

Wolfgang Klein, Nijmegen/Christiane von Stutterheim,
Heidelberg

Text structure and referential movement¹

Contents

1. Global and local constraints in text structure
2. Main structure and side structures of a text
3. Referential movement
4. Conclusions

1. Global and local constraints in text structure

A text², such as a route description, a personal narrative, a judge's opinion during a verdict, differs in two respects from an arbitrary collection of utterances:

A. It obeys certain global constraints which primarily result from the fact that the utterances in their entirety serve to express, for a given audience and to a given end, a complex set of information, a *Gesamtvorstellung*, as we shall say³. The components which belong to this *Gesamtvorstellung*, as

¹We want to thank Clifford Hill and Clive Perdue for helpful comments on an earlier draft of this paper.

²In what follows, we shall only consider monological texts. As we shall see, the approach taken here treats all texts as answers to (explicit or implicit) questions, and to that extent it is dialogue-oriented. But the speaker change which characterizes dialogue leads to some additional constraints in text structure which are not relevant to our present concerns.

³There is no received term for the complex cognitive structure which underlies a text. Indeed, there is not one such structure but various levels of representation, and text production is but the last step in a series of transductions from one level to the next. Take, for example, a narrative of an event, in which the speaker was involved. There is first the "real" event (level 0), which is experienced and perceived, hence transformed into some "percept" (level 1) and then somehow stored in long-term memory (level 2), where it quietly warps. It is then recalled on a given occasion (level 3), components of it are selected, linearized and possibly enriched by fictitious additions (level 4), and eventually, the resulting "discourse representation" is put into words (level 5, text). On the comprehension side, the listener extracts the meaning of these words (level 6), combines it with contextual information (level 7), and enriches the resulting interpretation by all sorts of inferences, based on his world knowledge (level 8). The number and nature of these levels as well as the ways in which they are related to each other are a matter of much dispute. We will not go into this controversial issue here but simply speak, with deliberate vagueness, of *Gesamtvorstellung*, which is meant to include all levels of representation on

well as the relations which obtain between them may be of a different nature. Thus, the judge's opinion is based on a *Gesamtvorstellung* whose individual components - the specific facts of the case, generic statements of the law, previous decisions, moral evaluations, etc. - are essentially held together by (hopefully) logical relations. The *Gesamtvorstellung* which underlies a narrative primarily consists of singular events whose main relationship is temporal, in contrast to, for example, a room description, where the components are some physical entities which are spatially connected, and so on. The nature of the *Gesamtvorstellung*, on the one hand, and the specific purpose the speaker has in expressing it, on the other, impose specific constraints on the overall organization of the text.

B. The way in which the text proceeds from one utterance to the next obeys local constraints, depending on which information is introduced, maintained or elaborated on. This *referential movement* from utterance to utterance becomes apparent in the choice of specific linguistic means, such as the use of definite vs. indefinite noun phrases, anaphoric elements, word order, intonation or lexical items like "too".

Each utterance selects a segment from the *Gesamtvorstellung* and puts it into words. The way in which this is done depends not only on what has to be expressed but also on what can be taken over from the preceding utterance(s) and what must be freshly introduced. The most obvious and best studied kind of referential movement concerns the introduction and maintenance of participants, for example the characters of a narrative (see, among many others, Marslen-Wilson et al. (1982), Givón (1983)); but clearly, it applies to other possible domains of reference, such as time, space and others.

Both global and local constraints⁴ can be violated. Thus we can easily imagine a text with the overall structure of a narrative, where adjacent utterances do not fit together, and we can imagine a text where each utterance is appropriately hooked up to the adjacent one but where there is no higher organization. This suggests that both aspects of text structure may be studied independently, and this indeed has been done in most, though not all of the literature. Global constraints, for example, have been

the speaker's side. The term *Gesamtvorstellung* (entire representation) is borrowed from Wundt (1912), although Wundt uses it in a narrower sense.

⁴We do not claim, incidentally, that there might be "intermediate" constraints, especially in longer texts. Our point here is merely that all well-formed texts obey at least these two. The terms "global" and "local", also used by other authors (for example van Dijk (1980)) are not meant in any technical sense here; they are used as loose labels for phenomena to be analysed below.

stated in terms of scripts, frames, "story grammars", macro structures, etc. (for a recent survey, see Hoppe-Graff (1984)). A great deal of what is named here local constraints has been investigated under labels such as "coherence", "cohesion", "thematic progression", and many others (see, for example, the still impressive work by Halliday/Hasan (1976)). What has been much less studied is how in a well formed text, both types of constraints in a text interact or, more specifically, how the local constraints follow on from the global ones. This is the issue of the present paper whose basic idea may be roughly stated as follows: Global constraints, which result from the *Gesamtvorstellung* and the "text question" (in a sense to be explained), can be stated as restrictions on possible referential movement and, as a consequence, of the appropriate language-specific means to express this referential movement.

2. Main structure and side structures of a text

Narratives of personal events belong to best-studied text types⁵, and so we will take them as a starting point for our considerations. They give an account of some event that happened somewhere sometime to the speaker (or to a third person). The *Gesamtvorstellung* (henceforth GV) consists of a set of sub-events which are above all temporally but also spatially or casually related to each other and which in their entirety constitute a singular event constellation, that can be situated in time and space. The single utterances which constitute the narrative, answer in their totality a - real or fictitious - question, the *quaestio*⁶ of the text, roughly: "What happened (to you) at this time at this place?" In answering this question, the speaker has a certain freedom in selecting the sub-events he wants to report, and in the way in which he arranges them. But in any event, the text must render a certain *event structure*, i.e. a set of sub-events and the temporal relationship between them. This is the central characteristic of a

⁵See, for example, Labov/Waletzky (1967), Labov (1972), Quasthoff (1980), Ehlich (1980), to mention but a few.

⁶The idea of characterizing the overall function of a text by such a *quaestio* is clearly not new. At least for argumentative texts, it is found in ancient rhetoric. But it also applies to many other texts, e.g. route directions ("How do I get from X to Y?"), to apartment descriptions ("What does your apartment look like?"), to recipes ("How to make an onion tart?"), etc. Three points should be noted however. First, the *quaestio* of a text need not be identical with the real question which may have elicited the text in the given case, although it may be, of course: but very often, there is no explicit question at all. Second, it may be more appropriate for some texts to characterize them by a pair or even a triple of *quaestiones*, rather than by one. And third, some texts of a more loose nature (small talk, for example) may not be characterizable by a *quaestio*, at all. This corresponds to the fact that they have no, or only a very weak, global structure: their organization is merely local.

narrative. The speaker may choose to add supplementary material. Thus, a narrative most often contains information about time and place of the entire constellation (“orientation”) as well as comments, explanations, evaluations, etc. These general characteristics of any narrative may be reflected in global constraints on its structure, for example:

A1: Whatever the selection of sub-events may be, they must be presented in the order in which they happen (“principle of chronological order”).

A2: At the beginning, the event constellation must be situated in space and time.

A3: Usually, evaluations etc. must be inserted immediately before or after the sub-event to which they belong.

Constraints of this sort have been stated by various authors in various forms⁷. They impose on the text a fixed overall scaffold - *a main structure* (A1) - which may be completed at designated points by various *side structures* (A2, A3).

There have been several attempts to capture the observations which give rise to this distinction, the best-known being the contrast between utterances which provide “foreground information” and those providing “background information” in a narrative. In the section, we will discuss these concepts.

