belonging to the main structure and obeying the relevant constraints.⁹ The adult's text differs distinctly from the child's in the narrative passage. It violates the constraints in many instances. Depending on the type of violation, the ss utterances are attached differently to the ms utterances.

Generally, one can distinguish ss which remain inside the given spatio-temporal frame (or situation) from those which fall outside this frame and which then serve to introduce new T/L and perhaps even M parameters. Examples of the first type are utterances 13/14, 17/18, 27, 33, 41 in AI. All of these ss utterances are similar in that they do not shift the temporal reference and that they do not refer to the protagonist in the p domain. They introduce unbounded states, giving information about the circumstances and results of the events.¹⁰

The referential link between ms and ss may be left implicit in that the spatio-temporal reference given in ms has to be maintained for the following ss utterances. This is the case in 13/14 and in 17/18. But the relation may also be established explicitly in one of the referential domains. In 27 and 33 a spatial reference is given in first position which establishes the specific relation between the ms and ss utterances.

Utterances 13/14 (relating back to the utterances 3 and 5 in the descriptive passage) and 17/18 show that ss utterances can establish their own referential net, running parallel to ms events. In 17 the temporal adverb "still" relates back to the mentioning of that particular fact in 14. The second type of ss can be found in utterances which introduce a shift in perspective from the protagonist to the story-teller. In these cases a new spatio-temporal frame is introduced by reference to the 'here and now' of the act of speaking. The deictic anchoring need not be made explicit for all the domains: Explicit reference to one of the domains usually implies shift to the deictic centre for all the other ones. Typical expressions serving this function are "man sieht" (one sees), "man muß sagen" (one must say). But also modal adverb like "offensichtlich" (apparently) in "offensichtlich ist die Kuh aus Gummi", "anscheinend" (it seems) in "anscheinend tut es ihm weh", introduce a shift in perspective of the same kind.¹¹

The discourse function of these ss utterances is to insert a comment on the story or to give an explanation of reported events from an outside point of view. The referential link between those utterances and the ms is typically given by person- or object-references (e.g. AI in 23 and 29).

Let us sum up here. The child(ren) produce(s) a straight-forward piece of text, following the basic principles of referential movement in the two descriptive parts (dominantly spatially related) and the narrative part (temporally related). We do get three instances of ss of the first type (i.e. where the situational frame is maintained), but no case of the second type. This is not specific to one particular child. Other stories exhibit the same pattern. Even in cases where a shift in perspective is essential for an understandable account of the story, the child is unable to provide it. He cannot connect different referential frames as components of one overall text structure.

The adult text, on the other hand, consists of a network of ms and ss utterances. In the narrative passage, 15 of 36 utterances are ss utterances of different types. They all violate the topic condition for the T-domain in that they either maintain the preceding reference point or introduce a new temporal reference which is deictically anchored. Moreover, they contravene the condition for the referential filler of the predicate domain. All the ss utterances introduce states. But in all cases, the ss utterances do not affect the temporal thread connecting the ms utterances across one or more insertions by a shift-in-time relation.

Referential Movement and Devices of Expression

Among the linguistic devices which a particular language offers, some are variable according to contextual (and maybe other) factors, others are not. In English, for example, the article precedes the noun within an NP, and the object follows the verb, independently of any contextual factor. The first constraint is also mandatory in French or German, but the second is not: the object may also

precede its governing verb; in French, this possibility is confined to anaphoric objects whereas there is no such restriction in German.

