Carroll, M., Stutterheim, C. von, & Klein, W. (2003). Two ways of construing complex temporal structures. In F. Lenz (Ed.), *Deictic conceptualisation of time, space and person*. (pp. 97–134). Amsterdam: Benjamins.

Christiane von Stutterheim, Mary Carroll and Wolfgang Klein

Two ways of construing complex temporal structures

1. Introduction

When faced with a complex verbal task, such as giving an account of the most unpleasant event in your life, describing the face of your dentist, or explaining why jogging is healthy, there are many options open to the speaker in accomplishing this task. Despite the wealth of variation which one may encounter, the result in each case is a text – a sequence of utterances which is structurally related. The shape of the text is to a large extent determined by the *quaestio*, that is, the explicit or implicit question which the speaker sets out to answer. Thus, the eye witness of a car accident will come up with very different texts, depending on whether the quaestio is 'What did you see?', 'What happened? or 'Who, to your mind, was responsible for the accident?'. The quaestio partitions the text into a main structure, i.e., those utterances which directly answer the quaestio, and side structures, i.e., utterances which supply kinds types of associated information. It also determines the topic-focus organisation of main structure utterances and thus imposes guidelines on the way information is selected and coded at utterance level (see, in detail, Klein and von Stutterheim 1987, von Stutterheim 1997). Although the quaestio clearly channels some of the main coding options available to the speaker, it does not take care of all decisions which the speaker will make in constructing a coherent text. The information organisation required to meet such demands can be described with respect to different structural aspects, the two most crucial being thematic continuity (manifested in terms of causal/temporal relations in the case of dynamic situations) and referential framing[1]. By referential framing we mean something which – in the basic sense of this notion – has been described in classical poetics as the unit of time and space. Phrased in modern semantic terms, propositions (predicate/argument structures) have to be referentially bound with respect to times and worlds and possibly also spaces. In order for a sequence of propositions to be coherent, these referential properties have to be related consistently across utterances. There is long tradition of studies which deal with these features of text structure under such headings as *macrostructural planning*, *anaphoric linkage, mental models,* etc. (Kintsch & van Dijk 1978, Brown and Yule 1983, Garrod & Sanford 1988, Ochs et al 1992, Goldberg 1996).

So, in addition to the guidelines set by the quaestio, what determines the actual linguistic options chosen when constructing a coherent text – is it just personal preferences, cultural or social habits, or do we have to take into account the influence of the specific encoding system, of the particular language which the speaker uses? The question is, do speakers of English and speakers of German, for example, having witnessed the same accident, *systematically* follow different strategies in text construction when they set out to answer the question 'What happened?'. This is the issue we investigate in this paper. More specifically, we shall look at the

way in which English and German speakers typically organise and convey temporal information when answering the text quaestio 'what happened?', since temporal information is a crucial part of the complex verbal task.

Roughly speaking, temporal information involves two components. First, the 'entire event', here the entire car accident, must be embedded in space and time. In the present context, we are only interested in the temporal side of this embedding or framing. This is the place where *deixis* comes in: tense and (among others) deictic adverbials relate the entire event to the (here-and-)now of the speech situation. Second, the 'entire event' has a complex temporal structure in itself: it consists of a number of sub-events and states which are temporally related to each other: they may follow each other, be contained in each other, partly overlap etc. In order to convey its temporal make up, the speaker must split the entire event into subevents and phases within these subevents and encode these, and he or she must mark what comes when, or during what other event. This is the place where *aspect* and *Aktionsart* as well as various types of *anaphoric linkage* become relevant. The mere fact that English and German differ with respect to these linguistic devices might be seen as an answer to our initial question. German, for example, has no progressive, that is, no formal means such as the *-ing* (as in *they are falling)* to code systematically the ongoing aspect of an event on the verb, and this should have consequences for the way in which German speakers present temporal information. But first, the absence of a progressive marker does not mean that German speakers cannot present events as ongoing (or whatever the overall function of the English progressive may be). And furthermore, it does not mean that within a given language, speakers cannot follow different strategies in the way in which they approach the temporal organisation of their complex verbal task.