2.1. Foreground and background

The foreground-background distinction comes from gestalt theories of perception, and although its application to texts is not new (see, for example Weinrich (1964), Fleischmann (1973)), it is only recently that it became more popular. Hopper (1979, p. 213) states:

“It is evidently a universal of narrative discourse that in any extended text an overt distinction is made between the language of the actual story line and the language of supportive material which does not itself narrate the main events. I refer to the former - the parts of the narrative which relate events belonging to the skeletal

⁷The principle of chronological order, for example, was explicitly stated in ancient rhetoric, albeit negatively: the “*hysteron proteron*”, i.e. the prior reporting of the later event, is considered to be a violation of regular text structure, which is only allowed for specific rhetorical effects.

structure of the discourse - as FOREGROUND and the latter as BACKGROUND”.

Similarly, Reinhart (1984, p. 781/2) writes:

“ ... narrative texts are organized obligatorily along a temporal axis, which is the sequence of narrative clauses (or units). The temporal axis is called by Labov the ‘narrative’ skeleton of the text. The non-narrative ‘flesh’ which is organized ‘around’ this skeleton provides the necessary details for reconstructing the represented world and for determining the meaning and the purpose of the text ...”

We will call what Labov had defined as the “narrative skeleton” the *foreground* material of the text, and the non-narrative material (under this definition) - *the background*.

Foreground and background, as defined here, relate to the overall organization of a narrative - to its “skeleton” and its “flesh”. But they also bear on the way in which the information is displayed in the individual utterance. For example, foreground clauses are regularly marked by perfective aspect, background clauses by imperfective aspect, where there is such a distinction in the language; subordinate clauses regularly contribute to the background; some languages use different word order for foreground and background clauses; others may indicate the difference by specific particles, etc. (For a survey and discussion, see Hopper (1979); Reinhart (1984); Thompson (1984))⁸.

The distinction between utterances that push forward the action and those that don't allow to interrelate global and local constraints within a narrative text. The overall structure - the narrative skeleton - is given by those utterances which, roughly speaking,

- refer to a singular event that normally can't be maintained from a previous utterance, and
- move ahead the time compared to the time of the previous utterance (if there was a previous utterance).

⁸There is a certain danger, not always escaped in the literature on foreground and background, to *define* these concepts in terms of the linguistic means which characterize the corresponding utterances (or clauses). Then, obviously, the distinction is circular, if it is understood as a hypothesis about the relation between the global and local structure, or becomes meaningless otherwise.

All other utterances are only “locally” connected to their environment, i.e. background clauses may be hooked up at any place of the skeleton. (This, admittedly, simplifies the picture, because there may also be global constraints where specific types of background clauses may appear; but this does not affect our general point).

This way of relating global and local features of text structure to each other is clearly a step forward, when compared to, for example, “story grammars”, in which the local organization is largely connected, or to “Markovian” approaches, in which only anaphoric relations, topic continuity and similar features are considered. Still, it is insufficient in at least three respects:

- the foreground-background-distinction, as defined here, is not sufficiently general: similar distinctions obtain in non-narrative texts whose underlying GV involves no temporal order between its components;
- it is unclear how the global constraints - the existence of a skeleton on the one hand, of supplementary material on the other - derive from the nature of the GV with its particular properties.
- it is not clear, how the local constraints follow from the global ones, i.e. *why*, for example, background information is marked by imperfective aspects, by clauses in the passive or by subordinate clauses.

In the following three sections, we will discuss these problems and develop a somewhat different approach.

2.2. Foreground and background in non-narrative texts

The distinction between foreground and background of a narrative is not based on communicative importance although the connotation of these two words suggest this; it may well be, however, that the whole point of the story is in the background utterances⁹. An utterance belongs to the foreground, if and only if it belongs to the “narrative skeleton”. All other utterances belong to the background. Hence, background utterances form a quite heterogeneous class. There is a more serious problem, however: This

⁹Literary texts, as a stylistic device, sometimes reverse the relative weight of communicative importance, which we tend to assign to both components: there is a plot line with a series of sub-events, but what is really interesting happens in the background. [Quite a typical case are almost all the novels by the Russian writers B. and A. Strugatzki.]

definition of the foreground cannot be extended to other types of texts, either because they have no temporal structure at all (such as opinions, arguments, picture descriptions, etc.), or because they have a temporal structure whose individual components, rather than being singular events, are generic events, states, possible happenings, etc., for which it is less obvious to follow a rigid chronological thread - for example those which answer questions such as "What do you remember from your childhood?", "How was the wedding party last Sunday?", "What are your plans for the future?", etc.

In all of these cases, however, there is a *main structure* which functions as a scaffold for the whole text, and various *side structures* which are hooked up at different points of the main structure and which may have different functions. Let us illustrate this by three examples.

(a) "Future plans"

Quaestio: "How do you imagine your future?"

Text: "I have no clear plan yet. Well, first I will finish high school. *Actually that's not sure, because what I would really like to do is to become a musician; but my father won't allow me to.* So, I will go to the university and study something, probably French. And then, I will become a teacher, *although the chances are bad right now.* And then of course, I will marry and have children. *I am very traditional here, and love babies.* There is something else I definitely will do: travelling through East Asia, for at least a year. Maybe I can do this between High School and University."

The quaestio specifies which kind of information has to be given in the text, albeit in a not very restrictive way: what has to be expressed, are events, activities, states which are all in the future and which are more or less certain: in short, possible or desirable events rather than singular, real ones, as in a narrative, and the relation between them is, or may be, temporal; but the quaestio doesn't impose a strict chronological order. Nevertheless, the text clearly contains utterances which directly relate to the quaestio and hence belong to the main structure, on the one hand, and others which give additional, though often important information, on the other; these are underlined in the text above. In this case, the distinction between main structure and side structure is not always straightforward, because the constraints which the quaestio imposes on the answer are less obvious than in narratives with their clear temporal sequencing.

(b) route directions¹⁰

Quaestio: “Can you tell me, where the Goethehouse is?”

Text: “*Yes, but let me think for a moment. I was there myself last week. Yeah, you go down here about three hundred meters, then turn left behind the church. Then, after another three hundred meters, you will come to a square, a very beautiful square. You cross it, carry on and then turn right. You really can't miss it. Then it is the second street to the left, and there you can see it. It is yellow, or yellowish. Okay?*”

The GV on which a route description is based is again not a temporally ordered set of singular, real events but a spatial configuration - a “cognitive map” on which some salient spots (“landmarks”) are spatially interrelated. An appropriate selection of these landmarks (including the deictic origin and the target) and the appropriate arrangement constitute the backbone - the main structure - of a route description. Both selection and arrangement of the landmarks follow certain principles, which we will not discuss here (cf. Klein (1979)). In any event, the backbone is completed by additional information (underlined in the text above). These side structures may have quite different functions; the introductory *yes*, for example, indicates that the speaker is willing and able to answer the quaestio; *there* are comments on the difficulty of the task, a control question (*okay?*), etc. Obviously, side structures of this type may have an important communicative function within the total exchange that constitutes a successful route direction; but they do not belong to the main structure as induced by the quaestio to the text.

(c) opinion (of a verdict)

Quaestio: “Why is the prior court's decision to make Mrs K.'s driving license unfounded?”