In this section we shall deal with the question of how referential movement influences the choice of a particular linguistic form, if the syntactic corset leaves an option. Two such cases will be considered—first some examples of ellipsis, and second optional word order of temporal and spatial adverbials.¹²

Ellipsis Phenomena

In referring to a previously introduced person (or object) the speaker has basically three options: a full (lexically specified) NP, a personal pronoun, or zero anaphora. The last option is typically most constrained, in order to ensure identification of the referent. Thus, English tends to confine it to the subject; a maintained referent in object position can't be left implicit, although it may be perfectly clear to whom or to what is referred. It is hard to say whether this asymmetry is due to the difference in syntactic properties ("subject-object") or to word order (before V-after V), since function and word order normally correlate. In German, where word order is much freer, zero anaphora is possible for object and subject, but for both of them only if they appear in initial position. Hence, the factors which determine word order (subject first vs. object first)¹³ and those which determine the choice of NP (full NP vs. pronoun vs. zero anaphora) interact in a subtle way. This means that in German the choice of zero anaphora is not independent from the organization of the information in the other referential domains. If the speaker chooses the temporal (e.g. "then") or spatial (e.g. "next to") domain for establishing the dominant referential link then zero anaphora is impossible.

It turns out that object ellipsis, though structurally possible, is absent in our texts, whereas omission of maintained referent in subject position (henceforth ØS) is quite frequent. Its appearance—in contrast to full NP or pronoun—depends on two factors:

1. The absence of explicit mention in an uninterrupted sequence of ms utterances serves a *chunking* function.¹⁴ Typically, the activities around one entity are more closely tied together in this way (cf. for example CI 6-7, 9-10, 12-15). The beginning of a new episode—or of a comparable segment in descriptive texts—is marked by an explicit reference to the (maintained!) protagonist, and in most cases also by a new mention of the dominant referential link in the text—the temporal shifter "(und) dann".

This factor is observed both in the adults' and in the children's text.

2. The second factor, in contrast, is only operative in the adult data. The possibility of ØS depends on whether the utterance in question belongs to ms or to ss: It can't go from ms to ss, or from ss to ms¹⁵ it is possible from

ms to ms (typically two subsequent events involving the same protagonist) or from ss to ss (typically two states involving the same protagonist).

Whenever there is a shift from ms to ss, or vice versa, and the referent is maintained, adult speakers use a pronominal form. Here are some examples:¹⁶

- (a) ms: er geht in ein Schaufenster
 he goes up to a store window
 ms: und ØS schaut sich an
 and looks at himself (his reflection)
 ms: ØS nimmt seinen Hut hinunter
 takes off his hat
 ss: er hat eine Glatze ...
 he is bald

Having ØS instead of "er" in the last utterance would sound very odd here: it would suggest the idea that having a bald head is part of the ms, the event following the taking off of the hat; having "er" indicates a shift from ms to ss.

- (b) ms: es blubbert
 it bubbles
 ss: er sieht sich schon mit dem herrlichsten Schmuck im Geiste
 he sees himself already with beautiful headdress in his mind
 ms: und er nimmt von dieser Flüssigkeit
 and he takes some of this liquid
 ms: und ØS streicht sie über seinen kahlen Kopf
 and smooths it over his bald head

Here, the "er" in the prefinal utterance could be replaced by ØS; but this would change the status of the preceding utterance: it would make it the first of a series of events, hence as a part of ms. In the following example, maintenance expressed by ØS "jumps" over an inserted ss with "er":

- (c) ms: er kriegt dabei ein paar Tropfen in die Augen
 he gets, because of that a couple of drops in his eyes
 ms: und ØS hat plötzlich ziemlich üble Augenschmerzen
 and has suddenly rather bad eye-pains
 ss: er kann gar nichts sehen
 he can see nothing
 ms: Ø will schnell nach einem Handtuch greifen
 wants to quickly reach for a towel

In this case, it can't be excluded that the third utterance expresses an event following the incident described before. But it seems much more plausible to assume that it describes a state concurrent with the eye pains, and that the narrative sequence goes on with the last utterance. Consider a last example which illustrates such a jump:

- (d) ms: und ØS mischt dann irgendein Mittel
 and mixes then some kind of concoction
 ss: das brodelt und bruzelt ganz schön
 that bubbles and simmers quite nicely
 ms: und ØS denkt sich jetzt also
 and thinks to himself then
 ss: jetzt wird er ein ganz Schöner
 now he becomes a very good-looking man
 ms: und ØS schmiert es sich auf den Kopf ...
 and smears it on his head