The ways in which languages encode time is surely one of the most-researched areas in linguistics. But most of this work deals with the temporal structure of individual sentences, and here in particular with what verbal categories such as tense, aspect and Aktionsart contribute to this structure (Reichenbach 1947, Partee 1983, Bäuerle 1979, Comrie 1976, Fabricius-Hansen 1986). Research which goes beyond the individual sentence, for example in Discourse Representation Theory (Kamp and Reyle 1993, Roßdeutscher 2000), is still the exception rather than the rule. Moreover, it is usually confined to the anaphorical relations between two or three adjacent sentences, and does not therefore address the temporal makeup of a whole 'event' and how it is rendered in the entire text. There is a further major problem in the research tradition on temporality: it has a strong bias towards 'individual events in reality', and canonical definitions of tense such as 'a verbal category which localises the event before, after or around the moment of speech' simply do not work for many text types. This and some other problems of the research tradition will be discussed in more detail in the next section. In conclusion, there is as yet no satisfactory analytical framework for the kind of analysis we have in mind here. We will largely adopt the framework proposed in Klein (1994, 2000a), together with some adaptations to be explained in the next section. In sections 3 and 4, we will report on two empirical studies, showing that there are substantial differences in the way in which speakers of English and speakers of German approach the task. The concluding section relates the empirical findings to the grammatical structure of these two languages.

2. Three temporal parameters

Ever since Aristotle, a tensed verb form is assumed to serve two functions: it describes a 'situation' (i.e., event, process, state), and it localises this situation in time by relating it to the moment of speech: the situation may be before, after, or – more or less – simultaneous with the moment of speech. Hence, there are two temporal parameters – the time of speaking and the time

of the situation – and three temporal relations between them: *before, after, simultaneous to*. This idea may be right, but it is surely insufficient as an account of a fully developed tense system. Classical Greek, for example, has more than three tense forms, and so do most Indoeuropean languages. One way of dealing with this problem is to assume a third temporal parameter, thus allowing more than three temporal relations. This idea was very common in 19th century linguistics (see, for example, Paul's influential *Principien der Sprachgeschichte* from 1882). Nowadays it is often associated with the name of Hans Reichenbach and his three time points: the 'point of speech' S, the 'point of event' E, and the 'point of reference' R (Reichenbach 1947). Reichenbach used his system for English; but it has found wide popularity in the analysis of other languages as well.

This is remarkable since it raises many problems. One of them has to do with the idea that Reichenbach speaks of 'points' rather than of 'time spans' or 'intervals' which can occupy a certain stretch of time, as is obviously the case for events as well as for the time of speech (the case is less clear for the 'reference time'). Most followers of Reichenbach adopt his S, R and E, but interpret them as intervals, which can have a certain measurable duration, instead of points. A second set of problems has to do with the concrete analysis of tense forms in a specific language. Should the German Perfekt be analysed as 'E before (S simultaneous to R)', like the English present perfect, or should it be analysed as '(E simultaneous to R) before S', like the German Präteritum? In the present context, we are less interested in these specific questions than in a number of problems that are quite independent of particular forms and what they mean. These problems have to do with the definition of the three temporal parameters (interpreted as intervals) and hence touch upon the very foundation of such an account. They will now be dicussed in turn.

2.1 Problems with S, R and E

There are at least four substantial problems with the notion of S, the time of speech.

(a)   In the first place, S seems to be defined for the 'canonical speech situation', in which speaker and hearer are present in a well-defined spatio-temporal setting. But what counts as S in written texts – the moment of writing or the moment of reading? There is no uniform answer. In written instructions, for example, it is typically the time at which a specific reader reads the instructions (*now, you turn the handle ...*); on a postcard sent while on holidays, it is typically the time at which he postcard is written (*tomorrow we will visit the Palazzo*).