Text: “*The court is not in agreement with the previous decision ... Our expert's calculations convincingly show that the speed at which witness L. was driving was at least 85 km/h and at most 95 km/h at the time of the collision. Had the driving of L. been more careful, then Mrs. K would have been at least 5 m over the crossroad. L's driving must be regarded as a gross violation of any traffic rule, and it is indeed disputable whether the fine of 500.- which was inflicted is appropriate to this totally irresponsible*

¹⁰The following example is adapted from Klein (1979), where also a more detailed analysis of route directions can be found.

conduct. On the other hand, Mrs. K doesn't need to reckon with the possibility that a driver exceed the speed limit by 40 km/h, especially in a narrow street”

The GV on which such an opinion is based consists of propositions¹¹ which are primarily connected by logical relations. The propositions themselves are of somewhat different nature. Some concern the matters of the case at issue and hence normally to some real happening in the past, which in turn may consist of a complex of subevents. Others concern attitudes and evaluations of the people involved; the attitudes may also be relevant for the verdict (was the behavior intentional or just careless? Are base motives involved?, etc.) Others concern generic, normative givens - such as the legal regulations, including their interpretation on other occasions. This makes the GV quite heterogenous, and since the quaestio is not too restrictive, either, it is sometimes difficult to decide which utterance directly contributes to answer the “why” and hence belongs to the main structure induced by that quaestio. Nevertheless, most utterances in an opinion can be easily assigned either to foreground or to background (the latter are again underlined in the text above).

Let us briefly sum up this point. The foreground-background-distinction, as it is usually defined, turns out to be a special case of a more general distinction between the “main structure” of a text and various side structures whose form and function may vary. This immediately raises the problem¹² of how to define these two types of structure - especially the main structure, since the side structures are a heterogenous class anyway - if we can't have recourse to the chronological principle, as in the foreground definition of Hopper, Reinhart and others¹³.

¹¹The reader may have noticed that our terminology is somewhat loose here. We will try to clarify it in section 3.2. below. Let us just say that “proposition” corresponds to the content of a declarative sentence.

¹²There is no problem, of course, if foreground and background are defined in terms of “communicative importance” (however this may be done) or of specific linguistic means, such as different aspectual marking. But then, the whole point of the distinction is lost.

¹³There are also some problems with the chronological principle within narratives, for example when two sub-events are explicitly marked as being simultaneous. Thus, “Charles opened the door. At the same time, the phone started ringing” are clearly narrative clauses which belong to the narrative sequence; but the corresponding events do not follow each other, and hence, the utterance violates the Labovian criterion. It is far from trivial to adapt the definition accordingly, because a more liberal definition which would also admit simultaneous events immediately runs into trouble with typical background-foreground sequences such as “We were sitting in the office. The telephone started ringing”. In other words: Two utterances which express (totally or partly) simultaneous events, may both belong to the foreground, or one may belong to the background, the other to the foreground. (They may also both belong to the background, obviously). These possibilities are regularly distinguished by different forms, such as

In what precedes, we have already suggested a possible solution which we will pursue now in more detail.

2.3. Quaestio, focus condition, topic condition

Defining the main structure in terms of a chronological thread of events, is only possible when the nature of the GV allows for such an ordering¹⁴, as is indeed the case for narratives.

But this presupposition is not sufficient. Someone's childhood, for example, also consists of temporally ordered events (and the corresponding feelings and experiences). But a question such as "What do you remember from your early childhood?" elicits quite a different kind of text than from a question such as "What happened to you there and then?", as in the case of a narrative (although, of course, the listener may *interpret* the first question in such a restrictive way that it amounts to the second question). Similarly, a route direction and a sight-seeing description of the same spatial area have different text structure, although they draw on the same stored spatial information: they "foreground" (and "background") different components of the same GV.

The main structure and, as a consequence, the structures of a text are determined by the nature of the underlying GV, on the one hand, and the quaestio of the text, on the other. The quaestio marks specific components as particularly pertinent for the text to be produced. As was mentioned above, the quaestio need not to be asked explicitly; it may result, for example, from the whole communicative context. But even if there is an explicit question, then it may be relatively unspecific, and the real quaestio at issue results from both what is explicitly asked, on the one hand, and additional contextual constraints, on the other.

The function of a quaestio in relation to a text is in principle not different from the function of a normal question in relation to an appropriate answer on utterance level. We may illustrate this function with an old example of Hermann Paul's (1891, p. 218). A sentence such as

- (1) Peter went to Berlin yesterday.

different aspect marking, but this *indicates* the difference and can't be the base of the definition. So, this shows again that the chronological principle is just a special - and often very useful - instance of more general principle.

¹⁴As we shall see in section 2.5., there is a possibility of introducing a subsidiary temporal structure, even if the nature of the information to be expressed is not temporally organized in itself; but this does not affect the present argument.

may be used to answer different questions, and while its grammatical structure (except intonation) remains constant, its “psychological structure”, to use Paul’s term (there has been some terminological progress since) changes according to the question which it answers:

- (2) (a) Where did Peter go yesterday?
- (b) When did Peter go to Berlin?
- (c) Who went to Berlin yesterday?
- (d) What did Peter do yesterday?
- (e) What happened?

After each of these questions, (1) decides on an alternative at issue (the term “alternative” taken in a broad sense: it may comprise more than two candidates). What is different, is the alternative which has to be decided on and actually is decided. After (2a), the alternative is the set of places to which Peter could have gone yesterday, and this alternative is specified by “Berlin”, after (2b) it is the set of (contextually relevant) time spans at which he could have gone to Berlin, and this alternative is specified by “yesterday”, and after (2e), it is the set of (contextually relevant) incidents that could have happened at a certain (contextually given) occasion - this alternative being specified by the whole utterance. Such a set of candidates from which one has to be specified we will call the *topic*, and the specification itself we will call the *focus* of a given utterance.

The terms “topic” and “focus”, as used here, refer to components of the entire information expressed by some utterance, rather than the words or constituents which express this information. In other words, we must distinguish between “topic” and “topic expression”, “focus” and “focus expression”. In the example above, the five questions already define what the five “alternatives at issue”, the five topics are, from which the focus has to be chosen. The corresponding answers express this topic again (except in the last case, see below) and then specify it - i.e. express the focus. Thus, the topic defined by the first question is the set of places to which Peter could have gone yesterday; in the answer, this topic is referred to again by the partial expression *Peter went yesterday* (= topic expression), and the focus expression *Berlin* specifies this topic by giving Berlin as the focus. Obviously, the topic expression in the answer is redundant here, and in fact, it could have been omitted. The relation between topic and topic expression (and similarly focus and focus expression) may be much less straightforward than in this case. In (2e), the alternative at issue is between several possible “happenings” at some relevant

occasion¹⁵, and all we know about these happenings is that they are in the past (due to the tense morpheme of *happened*). This component of the topic, namely being in the past, is expressed again in the answer, but there is no independent “topic expression” in the answer, unless one counts the inflection of *went* as such. We shall return to this problem in section 3.3.

In all of these examples, the topic of the utterance is explicitly raised by a general context, or its expression may be totally left to the utterance itself. All languages provide different devices to refer not only to a place, for example, but also to mark that this reference belongs to the focus (or is the focus) or to the topic, for example intonation, word order or specific particles. If there is a contextually given topic, then these means must be used in accordance with this contextual requirement, of course; otherwise, the utterance is contextually inappropriate. We may think, then, of all declarative utterances as being answers to some “quaestio”, which may be either explicitly given by a real preceding question, which may result in some other way from the context or which may be reconstructable from the way in which topic and focus are marked in the utterance itself.