Clearly, observations of this kind can't be accounted for by purely syntactic constraints on the use of ØS. They are related to the overall referential organization of the text. ØS is excluded when the *conditions for* the referential movement in the different domains are changed by an utterance (ss→ms; ms→ss). If an utterance contains ØS then no other element can be put in first position, indicating a new type of referential link. The ØS implies maintenance of the conditions for referential movement as established in the preceding utterance(s). Therefore, in a narrative structure ØS in ms implies a shift-relation in the temporal domain, in a descriptive text ØS goes along with maintenance in this domain.

We observe similar phenomena for other types of ellipsis. Omission of the finite verb, for example, is possible roughly when it agrees in form (person, number, tense) with a preceding verb which introduces the same information. But even when these conditions are satisfied, it is impossible to have ØV when the maintenance goes from ms to ss or vice versa. Like ØS, ØV implies maintenance of the conditions for referential movement as established by the preceding utterance. Since there are only a few examples of verb ellipsis in the text, we will not pursue it here.

Word Order

The order of major constituents is comparatively free in German. In declarative main clauses, the finite verb invariably takes the second position, but the one constituent before it can be the subject, any object, an adverbial or even the (infinite) verb. There is also a relative freedom in the arrangement of the remaining constituents after the finite verb. Which principles, other than purely syntactic constraints, determine word order in German? There are a large number of studies which explain word order regularities in German by means of

the pragmatic categories "given-new", "theme-rheme" or "topic-comment" (cf. Behaghel 1923/1932, Drach 1939, Grundzüge 1981, Danes 1974, Lenz 1977). According to these theories, the central assumption of these accounts is the notion that German follows the (wide-spread) principle "given precedes new" (or topic precedes comment, in one definition of these terms (cf. Lutz 1981)). But there remain a number of problems which cannot be solved by the given-new dichotomy:

- What are the factors determining word order if several constituents have the same status within the informational flow of a text (e.g. temporal and spatial anaphoric adverbs, and anaphoric pronouns)?
- What are the special functions of elements in first or last position in a sentence? These two positions, which can be regarded as marked in a psycholinguistic sense (cf. Slobin 1973), cannot be identified with topic or focus, respectively, for two reasons: 1. There can be more than one topic or focus element. 2. Topic and focus elements are not always put in the typical sentence-initial or sentence-final position.

The question "What do you see?", for instance, determines the specification of "what" as the focus component of the answer: "Ich sehe *eine Kuh*." (I see a cow). But one could also answer: "*Eine Kuh* sehe ich", which does not change the topic-focus structure of the answer sentence. In some studies the idea of a "hierarchy of topicality" was developed (e.g. Givón 1985, Li and Thompson 1976). But this does not lead any further in answering the question. The problem remains: How does the speaker decide which element to mark as "more topical" than the other(s).

Or to state the question from the point of view of comprehension:

- What is the difference in the interpretation of two sentences such as

"suddenly the alarm rings" and

"the alarm rings suddenly"?

Obviously we cannot even try to solve those problems here. Our aim is much less ambitious. By discussing two observations made in the data we will show that relating word-order phenomena to the referential structure of a text will add a dimension that might finally help to answer some of the open questions.

1. Within one type of referential structure the same element (e.g. spatial or temporal adverbial) is put in first position in one utterance and in last position in a following utterance.¹⁷ Examples of this kind are the utterances 2 and 3 in CI where the anaphoric adverb "daneben" differs

in position; in the same text, utterances 29, 32 and 33 where the spatial reference "in der Nähe" takes last position in the first utterance, the two following spatial references "überall" are put in utterance initial position. A third example of this kind is found in AI, utterance 13 and 17. Here, the temporal adverbial "immer noch" changes position. These variations in word order cannot be accounted for by syntactic factors nor by topic-focus regularities. In all examples, the elements in question are part of the topic component within the utterance. The word order pattern which has to be explained could be schematized as follows:

Focus — **Topic** (spatial or temporal
Topic — **Focus** reference)

The difference could have semantic reasons, in the first example, the subjects "cow" and "camper", respectively. The cow might be seen as a second actor in the story, whereas the camper is perceived as an attribute of the protagonist. This difference in animacy could be the reason for the difference in position (cf. Comrie 1981, Foley and van Valin 1985).