(b)   What counts as 'time of speech', if the speech is very long – what, for example, was S in Lincoln's 'Gettysburg address'? Does each utterance have its own S? Even a single utterance may require two different temporal anchors, as illustrated by sentences such as *From now, it is two seconds to now*. This problem often goes hand in hand with the first problem: which S has Proust in 'A la recherche du temps perdu?'. Or take the case of complex instructions, such as route directions: *Now, you take the first street to the right, continue on to the first traffic light, and now/then cross ...*. Note that the problem arises for the definition of the tense forms as well as for the interpretation of the temporal adverbials.

(c)   There are many text types for which the actual moment of speech (or writing) apparently plays no role whatsoever. A good example are legal or administrative texts. Take a simple sentence such as *The court will decide no later than two weeks after the application has been submitted*. It includes a future tense form and a present perfect, none of which is related to the moment of writing or the moment of reading this sentence (for a detailed argument on this point,

see Klein 2000b). An even more salient case are all sorts of fictitious texts, from historical novels to science fiction stories, from fairy tales to film retellings. Their plot is simply not 'in this world', and hence, it does not make sense to position them on the time line in relation to the moment of speech. Conventionally, most of them are told in the simple past; but the present tense is not infrequent. In either case, their temporal anchor cannot be the moment of speaking, writing or reading.

(d)　　Even with the 'most canonical use type', that is, when talking to someone about singular events in the past, the usual definition of tense forms is often wrong. Take for example the common definition of the English present tense: 'time of situation at or around moment of speech'. There are many cases, in which the situation is clearly in the past, and still, the present tense is used. Two well-known cases are the 'historical present', as in *And then along comes this guy and drops down just in front of the shop,* and the 'praesens tabulare', as in *In the year of 1809, he publishes 'Die Wahlverwandtschaften' and begins to work on the second part of 'Faust'.*

These problems shed considerable doubt on the classical understanding of 'deixis' – the 'here-and-now of the speech situation' in the analysis of temporal expressions, be it tense forms or temporal adverbials. It works in some cases, but there are at least as many in which it does not. Hence, we must replace the notion of deictic speech time by a somewhat broader notion of temporal anchoring (see section 3.2.).

　　Let us turn now to R, the 'time of reference'. The problems here are of a very different nature: Reichenbach never gave a definition of this notion; he considers it to be a time (in his case, a point), which is somehow given in context. This, incidentally, is in remarkable contrast to earlier work on temporality which operates with a third parameter, such as Paul, who considers it to be some time from which something is imagined. More surprising is the fact that the literature following Reichenbach did not bother to define it either. But it seems obvious that the potential role of the third temporal parameter cannot be evaluated without a meaningful characterisation of this notion; R cannot just be an arbitrary time span. The notion of 'topic time', to be discussed below, is such an interpretation.

　　This leaves us with E, the 'time of the situation'. Note, first, that Reichenbach did not consider E to be the 'event' but its 'time'. Tense forms do not localise events but temporal intervals, to which such an event is related. He did not explain the nature of this relation, except that the temporal interval associated with some event was considered to have no extension. In subsequent work on E this idea was for the most part abandoned. But the precise relation between 'the situation' and 'its time' is hardly ever addressed. In fact, it raises a number of problems, primarily because situations in themselves often have a complex internal temporal structure.

　　Consider, for example, a sentence such as *He planned to have left by five.* Clearly, the entire situation, which is selectively described by this sentence, involves a number of different temporal intervals – the time of his planning (t_1), the time of his leaving (t_2); and the time at which this leaving is over (t_3). There are the temporal relations between these intervals: t_1 precedes t_2 as well as t_3, t_3 follows t_2. There is no single 'situation time'. It is possible to define such a time, for example as the minimal interval ts, which includes t_1, t_2, and t_3. But such a notion is of little use, tense markings as well as temporal adverbials do not operate on an 'entire situation time' but on its individual components. Thus, the temporal expression *by five* does not specify ts but t_3, i.e., the time at which his leaving is completed. Note that this is different for *At five, he planned to have left the city,* where *at five* indicates the time of his planning, i.e., t_1. It is also different in *He planned to have left the city at five*, which is a bit odd under all three possible

readings, with a certain preference for t_1 (since the prospective nature of *to have left* in this context generally requires use of *by* and not *at*). The same point can be made for the relation between the 'situation time' and the temporal anchor, here the moment of speech S: it is his planning which is related to the moment of speech; the temporal relation of t_2 and of t_3, and hence of ts, to S is completely opaque.