Let us return now to the quaestio of a text, in contrast to the quaestio of an utterance. In principle, the function of a text quaestio is not different, except that it does not call for the specification of a single referent - for instance, one specific place, time, person, action - but for the specification of a whole structure of such referents and the specification of this structure is distributed over the utterances of the text. Each single referent is taken from the underlying GV, and the quaestio imposes restrictions on the possible referents and their arrangements: it narrows down the set of candidates which are admissible for specification within an utterance, and it restricts the way in which this specification of referents may proceed from one utterance to the next.

The main structure of a text is a restriction of referential movement. It results from the underlying GV, on the one hand, and the quaestio, on the other. This restriction has two components which we will call focus conditions (FC) and topic conditions (TC), respectively. For narrative texts, these conditions may be roughly stated as follows.

¹⁵In all of these cases, there are context influences of a more global nature. Thus, a question such as “What happened?” would be asked in a certain situation, and depending on this particular situation, only certain “happenings” would be acceptable as specifications of the topic. If your friend comes to your room, pale, trembling and covered with sweat, the question “What happened?” clearly means: “What happened that made you pale, trembling and covered with sweat?”. Cf. section 3.1. below.

Main structure of a narrative

- FC: Each utterance specifies a singular event which occupies a definite time interval t_i on the real time axis.
- TC: The time interval of the first event is explicitly introduced (unless contextually given); all subsequent ones follow chronologically, i.e. the interval belonging to the event reported in the n th-utterance is not before the interval which belongs to the event reported in the $(n-1)$ th-utterance.

To put it somewhat differently: The quaestio is a question-function q_i where i ranges over time intervals and each question corresponds to an “utterance-question”:

- “What happened (to you) at t_i ?”
“What happened (to you) at t_{i+1} ?”
“What happened (to you) at t_{i+2} ?”
...
...

and each of the subsequent utterances of the text contains an answer to “what” - it specifies one event from the set of those that could have happened (to the speaker) at that time. Thus, the primary restriction on the events is the definite time interval, although other restrictions are, of course, not excluded. Both FC and TC may be violated by an utterance. This leads to side structures of various types, depending on the kind of violation and the nature of referential linkage to the main structure which still exists. Let us consider some examples of such side structures.

An utterance (or a clause) may serve to specify a time interval in explicit terms, rather than have it simply given by TC. A typical case of this type of side structures are “background clauses” such as *Nous étions a l'étude*¹⁶ or, to use the example from 13 again, *We were sitting in the office*. Most often, subordinate temporal clauses serve exactly this func-

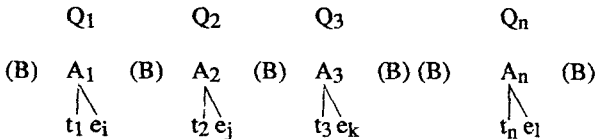
¹⁶As the reader will have noticed, this is the introductory clause of “Madame Bovary”. The full sentence *Nous étions a l'étude quand le proviseur entra (...)* is, by the way, a neat illustration of the fact that the subordinate clause may express foreground information, the main clause background information. This is a special case, however, in that this whole sentence introduces the story. Note that our formulation of TC is such that *Nous étions a l'étude* would not violate TC, since it refers to the first time interval. An alternative way may be to characterize a narrative by two questiones “When and where did the total event happen? What happened to you at t_i ?” where t_i are sub-intervals of the “when” of the total event.

tion, and this is the reason why they contribute to the background: They answer the question "What happened to some person at some time t_i ?"

Other structures don't violate TC, but they do not specify an event, as required by FC. Typical side structures of this type are comments and evaluations, such as *He wouldn't have thought so* or *It was creepy*, etc. There may be some argument here as to what counts as a singular event; for example, an utterance such as *The sky was all red* is normally interpreted as describing a state; but it may be used to refer to an event, as in *Suddenly, the sky was all red*. But neither ambiguities of this kind nor semantic problems of how to define events (in contrast to states, processes, etc.) affect the general principle.

Still other utterances may violate both conditions, for example generic statements inserted at some point in the narrative, such as *Well, that's how life is* or *There is always someone who wants to object*. This loose typology of side structures could be refined in various ways; but we think the general point is sufficiently clear. It is important to keep in mind that side structures may indeed have, and normally do have, some referential linkage to the main structure, as reflected in the use of anaphoric elements, of word order, or merely lexical relations.

We can illustrate what has been said so far in a diagram which shows the relation of main structure and side structures. Q_1, Q_2, \dots are the arguments of the questions function Q_i "What happened to someone at t_i ?", where t_i refers to definite time intervals; A_1, A_2, \dots are the corresponding answers, e_i, e_j, \dots the events. B refers to side structures, which are optional (with the possible exception of the first one, unless the first time specification is contextually given):



Note, incidentally, that the event specifications need not be different, although they normally are. But in *Peter rang. Then, he rang again* the information given about the event specification is the same.

So far, we have dealt with the first two inadequacies of the foreground-background distinction, suggested a more general approach, which seems to overcome these insufficiencies, and illustrated it for narratives. Let us turn now to the third problem, the question *why* main structure utterances are marked by specific linguistic means, side structures by others.

2.4. Referential movement and linguistic means

The constraints on referential movement, as induced by *quaestio* and *Gesamtvorstellung*, apply to the specification of specific contents, not to the linguistic means which are used to express them. But they have consequences for the expression, too: They indirectly restrict the choice of linguistic devices. The nature of these latter restrictions depends on the specific language and the linguistic means which it offers for expression. We shall illustrate this again for narrative texts.

The focus condition stated above requires that each utterance of the main structure has to specify a singular event, which occupies a definite¹⁷ time span on the real time axis - in contrast to habitual events or to states. An utterance which satisfies this condition and hence belongs to the main structure may then have the following linguistic features (depending, of course, on the language in question):

- it is normally marked by perfective aspects (or similar means), in order to indicate that the event occupies a definite time span on the time axis;
- the lexical verb (under the assumption that the lexical verb primarily denotes the event) is not in the scope of certain quantifying or modalizing operators, such as *usually*, *almost*, *not*¹⁸.
- the subject must not be marked as a generic, because otherwise, the whole utterance is normally interpreted as generic, too; this in turn may have consequences for article selection, word order, etc.
- the tensed verb must not be in the "habitual" form (if there is such a marking in the specific language);

etc., etc. Whenever these conditions are violated, the utterance in question is recognisable as contributing to the "background".

Consider another somewhat less straightforward example. Suppose a language has no syntactically determined constituent order but the constituent (or constituents) which corresponds to the topic comes first, the one

¹⁷For the sake of discussion, we will assume that a definite time span is characterized by a right and a left boundary, in contrast to indefinite events; this is clearly disputable; in particular, one boundary may often be enough. For a more detailed discussion of temporal movement in discourse, see v. Stutterheim (1986).

¹⁸There are cases, though, in which a negated verb can be interpreted to denote an event, in the sense of FC. Take an example such as "And then, he didn't show up", where the event is simply, that something which was expected or plausible did not happen.

(or ones) which corresponds to the focus comes last. If Latin were such a language, then the answer to the question *Quid facit Petrus?* would be *venit Petrus*, the answer to the question *Quid facit Petrus?* would be *Petrus venit*. In such a language, all utterances which belong to the main structure of narratives must have the word order *Petrus venit* (under the assumption that the verb refers to a singular event). This restriction on the word order of main structure utterances operates *within* the syntactically admitted word order patterns. Conditions such as FC and TC cannot outweigh obligatory syntactic rules, but they can use options left by these rules.

2.5. Temporal projection

Before elaborating on the idea of referential movement, we will briefly deal with a possible complication of the general picture - temporal projection.