This 'semantic explanation' does not hold for the other examples, however. There is no parallel between the subjects of the three sentence pairs. Another reason for reversing the order in the second utterance could be the position of the utterance with respect to the unfolding of a particular global structure. The examples in CI occur at the beginning of descriptive passages, in which the spatial domain is used for structuring the information. In AI-13 a shift from ms to ss takes place in which a stative temporal reference is established for more than one utterance. If a topic element, e.g. anaphoric spatial or temporal reference, is put in last position, which is typically the focus position, then this has the function of establishing a particular referential domain as cohesion creating for a following piece of discourse. The discourse function of the focus-topic order can be described as *cataphoric*, announcing a particular type of referential structure. This goes along with the intuition that changes in the unmarked topic-focus order often create a certain expectation on the listener's side about how the discourse will be continued.

The topic-focus order in the second and third utterances in CI follows the normal pattern by putting the element expressing the dominant referential link (topic condition) in first position. We do not claim the textual factors are solely responsible for explaining word-order variation. Word order, where it is variable, is the result of a *competition* between different, partly incompatible factors, the overall referential structure of a text being one of them.

2. The second phenomenon to be discussed is illustrated by the following examples:

- (a) ms: "...und geht *dann* weiter
and goes then farther
ss: er ist also nicht reingefallen
he has, after all, not fallen in
ms: und *dann* geht er weiter ..."
and goes then farther
- (b) ss: "es bleibt *schließlich* nur ein Hautfetzen übrig"
there remain finally only pieces of skin
ms: "*schließlich* besteigt er seinen Wohnwagen"
finally he climbs into his caravan
- (c) ms: "er geht nun auf die Straße
he goes then on the street
ms: und stellt *auf einmal* fest ..."
and realizes suddenly
- ms: "der Friseur setzt ihn wieder auf einen Stuhl
the barber sets him back on his chair
ms: und hat *auf einmal* eine Maschine in der Hand ..."
and has suddenly a machine in his hand
- ms: "*auf einmal* klingelt der Wecker"
suddenly the alarm rings
- (d) ss: "ein paar lange komische Kruselhaare hat er *nun* auf dem Kopf
a few long strange curly hairs has he now on his head
ms: "nun will er sich ein Handtuch holen ..."
now he wants to get a towel

The examples are all taken from narrative passages, a comparison with descriptive texts is (due to the limited amount of data) not possible. In narrative texts the temporal domain constitutes the dominant referential link, as determined by the topic constraint. In the examples, the temporal adverbial are in different positions. This leads us to the following two questions:

- How does shift from ss→ms and vice versa influence the position of the temporal adverbial?
- Does the change of position of the adverbial have different effects in ms and ss utterances?

The answer to the first question is relatively clear and can be substantiated by a number of examples in the texts. In (a) we have a case where the same ms-utterance is repeated in different order after a ss has been inserted. A change of order in the third utterance "und geht dann weiter" would lead to a strange sounding piece of text.

The general regularity behind this order phenomenon is this: If the ms is interrupted by a ss and if the following ms utterance contains an anaphoric temporal adverbial, then the shift back to the ms will be indicated by putting the temporal adverbial in first position—"reestablishing" the dominant referential link (cf. the "jetzt" in almost all ss→ms in AI) which was suspended for the ss-utterance(s). The answer to the second question is much more problematic. Intuitively, there is a clear difference between

- e.g. I ... und *schließlich* stellt er fest ...
 and finally he realizes
 II ... und (er) stellt *schließlich* fest ...

in the context of narration.