Now, a sentence such as *He planned to have left by five* has a compound verb phrase. But very similar problems arise with simple verb phrases, as in *She closed the window*: There must be a time t_1 at which she does something (e.g., turns a handle or pushes a button), a time t_2 when the window is open, and a time t_3, when the window is closed; t_2 must precede t_3, and t_1 must overlap with t_2 (but can also stretch into t_3). In cases like this, that is, when the different time spans are packaged into a single word, it is much more difficult to 'access' them individually; but still, it is possible. Thus, a temporal expression such as *for five minutes* can indicate the duration of t_3, but not, for example, the duration of t_1, nor the duration of the minimal interval which includes t_1, t_2 and t_3, i.e., the full situation of her closing the window.

Summing up, the situation, as described by the sentence, regularly involves a whole array of temporal intervals: the 'time of situation' is not a single interval but a 'temporal structure', with a single interval as a borderline case. This applies whether the verb phrase is simple or complex. In this section we have shown that each of the three parameters S, R and E raises massive problems. In the next section, we will try to accommodate some of them.

2.2 The characterisation of temporal structures

In what follows, we shall adopt some ideas developed in Klein 1994 (and further refined in 2000a). They make use of three time spans as well, called TU, TT and TSIT (for 'time of utterance', 'topic time' and 'time of situation', respectively); hence, these ideas continue the tradition of three-parameter approaches. The crucial feature is a rigid distinction between 'temporal structures', that is, time spans and the temporal relations between them, on the one hand, and the characterisation of individual time spans which belong to such a temporal structure, on the other hand (see Klein 1994, chapter 3).

A temporal structure as such is, so to speak, 'empty'. Temporal intervals have no properties, except duration and being before, after, or within other temporal intervals. But they can be assigned properties. This is either done explicitly, by the lexical content of verbs or other expressions associated with them, or implicitly, i.e., by all sorts of contextual information or world knowledge. Consider again a sentence such as *He planned to have left by five.* It minimally involves the three times spans t_1, t_2 and t_3 together with the temporal relations described above. Each of the three time spans is 'lexically characterised' as the time of his planning, the time of his leaving and the time of his having left. Contextual knowledge may enrich this characterisation (for example the knowledge that 'his leaving' concerns 'Peter leaving Heidelberg' rather than 'Napoleon leaving Moscow').

The situation, as characterised by the lexical content of the sentence, regularly involves a more or less complex cluster of time spans. In other words, the 'time of situation' is normally not a single interval but a 'temporal structure' (with a single interval as a borderline case)[2]. When we speak in the following of 'the time of situation (TSIT)', this term is to be understood in this sense: normally, it is not a single interval (let alone a point), but a temporal structure which may involve a number of subintervals. This applies whether the verb phrase is simple or complex. But compound constructions regularly add temporal intervals which are accessible to further specification by adverbials and other devices. Consider, for example, an action such as digging a hole – it clearly involves various subphases: it has, for example, an initial phase and a final phase.

These, however, are not directly accessible to further specification: no adverb added to this verb can specifiy just its initial phase. But it is possible to decompose the very same action by a more fine-grained description and thus create further 'accessible time spans'. One such possibility is 'phasal verbs', such as *to start digging*, which expresses an initial phase. Thus, in *He started digging at five*, the expression *at five* does not specify the time of digging but the time at which it was initiated, its initial phase. Another important possibility is the nucleus or intermediate phase, *to be digging*, which gives an inner subinterval. Hence, *He was digging* relates an inner subinterval to the temporal anchor, thus giving the impression of somehow being within the action: the action of digging is seen as ongoing, as not completed. This naturally explains the notion of 'imperfective aspect', which is typical meaning associated with the English progressive.