Every text transforms, in accordance with the requirements of the quaestio, a set of information taken from some GV into a linear sequence of utterances. Not everything from this GV need or should be expressed - either because the speaker is entitled to assume that the listener has access to the information anyhow, or because he thinks that, given the quaestio, this information is irrelevant. Whatever survives this process of *selection*, the speaker must in any event transfer a complex set of information into a linear sequence of utterances (*linearization*). How straightforward this linearization is, depends on the nature of the information. In the case of narrative texts, the relevant units are sub-events of a total event, and those sub-events are ordered along the time axis. Some complications aside (see note 13), this order "a then b" of events can be transferred to the order "a' then b'" of corresponding utterances, hence the nature of the GV itself suggests a straightforward linearization of the utterance. Linearization is much more problematic when the underlying GV, as in the case of route directions, apartment descriptions, etc., is a multi-dimensional spatial structure whose units are physical objects; in this case, a multi-dimensional arrangement must be projected onto a one-dimensional array of utterances. A convenient way to solve this problem is the introduction of an ancillary temporal structure. In route directions, this ancillary structure is an "imaginary wandering" (Klein (1979)), that is, a sequence of possible actions of a participant (for example, of the person who asks for route directions); these actions can be chronologically ordered and thus constitute a projection principle which allows the speaker to solve the linearization problem. A similar principle applies in apartment descriptions (Linde/Labov (1975)) and often, but not always, in room descriptions, except that in the latter case, the imaginary wandering is replaced by a -

real or imaginary - “gaze tour” (see Ehrich (1979), Ehrich/Koster (1982)).

The use of an ancillary temporal structure is virtually impossible in the case of essentially “logical” texts, such as an argumentation or an opinion (as it was discussed in 2.1.). There is no uniform principle of how linearization is achieved in these cases, although in practice, there are a number of guide-lines (see for argumentation Klein (1981), for linearization in general, Levelt (1982)).

If there is such a projection principle, whatever its precise form may be, it has to be included in the definition of the main structure: The global constraints, such as FC and TC explained above, do not directly result from the nature of the GV (its components and the relations between them), on the one hand, and the quaestio, on the other, but a special form of projection, such as temporalization, may intervene. How this is done in different types of texts is a matter of empirical research.

2.6. Summary

A text is based on a GV - a complex structure of various information - and a quaestio (or some questiones, cf. note 2), which impose restrictions on its global and local structure: They determine its *main structure* (“foreground”), with the possible intervention of some projection principle. The main structure includes two kinds of restrictions on *referential movement* (i.e., which information has to be specified in an utterance in relation to the preceding utterance): focus condition, and topic condition. Both conditions can be violated; this leads to different types of *side structures* (“background”).

Indirectly, focus condition and topic condition also bear on the choice of linguistic means in main structure and side structure utterances, although these means may vary from language to language.

3. Referential movement

The point of a text is the fact that the entire amount of information to be expressed is distributed over a series of utterances, rather than being patched into a single one. The distribution is not done at random, but is governed by several principles which impose a certain structure on the text¹⁹. In particular, they constrain which information is to be displayed within an utterance relative to the preceding one.

¹⁹There may be, and often are, non-accidental features of text structure above and beyond the principles studied here. For example, there may be cultural habits such as to begin

Let B be the utterance in question, A the preceding one; as a special case, A should also include the “empty” utterance, such that B is the first utterance of the text. Then, the “topic condition” (TC) states that, in the case of narratives,

- B must include a reference to a time interval t_j in the real time axis;
- this time interval t_j must be after the time interval t_i referred to in A (although not necessarily adjacent to that time interval);
- this time reference may be implicit; but if it is implicit, it must not be marked as contributing to the focus of B.

The focus condition (FC) states that

- B must include a reference to an event (in contrast to a state, for example);
- this event must be marked as singular and real (in contrast to generic or possible events);
- the event referred to must be marked as contributing to the focus.

Moreover, the general idea of information distribution over the utterances requires B to contain some *new* information with respect to A; B must achieve some progress, compared to the state reached after A. It should be kept in mind that all of these constraints may be violated, of course, thus giving rise to side structures.

The constraints which FC and TC impose on utterances of the main structure are of two sorts. Firstly, they prescribe or exclude specific contents in some domains of reference, for example temporal reference in this case; other domains of reference, such as reference to place or to

every text of a certain type with *Praise the name of the Lord*; Grimm type fairy tales often end with *Und wenn sie nicht gestorben sind, dann leben sie heute noch* (and if they didn't die, they're still alive). A more interesting case are prayers or magical formula whose underlying organisational principles are largely unknown. We simply do not know why, in a love magic, the utterances must be ordered in a certain way to achieve the intended effect. - To avoid misunderstandings, it should be emphasized that the constraints we are talking about here do not totally *determine* the text structure; they rather narrow down the options in a certain way, and depending on the case, these restrictions may not be particularly tight, anyway (cf. also note 2).

persons involved are not constrained, although this may be different in other text types than narratives. Secondly, they prescribe whether a certain component of the total content of the utterance goes to its topic or to its focus. Thus, both time and event referred to in a main structure utterance of a narrative cannot be maintained from the preceding utterance, but the shift in the former domain concerns the topic, whereas the new event specification goes to the focus.

In what follows, we will first have a look at the various domains which may be affected by these constraints (section 3.1.) and then at the various ways in which reference *within* such a domain may move from one utterance to the next (section 3.2.); finally, we will discuss a number of open problems (section 3.3.).

3.1. Referential domains and their interrelation in utterances

With every utterance, the speaker puts a segment of the underlying GV into words. This segment may include an event, a state of affairs, some spatial arrangement, or whatever - depending on the nature of the GV. We will call such a segment a *Sachverhalt* (borrowing a term from Wittgenstein's). We will not consider here what qualifies a specific clustering of components from the GV as a *Sachverhalt* which may be expressed in an utterance, although this is clearly not a trivial question. In any event, a *Sachverhalt* is some arrangement of various temporal, spatial, personal and other features which are interrelated in a specific way.

Not every ingredient of a given *Sachverhalt* is indeed expressed. Suppose the underlying GV is a witnessed accident, and the speaker has to give an eyewitness report of that accident (with the *quaestio* "What did you hear and see?"). Then the utterance

(3) She drove against the signpost.

renders a segment of the underlying GV - a *Sachverhalt*. But clearly, it does not express everything that belongs to that *Sachverhalt* and that the speaker could have packed into it. For example, it does not mention

- the speed at which she was driving
- the kind of vehicle
- the place where all of this happened
- the direction from which the driver came

- the shape of the signpost
- the approximate age of the driver
- the fact that she was not wearing sun glasses

and so on - to mention only a few features which could be relevant, indeed, in any eyewitness report. The speaker has taken a specific choice among the features he could refer to in his utterance²⁰. The choice depends on (a) what he thinks to be important that the listener know, (b) what he assumes not to be accessible to the listener from other sources of knowledge (context information), and (c) on the structural constraints of the language in question (English normally requires reference to a subject and to the event time, although these may be irrelevant or redundant).