The first sentence indicates a cut in the flow of information, a new step or episode. By the second sentence a much tighter link is established with respect to the preceding utterance. The reason for the difference could be that in the second sentence the topic condition for the temporal domain is implicitly maintained whereas in I it is explicitly replaced by reference to a shifted point in time.

For ss-utterances the effect of word order change is less obvious (cf. example (b) and (d)). In both cases the adverbs "schließlich" (finally) and "nun" (then) do not indicate a temporal shift. Due to the stative character of the predicate, they maintain the preceding reference point, and the adverbs get a kind of modal meaning. Changes in the position of the adverbs in ss therefore do not lead to the same consequence for the temporal structure as has been claimed for ms-utterances. It should be clear that these considerations are not an answer to the question as posed above; but they point in the direction in which we think the answer can be found.

5 Final Remark

The general framework outlined in the first part of this chapter integrates a wide range of issues which, though traditionally not studied together, closely interact in the constitution of a text. It provides us with a new perspective on a number of linguistic phenomena which we think have as yet not been sufficiently understood, for example ellipsis and regularities of word order.

A Appendix

A.1 Text CI—Nine Year Old Child

1. Albert liegt auf einem Liegestuhl in der Sonne
Albert is lying in a deck-chair in the sun
2. eine Kuh mit einer Glocke um den Hals sitzt daneben
a cow with a bell around the neck is sitting next-to-it
3. dann daneben ein Campingwagen mit einem schönen Sonnendach
then next-to-it a camper with a a nice sun blind
4. plötzlich ringelt der Wecker
suddenly the alarm clock rings
5. Albert steht auf, gähnt, streckt sich und reckt sich
Albert gets up, yawns, stretches
6. dann nimmt er das Sonnendach
then he takes the sun blind
7. rollt es auf schlipp-schlapp
rolls it up schlipp-schlapp
8. dann macht er mit seinem Liegestuhl krick-krack
then makes noises with his deck-chair
9. und er nimmt direkt die Treppe
and he takes immediately the stairs
10. und klappt sie wieder ein in den Campingwagen
and folds it again up into the camper
11. dann tut er seine Becher und sein Trinken einpacken
then does he pack his cups and his drinks
12. und dann geht er auf die Kuh zu
and then he approaches the cow
13. packt sie am Kopf
takes it by the head
14. macht mpf
goes mpf
15. und nimmt den Kopf ab
and takes off the head

16. dann zieht er sie an der Glocke
then he pulls it by the bell
17. und dann macht es pfff
and then it goes pfff
18. die Kuh wird immer kleiner
the cow becomes smaller and smaller
19. dann nimmt er ein Stück vom Rasen
then he takes a piece of the lawn
20. und rollt den Rasen ein
and rolls up the lawn
21. dann geht er einfach auf die Berge zu
then he approaches simply the mountains
22. dann plötzlich nimmt er ein Stück von der Sonne
then suddenly he takes a piece of the sun
23. und zieht daran
and pulls on it
24. sieh da—es ist ein Reißverschluß
look—it is a zipper
25. dann nimmt er das ganze Ding auseinander, nämlich eine Leinwand
then he takes the whole thing apart, namely a projection screen
26. und bindet sie zusammen
and ties it up together
27. dann zieht er sich normal an
then he dresses normally
28. und wo er war ist eine eingeödete Landschaft
and where he was is a waste-land
29. Fabriken in der Nähe
factories nearby
30. ein scheußlicher Sonnenuntergang
a horrible sunset
31. überall liegen die Fetzen, Fischgräten
everywhere rags are lying around, fishbones

32. und überall Dosen
and everywhere tin cans
33. und er fährt mit dem Campingwagen in die nächste Stadt
and with the camper he drives into the next city