TSIT is a temporal structure, which is more or less filled with content by the words of the sentence (and by contextual information). It is not 'embedded in time', i.e., it is not related to some other time span outside TSIT. Such an external time can be provided by the moment of speech. But as discussed in section 2.1, this notion is insufficient. We may now sum up the argument in three points:

1.   Temporal expressions are often linked to an interval which is not given in the text itself but must be taken from context. Let us call this interval the 'temporal anchor' (abbreviated as TA).

2.   Depending on a number of factors, TA can have different interpretations. One of these is 'the moment of speech'; but there are many other possibilities.

3.   The interpretation of TA may vary within a text, for example, from one utterance to the next; but it may also remain constant over some stretch of discourse. Hence, for a whole text, we typically have a whole array of TAs, which can be related to each other in various ways. A single TA is just a borderline case of such a 'TA-structure'.

It is a matter of terminology whether any type of temporal anchoring in this broad sense should be called 'deictic'. Since this term is traditionally used in connection with the here-and-now of the speech situation, such a broad understanding would be confusing, and therefore, we prefer to restrict the notion of 'deixis' to those cases in which the temporal anchor is indeed defined by the time at which the utterance is produced.

Note that the broad notion of temporal anchoring is not at variance with a deictic definition of tense (and temporal embedding in general). It is just a special case of a much broader phenomenon. For the sake of argument, let us assume in the following discussion that we indeed have only one utterance and a relevant 'deictic origin'. How are such a TA and a TSIT brought together?

Languages have various devices to this end. One of them is temporal adverbials. Another important device, found in many but not in all languages, is 'tense marking' on the finite verb. This is the point where the third temporal parameter, called here 'topic time' (abbreviated TT), comes in. A speaker, when formulating an utterance, must not only decide which descriptive content is to be included in the sentence; a decision must also be made on the time which he or she wants to talk about. These two decisions must be kept apart. Suppose the speaker, having witnessed an accident, decides to describe a scene in which a rider on a bicycle takes a left turn. In reality, this scene may have occupied a certain time span of, say, ten seconds; this is TSIT. For whatever reason, the speaker may choose to assert something not about these ten seconds as such but about a time which follows them, or about a time which precedes them, or a time which is properly included in them, or a time which includes them. English has grammaticised these

options. The speaker may say *The bike has turned left, The bike was going to turn left, The bike was turning left, The bike turned left*, respectively. As these examples show, English solves the task by creating a more sophisticated temporal structure TSIT', which also involves post times, pretimes, subintervals of the time which it takes the bike to turn left (the 'original TSIT'). For instance, *have left* differs from the bare *leave* by adding a 'post time'. The relation between the time talked about and the 'original TSIT' reflects various ASPECTS or ways of looking at the situation: thus, aspect is the grammaticalised relation between TT and TSIT. Other languages do not grammaticise the relation between TT and TSIT. In German, for example, there is no counterpart to the distinction between the progressive and simple form: German simple forms are neutral with respect to a 'perfective' and an 'imperfective' relation, just as the English copula *to be* in *It was green*. It is arguable whether the German Perfekt expresses a 'perfect aspect' in just the same way as the English perfect – an issue to which we shall return in section 5. It is the TT chosen by the speaker which is eventually related to the temporal anchor. In *The bike had turned left*, for example, TT precedes the temporal anchor, and in *The bike has turned left*, it includes the temporal anchor. In the construction *The bike is turning left*, an assertion is made about some time which includes the moment of speech (or some other temporal anchor), which is a subinterval of the bike's turning left. In *The bike was turning left*, the time talked about is also a subinterval of the bike's turning left, but it is said that this subinterval precedes the temporal anchor.

Summing up, we have three temporal parameters. These are

– the temporal anchor TA, which can, but need not be the moment of speech;
– the topic time TT, that is, a long or a short time span about which the speaker decides to speak;
– the time of situation TSIT; as a rule, TSIT is not a single interval but a complex temporal structure, i.e., a set of intervals which are temporally related to each other; each of these intervals is characterised by the lexical content of the sentence (and by contextual information; many, though not necessarily each, of these interval can be subject to further grammatical processes, such as adverbial modification).