The listener may know some of the non-mentioned features, though surely not all of them. For example, he may know from previous utterances what the place of the whole event is and that *she* refers to an elderly lady; similarly, he may infer from the whole context that she was driving a limousine, rather than a bulldozer. It seems useful to distinguish two types of contextual information which complete the listener's interpretation of the utterance, above and beyond what is made explicit by linguistic means. First, there is contextual information which is directly linked to context-dependant verbal elements in the utterance, such as deixis, anaphora, ellipsis. The interpretation of an utterance such as *Me, too* is based on knowledge of the meaning of deictic words and the rules of ellipsis in English, on the one hand, and on access to the necessary contextual information, on the other (roughly, the listener must be able to identify who is speaking, and must have heard the previous utterance). In these cases, we will talk of *structure-based* or *regular* context-dependency. The integration of linguistic information proper and of what can be derived from structure-based context-dependency provides the listener with a first interpretation, which we will call *proposition*. In addition, the listener may *infer*, with various degrees of certainty, other features of the Sachverhalt, such as the type of vehicle or the appropriate speed; this reference is not directly linked to structural means but related to the proposition in a less explicit way. Therefore, *inference* is less accessible to linguistic analysis than structure-based context-dependency; but it is no less important for text organization and more specifically, for referential movement. Consider a sequence of two utterances such as

²⁰Note that this is in principle not different for fictitious Gesamtvorstellungen, where the underlying information, or parts of it, do not stem from perception and memory, but from imagination.

- (4) Yesterday, I went to Heidelberg. My parents-in-law celebrated their silver wedding.

The first utterance introduces, among other things, a place - the target position of the movement. The second utterance contains no spatial reference at all. Still, we tend to infer that this wedding party is in Heidelberg: The spatial reference, taken from the focus of the previous utterance, is maintained. This inference is not certain (the second utterance could continue ... *I tried to escape the party*), and if the speaker had wanted to avoid this uncertainty, he could have chosen another maintenance technique, for example by adding a spatial anaphor like *there*, thus relying on structure-based context-dependency rather than on inference.

Let us sum up what has been said so far: When talking about the meaning expressed in an utterance within a text, we must distinguish between four complex clusters of various temporal, spatial, personal and other than *linguistic meaning* based on the lexical meaning of the words on the way in which these words are fused into higher units (i.e. syntax), the *proposition*, where structure-based contextual information is added, the *utterance interpretation*, which enriches the proposition by all sorts of inferable information, and the *Sachverhalt*, which comprises also non-inferred information²¹. In a diagram:

- (5) Sachverhalt (all features)

utterance interpretation
proposition

non-inferable features
contextual features II
(derived through inference)

linguistic meaning
(derived from lexical
meaning and syntax)

contextual features I
(derived through structure-
based context-dependency)

Referential movement, as understood here, is on the level of utterance interpretation, and in order to understand how referential movement works, we must look at the way in which this interpretation proceeds from

²¹There is a familiar distinction between sentence meaning and utterance meaning, where the former roughly corresponds to our "linguistic meaning" and the latter to our "proposition". We have avoided this terminology, especially the term "sentence meaning", since we also want to include the meaning of utterances such as *Me, too* or *She him* or *Why four?*, which one would not consider to be sentences, but which may function as perfect utterances in a text.

utterance to utterance. In what follows, however, we shall not particularly be concerned with those processes which lead from the proposition to the utterance interpretation since they are more on a cognitive than on a linguistic level (for a selection of recent studies of inference processes, see Rickheit/Strohner (eds.) (1985)). Whenever necessary, we will briefly say "by inference". So, we will be mainly concerned with the transition from proposition to proposition.

A proposition selects some of the many features which constitute the Sachverhalt and integrates them in a particular way. Consider, for example, the proposition resulting from uttering (6) in some context:

(6) Yesterday, the Millers left for Heidelberg.

The proposition contains a specific event which includes, among others, the following features:

1. A time interval, within which the event happened (expressed by *yesterday* and verb inflection);
2. A participant (the Millers, encoded by the grammatical subject);
3. the "activity as such", i.e. the leaving for Heidelberg, which is in turn compound of various features - minimally a target position (Heidelberg) and a change in position.

In addition, there are several other features which belong to the Sachverhalt and perhaps to the utterance interpretation but which are not referred to in (6), for example.

4. the place at which the activity begins (the source position of the movement);
5. the specific circumstances under which it happens, such as causes, means, purposes, etc.

In what follows, we shall assume that these five types of features represent five *referential domains*, and that a proposition is made up from a combination of features from these domains.

Not all utterances express specific events. They may also render specific states (as *Yesterday, the Millers were in Heidelberg*), property assignments (*The Millers are sweet*), generic or habitual events (*During the winter, the Millers live in Heidelberg*), and maybe others. To account for this, we need two refinements. First, we will replace the referential domain

“activity” by the more general “predicate” which will also include property assignments, states, processes, etc. Second, we shall assume that an utterance also contains a reference to a modality; roughly speaking, it is somehow related to a real, a fictitious, a hypothetical world. Admittedly, this is simply a way to circumvent a whole range of complicated problems, but it will do for our present purposes²². This leaves us with six referential domains:

1. R_t: reference to time interval
2. R_e: reference to places
3. R_c: reference to circumstances
4. R_p: reference to participants
5. R_a: reference to predicates of various types
6. R_m: reference to modality (real, fictitious, etc.)

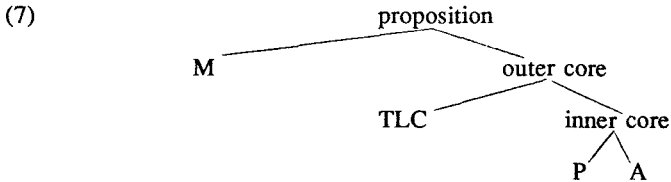
An utterance refers to selected features from these six domains and integrates them into a proposition. This may be done in a multitude of ways; in particular:

- not all domains must be represented; in utterances such as *It was raining* or *There will be dancing* no participant is referred to although there may be one in the Sachverhalt (at least in the latter example);
- features from one domain may show up several times in either the same or different functions; cf. utterances such as *Yesterday, they left at five* or *In Heidelberg, they walked from the station to the castle* or *The man at the castle looked like an alien*;
- reference to time, to place, to circumstances may be conflated in one concept, as in *On many occasions, there was dancing*

²²There is another somehow related category which we will not consider here: the generic, habitual or specific status of a proposition. The “generic nature” of utterances like *Germans live in caves* or *Peach ice cream tastes worse than you would have thought* is not independent of other referential domains; it results from non-specific participant reference, or from non-specific time reference. But pursuing this issue would lead us away from our central problems.

to mention only a few of the complications. On the other hand, rigid syntactic rules and other ossifications of expressive power may predetermine or at least bias the speaker's freedom in selecting and combining various features. The structure of English, for example, makes it difficult to produce an utterance without referring to a participant and a time, because subject and tense marking of the verb are obligatory. We cannot pursue all of the ramifications of these problems here. We will start from a sort of "basic structure" and then come back to some complications in section 3.3.

Traditionally, it is most often assumed that reference to a participant P (often encoded by the grammatical subject) and reference to a predicate A (often encoded by the grammatical predicate) constitute something like the "inner core" of a proposition, which is then further characterized by a time T, a space L, and maybe some circumstances C; the resulting structure, the "outer core", is then related (M) to some real or fictitious world. We will adopt here this conventional picture, arguable as it may be. This gives us a "basic structure":



Depending on the nature of the underlying Sachverhalt, this basic structure may be reduced by one or several components. A mathematical theorem, for example, does not have a time or a place to be referred to; so, its basic structure is reduced by at least two parameters, compared to (7). This is not to be confused with basic structure where some domain is not *explicitly* referred to, although the Sachverhalt as such would allow such a reference.

Compare again the propositions expressed by the utterances²³ *It was raining* and *There was dancing*.