A.2 Text AI—Adult

1. A. E. liegt gemütlich auf einer Liege vor seinem Wohnwagen
A. E. is lying comfortably in a deck-chair in front of his camper
2. und liest ein Buch
and is reading a book
3. die Sonne scheint am blauen Himmel
the sun is shining in the blue sky
4. vor seinem Wohnwagen hat er's wunderschön mit Blumen
in front of his camper he is having it very nice with flowers
5. und auf dem grünen Rasen ist eine Kuh zu sehen die ihm zuschaut
and on the green lawn a cow is to be seen which is watching him
6. mit einmal scheint seine Ruhezeit zu Ende zu sein
all of a sudden his rest seems to have come to an end
7. der Wecker läutet
the alarm clock rings
8. und er steht auf
and he gets up
9. und reckt sich
and Stretches
10. und denkt
and thinks
11. "jetzt is meine Zeit hier leider bald um"
now my time here is unfortunately soon over
12. und er legt den Liegestuhl zusammen
and he packs up the deck-chair
13. die Sonne scheint immer noch
the sun is still shining

14. und die Kuh schaut ihm zu
and the cow is watching him
15. er verpackt sein Zeugs im Wohnwagen
he packs his stuff into the camper
16. eins nach dem andern
one after the other
17. und immer noch schaut ihm die Kuh zu
and the cow continues to watch him
18. und die Sonne scheint
and the sun is shining
19. jetzt geht er auf die Kuh los
now he approaches the cow
20. packt sie am Kopf
grabs her by the head
21. ach du liebe Zeit
oh my goodness
22. und reißt ihr den Kopf ab
and tears off her head
23. da sieht man, daß das gar keine echte Kuh ist, sondern daß das nur eine
Attrappe ist, eine gemachte Sache
from this one sees that it isn't a real cow, but that it is only a dummy, a
faked thing
24. jetzt nimmt er die Glocke ab pffftt
now he takes of the bell pffftt
25. und die Luft geht aus diesem aufgeblasenen/anscheinend dieser aufgeblase-
nen Gummikuh
and the air comes out of this inflated, this apparently inflated rubber cow
26. ach du liebe Zeit
oh, my goodness
27. da liegt der Rest der Kuh auf einem Haufen
there lies the rest of the cow in a heap
28. und jetzt rollt er sogar den Rasen mit allen Blümchen zusammen
and now he even rolls up the lawn with all the little flowers

29. und man sieht, daß das auch nur eine künstliche von ihm ausgerollte Matte war
and one sees that this was also only an artificial mat rolled out by him
30. jetzt hat er das schon alles zusammengerollt
now he has finished rolling all of this up
31. ja und was is denn das hier
yes and what is this then
32. jetzt nimmt er von einer Leine eine große/ einen großen Vorhang, einen großen blauen Vorhang ab
now he takes from a line a big curtain, a big blue curtain
33. und darauf ist die Sonne und die Wolken und der blaue Himmel gemalt
and on this is painted the sun and the clouds and the blue sky
34. das legt er jetzt auch zusammen
this he now also folds up
35. jetzt zieht er sich sein Jacke an
now he puts on his jacket
36. und is ganz fertig
and is totally ready
37. und was sieht man
and what does one see
38. er fährt mit seinem Wohnwagen davon
he drives off with his camper
39. hat alles aufgeschnallt
has everything buckled up
40. den Rasen und den Himmel
the lawn and the sky
41. und zurück bleibt eine scheußliche fast Müllhalden-Landschaft
and left behind is a horrible almost ash pit-land
42. und er fährt zurück in seine große Stadt mit hohen Häusern
and he drives back to his big city with tall buildings

References

- [1] Behaghel, O. (1923, 1932): *Deutsche Syntax*. Heidelberg.
- [2] Comrie, B. (1981): *Language Universals and Linguistic Typology*. Chicago.
- [3] Danes, F. (1974): Functional Sentence Perspective and the Organization of Texts. In: Danes, F. (ed.): *Papers on Functional Sentence Perspective*, 106-128.
- [4] van Dijk, T. (1977): *Text and Context*. London-New York.
- [5] Drach, E. (1939): *Grundgedanken der deutschen Satzlehre*. Frankfurt a.M.
- [6] Ehrich, V. and Koster, C. (1983): Discourse Organization and Sentence Form. The Structure of Room Descriptions in Dutch. *Discourse Processes* 6, 169-195.
- [7] Foley, W. A. and van Valin, R. (1985): Information Packaging in the Clause. In: Shopen, T. (ed.): *Language Typology and Syntactic Description I*, 282-364. Cambridge.
- [8] Givón, T. (1983): *Topic Continuity in Discourse*. Amsterdam.
- [9] Givón, T. (1985): *Syntax: A Functional-Typological Approach*. Amsterdam.
- [10] *Grundzüge einer deutschen Grammatik*. (1981). Autorenkollektiv. Berlin.
- [11] Herring, S. (1985): *Marking and Unmarking via the Present Tense in Narration. The Historical Present Redefined*. Ms. University of California. Berkeley.
- [12] Hickmann, M. (1980): Creating Reference in Discourse. A Developmental Analysis of Linguistic Cohesion. In: Kreiman, J. and Ojeda, A. E. (eds.): *Papers from the Parasession on Pronouns and Anaphora*. Chicago Linguistic Society. Chicago.
- [13] Hopper, P. (1979): Aspect and Foregrounding in Discourse. *Syntax and Semantics* 12, 213-241.
- [14] Klein, W. and Perdue, C. (1988): The Learner's Problem of Arranging Words. In: Bates, E. and MacWhinney, B. (eds.): *Crosslinguistic Studies in Language Acquisition*. Cambridge.
- [15] Klein, W. and v. Stutterheim, C. (1987): Quaestio und referentielle Bewegung in Erzählungen. *Linguistische Berichte* 109, 163-183.

- [16] Labov, W. and Waletzky, J. (1967): Narrative Analysis. In: Helm, J. P. (ed.): *Essays on the Verbal and Visual Arts*, 12-44. Seattle.
- [17] Labov, W. and Linde, C. (1975): Spatial Networks as a Site for the Study of Language and Thought. *Language* 51, 924-939.
- [18] Lang, E. (1977): Semantik der koordinativen Verknüpfung. *Studia Grammatica* XIV. Berlin.
- [19] Lenerz, J. (1977): *Zur Abfolge nominaler Satzglieder im Deutschen*. Tübingen.
- [20] Li, C. N. (1976): *Subject and Topic*. Academic Press. New York.
- [21] Li, C. N. and Thompson, S. (1976): Subject and Topic: a New Typology of Language. In: Li, C. N. (ed): *Subject and Topic*. New York.
- [22] Lutz, L. (1981): *Zum Thema "Thema-Rhema"*. Hamburg.
- [23] Lyons, J. (1976): *Semantics* 1/2. Cambridge.
- [24] Marslen-Wilson, W., Levy, E. and Tylor, L. (1982): Producing Interpretable Discourse. The Establishment and Maintenance of Reference. In: Jarvella, R. J. and Klein, W. (eds.): *Speech, Place, and Action*. Chichester.
- [25] Palmer, F. R. (1986): *Mood and Modality*. Cambridge.
- [26] Quasthoff, U. (1980): *Erzählen in Gesprächen*. Tübingen.
- [27] Rehbein, J. (1984): Beschreiben, Berichten und Erzählen. In: Ehlich, K. (ed.): *Erzählen in der Schule*, 67-124. Tübingen.
- [28] Reinhart, T. (1984): Principles of Gestalt Perception in the Temporal Organization of Narrative Texts. *Linguistic* 22, 779-809.
- [29] Slobin, D. (1973): Cognitive Prerequisites for the Development of Grammar. In: Ferguson, C. A. and Slobin, D. I. (eds.): *Studies of Child Language*. New York.
- [30] Weinrich, H. (1964): *Tempus. Besprochene und erzählte Welt*. Stuttgart.
- [31] Weissenborn, J. (1984): La Genèse de la Reference Spatiale en Langue Maternelle et en Langue Seconde: Similiarités et Differences. In: Extra, G. and Mittner, M. (eds.): *Studies in Second Language Acquisition by Adult Immigrants*, 262-272. Tilburg.
- [32] Werlich, E. (1979): *Typologie der Texte*. Heidelberg.