In the first example, there is no participant involved, hence the basic structure is reduced by this component; in the latter case, there is a participant involved - the dancer or dancers -, but it is not specified; the basic structure, however, is that of (7).

²³Actually, each of these utterances may refer to infinitely many propositions, since they contain a deictic component (tense marking) and hence the time referred to may vary, depending on speech time. But we will ignore this complication here.

Note that (7) refers to the way in which the *proposition* is organized, not to the way in which the utterance is constructed. The way in which time, place, participant, etc. are indeed referred to, may be very different, depending on the language-specific means. It may also be that the expression which has this function is very complex and uses features from some other domain. For example, reference to the participant may use spatial or temporal information, as in *The man at the corner* or *Poets from the 19th century*. We will return to this point in a moment. A most simple realisation of a basic structure like (7) would look like

(8) There and then and thus, she did such and such.

where *there* refers to the place, *then* to the time, *thus* to the circumstances, *she* to the participant, *did* to the modality (and the time and, perhaps, the predicate, too) and *such and such* to the predicate. In this case, the linguistic meaning contributes not very much to the proposition; this does not mean that the proposition itself is poor in content; rather, most of it stems from structure-based context-dependency. Normally, the linguistic contribution is richer, of course, and we shall return to this issue in section 3.3. below.

3.2. Types of referential movement

We may think of the components of a basic structure as “open slots” to be filled appropriately in order to yield a proposition: a place must be referred to which specifies position L of the basic structure; similarly a time which specifies position T of the basic structure, and so on. This specification may be introduced in this utterance for the first time, or it may be maintained from the specification of a preceding utterance. It is a simplification, however, just to talk about introduction and maintenance of reference. In what follows, we will give a somewhat refined typology of referential movement. This typology does not relate to the specific linguistic forms which express the reference in question, but to content only.

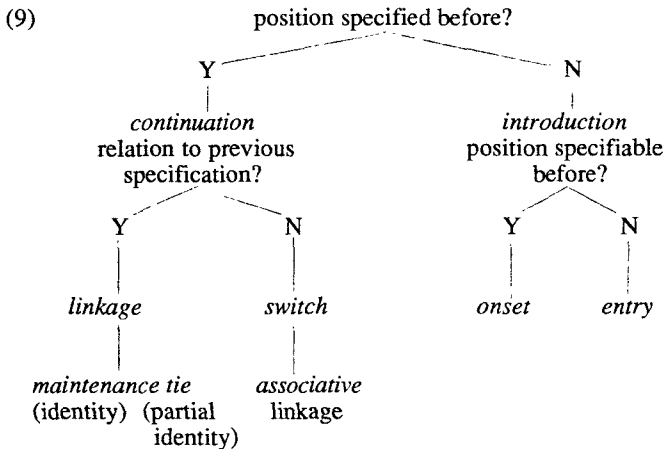
First, we must distinguish as to whether a certain referential domain, say P, was specified in the preceding utterance or not. In the former case, we will talk of *continuation*, in the latter, of *introduction*; note, that continuation does not necessarily involve identity of a referent: it just means that the domain in question, for example place, was specified before, no matter how. In the case of an introduction, it may be that a specification in the previous utterance was impossible, because the *Sachverhalt* there did not allow it - there is no position P in the basic structure, for example - or because there is such a position but for some reason, the speaker did not

specify it. The latter case we will call *onset*, the former *entry*; in practice, this distinction is of minor importance, however.

Consider now the various possibilities of continuation. There is again an important distinction between *linkage* and *switch*. In the former case, the specification is related to the content of the previous specification, although this relation need not be identity; in the latter, there is a change of specification without referring back, or using the previous specification. There are at least three types of linkage. First, the reference specified may be indeed identical; this is the pure case of *maintenance*; note, again, that this term refers to maintenance of a referent, not of an expression. Maintenance may be expressed indeed by repetition of an expression (*Strauss-Strauss*), but also by an anaphoric term (*Strauss-he*) or even by another descriptive expression which refers to the same entity (*Strauss-the king of the Vienna waltz*). Next, it may be that there is an anaphoric linkage, but still, a new referent is introduced; we shall call this type *tie*. Such a tie may be expressed by words such as *thereafter* or even *then* in sequences such as *He closed the door. Then, he opened it again*, where *then* means something like “at a time t_j after the time t_i referred to before”.

Third, there may also be a vague linkage which we will call *associative*; it shows up in cases where, for example, a mountain is introduced and the second utterance refers to *the valley* or *the summit*. Linguistically, this type of linkage is hard to grasp; but its importance for referential movement and for text structure in general is obvious. A *switch*, finally, is in a sense comparable to an introduction, except that the position in question was specified before. Therefore, a switch often has a contrastive function. Thus, in a sequence such as *It was strange. Peter cleaned the dishes*, the reference to the participant “Peter” is an introduction (more precisely, an entry), whereas in *Mary slept. Peter cleaned the dishes*, it is a switch.

Let us sum up this typology in a diagram:



Among the possible continuations, we group all of those together (as *linkages*) which maintain at least part of the preceding specification. We might as well take those together in which at least part of the information is changed - tie, associative linkage, switch - and label them as *shifts*, in contrast to full maintenance.

3.3. Some complications

Referential movement is from one proposition to the next, each proposition being a complex web of features which belong to different referential domains; this web results from an interplay of linguistic meaning - lexical meaning of the words and the way in which these are put together - and context. In an utterance such as (8), the various referential domains are quite distinctly teased apart, that is, there is one expression (*she*) for reference to the participant, one expression for reference to the place (*there*), etc., and only *did* merges some of the referential domains. These domain-specific expressions have almost no descriptive content in this case. The lexical meaning of *there*, for example, makes clear that the referent is a place, and if this reference is understood indeed, then this is only due to the fact that the place in question was referred to before. Normally, successful reference needs much more descriptive information. This information is provided by words with a richer lexical content or by syntactically compound expressions, or both. Then, however, the relation between expressions and features expressed becomes much less straight-

forward than in (8). This has many consequences for referential movement, three of which will be discussed in the sequel.

3.3.1. Simple and compound reference

An expression such as *there* is simple in two ways. First, it is one word, in contrast to syntactically compound spatial expressions, such as *at the castle*, *in front of the house* or *between here and there*. Second, it contains only spatial features, in contrast to for example a verb such as *come*, which contains spatial, but also other features; *there* is homogenous with respect to the referential domains which we have distinguished above, whereas *come* is heterogeneous in this respect. Such a “clustering” of features also appears in syntactically compound expressions, and this often constitutes a problem for referential movement.

An expression such as *at the castle* is syntactically compound, but homogeneous²⁴: it refers to a place. This reference may “fill” the appropriate position of the basic structure. But it may also be used to support reference to a participant - for example in a complex expression such as *the man at the castle*, which is heterogeneous. In this case, *at the castle*, while still being a reference to a place, cannot fill the place coordinate of a basic structure, and hence cannot be maintained as the place reference of some subsequent proposition, for example by the use of anaphoric *there*:

- (10) The man at the castle was better informed than our travel guide.
There, ...

There is appropriate here only if it is clear from some other contextual information that the “locus” of the whole action is at the castle, but not as direct anaphoric maintenance from the first of the two utterances. It is not true, however, that anaphoric linkage could not cross the referential positions of the basic structure. Consider an example where a place reference functions as a part of the predicate reference, as in the compound predicate *being at the castle*:

- (11) We were at the castle. There, ...

Here, anaphoric linkage is clearly possible, or, to put it slightly differently, the place introduced in the first utterance, where it is part of the predicate, is *accessible* to anaphoric maintenance within the basic structure. This is quite typical for compound predicates. It is difficult to

²⁴This is arguable, since one might argue that *castle* contains more than just spatial features. But this does not concern the point of the argument.

say what is responsible for these differences in accessibility as exemplified by (10) and (11). The type of compoundedness is one factor, but clearly not the only one. Moreover, accessibility to anaphoric maintenance often correlates with accessibility to other semantic processes, such as the possibility of being marked as topic or modifiability by an adverb, to which we will turn now.

3.3.2. Accessibility of information to semantic processes

The Sachverhalt rendered (in an appropriate context) by

(12) In 800, Leo crowned him.

may also be rendered by

(13) In 800, Leo put the crown on his head.

Time and participant are the same, and the grammatical predicate refers to the same action; but in the second case, some of the semantic features implicitly contained in *crowning him* are singled out and referred to explicitly. These are the object which is moved from some initial position to some target position (=the crown) and the target position (=on his head). Other components of the predicate are unaffected, such as the person involved as a receiver and denoted by *him/his*. The “singling out” of two components makes them accessible to anaphoric processes. Thus, (13) but not (12) allows the continuation:

(13) It looked splendid there.

Similarly, the decomposed predicate allows a more subtle topic-focus-assignment of features. In (12), all features of the predicate except the receiver belong either to the topic or to the focus. Thus, it cannot be used to answer the question *Where did Leo put the crown?* or *What did Leo put on his head?* whereas (13) allows for a much more selective assignment of features to topic or focus; hence it may be appropriately used to answer all of these questions (and some more).

Thirdly, when features are “encapsulated” in a single lexical item, they offer limited access to further modification. Thus, the *crowning* from (12) implies a crown, as is evidenced by the possible continuation *The crown was splendid* (with a definite article). But this implicit crown cannot be further specified so long as it is only implicit. This is not to mean that no feature within *crowning* is accessible; adverbials, such as *rapidly*, may easily address temporal characteristics of the predicate. Coming back to

referential movement, it does not only matter whether a referent is introduced at all but also how - as an implicit feature of some lexical item, such as *the crown* in crowning, as a syntactically separate part of some semantically “heterogeneous” expression, such as the spatial reference by *at the castle in the man at the castle* or finally as some independently accessible, explicitly specified referent, as *the crown* or *in 800* in (13).

There is a third, in a sense complementary, problem with maintaining reference:

If several features are available for anaphoric maintenance, which among them are picked up and maintained by a specific anaphoric device? We will briefly discuss the “bundling” of features.

3.3.3. Bundling of maintained features

An utterance such as

(14) Yesterday, the Millers left for Heidelberg.

introduces a complex proposition, including a time, a participant, and other information which is then - with the restrictions mentioned above - available for maintenance and further elaboration. An anaphoric term may pick out some referent in a selective way, such as *there* for place, *they* for the participant, etc. But there are also anaphoric terms which bundle various types of information, for example *this*. Consider the following four possible continuations:

- (15) (a) We may do this, as well.
(b) This was a surprise. We thought they would go to Saarbrücken.
(c) This was a surprise. We thought they had already left last week.
(d) This was a surprise. Everyone thought they would stay in München.

In all of these cases, *this* picks up a different bundle of features among those which were introduced before. Thus, it is quite unselective with respect to referential movement: *this* maintains the “central” feature, or features, of a proposition, which are contained in the predicate, and an arbitrary share of “peripheral” features, namely all of those components of the basic structure which are *not* freshly specified.

4. Conclusions

Any coherent text has its underlying quaestio (or quaestiones) which it is meant to answer. This quaestio not only defines main structure and side structures of the text - its "foreground" and its "background"; it also imposes constraints on what belongs to the topic and what to the focus within an individual utterance of the main structure and how reference within certain domains proceeds from one utterance to the next. Any concrete attempt to state the conditions of referential movement has to face a number of problems which result essentially from the fact that there is no simple one-to-one mapping between these elements of the content, for which referential movement and topic-focus-assignment are defined, and the linguistic means which serve to express these content elements.

Bibliography

- van Dijk, T. A. (1980), *Macrostructures*. Hillsdale, N.J.
- Ehlich, K. (1980) (ed.), *Erzählen im Alltag*. Frankfurt.
- Ehrich, V. (1979), *Wohnraumbeschreibungen*. *Zeitschrift für Literaturwissenschaft und Linguistik* 33, 58-83.
- Ehrich, V./Koster, Ch. (1982), *Discourse organization and sentence form. The structure of room descriptions in Dutch*. *Discourse processes* 6, 169/195.
- Fleischmann, K. (1973), *Verbstellung und Relieftheorie*. München.
- Givon, T. (1983) (ed.), *Topic continuity in discourse: a quantitative cross-language study*. Amsterdam.
- Halliday, M. A. K./Hasano, R. (1976), *Cohesion in English*. London.
- Hoppe-Graff, S. (1984), *Verstehen als kognitiver Prozess. Psychologische Ansätze und Beiträge zum Textverstehen*. *Zeitschrift für Literaturwissenschaft und Linguistik* 55, 10-37.
- Hopper, P. I. (1979), *Aspect and Foregrounding in Discourse*. In: *Discourse and Syntax*, ed. by T. Givon, New York, 213-241.
- Klann-Delius, G. et al. (1986), *Untersuchungen zur Entwicklung von Diskursfähigkeit am Beispiel von Spielerklärungen*. *Linguistische Arbeiten und Berichte* 21. Berlin.
- Klein, W. (1979), *Wegaukünfte*. *Zeitschrift für Literaturwissenschaft und Linguistik* 33, 9/57.
- (1981), *Logik der Argumentation*. In: *Dialogforschung. Jahrbuch 1980 des Instituts für deutsche Sprache*. Düsseldorf, 227-264.
- Labov, W. (1972), *The Transformation of Experience in Narrative Syntax*. In: *Labov, W. Language in the Inner City*. Philadelphia, 354-396.

- Labov, W./Waletzky, J. (1967), *Oral Versions of Personal Experience. Essays on the verbal and visual arts*, ed. by J. Helm. Seattle.
- Levelt, W. J. M. (1982), The speaker's linearization problem. *Philological Transactions of the Royal Society in London, Series B*, 295, 305-315.
- Linde, C./Labov, W. (1975), Spatial networks as a site for the study of language and thought. *Language* 51, 924-939.
- Marslen-Wilson, W. et al. (1982), Producing interpretable discourse: the establishment and maintenance of referents. In: *Speech, Place, and Action*, ed. by Robert Jarvella and Wolfgang Klein, Chichester, 339-378.
- Paul, H. (1896), *Prinzipien der Sprachgeschichte*. (3rd printing). Halle.
- Quasthoff, U. (1980), *Erzählen in Gesprächen*. Tübingen.
- Reinhart, T. (1984), Principles of gestalt perception in the temporal organization of narrative texts. *Linguistics* 22-6, 779-809.
- Rickheit, G./Strohner, H. (1985) (eds.), *Inferences in Text Processing*. Amsterdam.
- von Stutterheim, Ch. (1986), *Temporalität im Zweitspracherwerb*. Berlin.
- Thompson, Sandra (1984), *Subordination and narrative event structure*. Ms.
- Weinrich, H. (1964), *Tempus. Besprochene und erzählte Welt*. Stuttgart.
- Wundt, W. (1912), *Völkerpsychologie*, Bd. 1: *Die Sprache* (2nd printing). Leipzig.