Footnotes

- ¹ This article is based on two papers given at the Symposium "Language processing in social context", held at the University of Heidelberg in 1986. They have been considerably revised. We wish to thank M. Carroll, U. Kohlmann, U. Scharnhorst, J. Edwards, and the participants of the Symposium for helpful comments on earlier versions.
- ² Obviously, a text may contain utterances which cannot be readily classified under the one or the other type of structure. For example, the speaker could have said: "Then, I saw that he had a thin, black moustache." Superficially, this utterance reports an event; but the point of the utterance is clearly to provide a descriptive feature of the type asked for. Thus, on a deeper level, it contributes to the main structure. It is a description in disguise. For present purposes, we will ignore this and similar complications.
- ³ For a more extensive discussion of "background—foreground" on the one hand, "main structure—side structures" on the other, see Klein and von Stutterheim (1987), section 2.
- ⁴ It should be evident that there is no clear-cut boundary between descriptive texts, narrative texts, reports or other text types. On the other hand, there is no doubt that there are constraints on text structure above and beyond those imposed by the *questio*. There may be particular culturally conditioned patterns of telling a story, there may also be special cognitive "schemas", like the ones described in "story grammars".
- ⁵ There are languages and cultures which make a systematic difference between these two types of narratives, for example by a special morpheme (like the Turkish "evidential"). In this case, the statement above has to be qualified accordingly.
- ⁶ There are exceptions to this constraint. Under specific circumstances, simultaneity of two events, hence maintenance of the temporal referent, does not lead to a side structure; but then a number of relative constraints come into play (cf. Klein and v. Stutterheim 1987).
- ⁷ Note that in descriptive texts, connectors such as "and then" or "next" serve a different function compared to that served in narrative texts. They do not shift the temporal reference, but often the "order of describing", such as in "And then, there is a book on the table" or "Next, there are three sideboards at the opposite wall".
- ⁸ AI is the only one who refers to the situation by means of only a single sentence.
- ⁹ The internal segmentation of the narrative will be discussed below in the context of questions related to the linguistic representation, because it is not determined by specific features of the referential structure.
- ¹⁰ These are typical instances of what has been analyzed as "background information" in Labov and Waletzky 1967, Hopper 1979, Reinhart 1985.
- ¹¹ It is interesting that whenever a *ss* of this type is inserted, the modal adverb is always put in first position. We will come back to the role of word order below.
- ¹² Ehrich and Koster studied word order variation in room descriptions. They related different word order patterns to different discourse strategies of the speakers and distinguished two strategies: spatially oriented discourse organization and object oriented organization.
- ¹³ If some other constituent, like an adverbial, comes first, in this case, both object and subject follow the verb, and zero anaphora is (normally) excluded. There are some arguable cases, though; for a full discussion, see Grundzüge 1981.
- ¹⁴ This textual function of zero anaphora has been described for children and adults in Quasthoff 1980, Marslen-Wilson, Levy and Tylor 1982.

Cf. the study on ØS in L2-discourse by Klein and Perdue 1986, who observed similar regularities.

We illustrate the rules taking as an example a narrative text type. The same rules are valid for descriptions, but there the characteristics of ms and ss have to be reversed.

See the discussion in Ehrich and Koster about different word order types. They state the following "formulation maxim":

FI "Encode information being related to the overall organization principle of the discourse in first position." (p. 189)

Different organization principles can be found in object-oriented versus path-oriented descriptions. Although the maxim FI explains a number of word order phenomena, it cannot be the whole story. As our examples show, texts are not always as consistent in terms of word order as FI would require.