All of this applies to individual utterances. But what about the entire text to which such an utterance belongs, say the full account of the accident? As for 'situation time', we must obviously distinguish between the time of the entire event and the times of the various subevents which the speaker has chosen to map into a single utterance. In other words, we have a 'big TSIT' and many 'small TSIT's; each of which is normally not a single interval but a 'temporal structure', in the sense explained above. Similarly, we must distinguish between the topic time of the entire text and the topic time of the individual sentences, and, finally, between the overall temporal anchor and the temporal anchors of individual sentences. Let us call the 'text time parameters' TA*, TT* and TSIT*, respectively. The crucial question is, of course, how these are related to the corresponding parameters of the individual utterances. If the entire text consists of just one sentence, they are simply identical. If the text is more complex, there are many possibilities. In the case of TSIT*, the individual TSITs are coherent substructures of TSIT*, chosen at the speaker's discretion. The case is more difficult for TT* and TA*. It may well be that the topic time is kept constant throughout the text, as in a picture description or in a mathematical argument, for example. The claim that 911 is a prime number is not just confined to a small time span, it is claimed to hold forever, and hence constant for all such arguments. But the speaker may also choose to change topic times while constructing the text. The same is true for the temporal anchor: there may be one temporal anchor which remains constant throughout the text; normally, however, there is a steady shift. This is clearly the case when 'the moment of speech' is chosen as a temporal anchor, thus giving rise to what in research on narratives is sometimes

called 'the order of mention'.

It is these choices which lead to different temporal perspectives for one and the same complex verbal task, as defined by the initial quaestio. We may now return to our first question: taking English and German speakers, are there systematic differences in the choices made in the same communicative situation? Before turning to the empirical data, we will take a look at some of the possibilities.

3. Variation in perspective

The complex verbal task which we investigated was the retelling of a short silent film, called *Quest* (see, in detail, Carroll and von Stutterheim in press). In this setting, TSIT* is the highly complex temporal structure of the film from which the speaker has to select substructures that will form the individual TSITs of each utterance.[3] There is also a well-defined 'text speech time' – the temporal interval during which the speaker produces the text. But there is no obvious way in which this speech time – the potential deictic root – could be related to TSIT*, since the situation, the plot of the film, is not in the real world. The events, processes and states shown in the film are neither in the past nor in the present nor in the future, as defined by the time of speech. Now, under the analysis of tense suggested above, there is no direct relation between TSIT and TA, anyway. It is always mediated by TT* (and the individual TTs, respectively). But the problem remains: no matter how large the speaker sets the entire TT* and how it is defined for the individual utterances – the times talked about cannot be related directly to the time of speech, on the one hand, and the temporal structure defined by TSIT*, on the other. There are various possibilities in which the speaker can solve this particular problem, as well as the general task of retelling the film; we will briefly sketch three of them.

First, the speaker may treat TSIT* as well as the individual situations *as if they had really happened*. This allows a 'deictic' treatment in the usual sense. The actual production time is considered to be the temporal anchor TA*; this time is split into subsequent smaller 'times of utterances', which serve as the temporal anchors for the various topic times. These precede their temporal anchor, hence the tense is the past. Depending on how the individual topic times are chosen in relation to the individual situations described by a sentence, various aspects are possible. In the typical case, all main structure utterances, – those utterances which describe what happened rather than giving background information – are in the simple past. Another factor which may play a role here is the 'length' of some topic time: if the speaker considers it to be very short, the 'time window' on the alleged reality is very small, and only a part of the situation is visible. In this case, the speaker might be forced to 'zoom in' on a special phase of the situation, for example by saying *He started to dig a hole*, instead of *He dug a hole*.

Second, the speaker may indeed treat the film events as fictitious. Then, a different type of temporal anchor TA* must be chosen. One way of doing this is to take 'the time of experiencing the film' as the temporal anchor. Note that this time, in turn, is related to the moment of speech. But this need not have any relevance for the way in which the text is embedded in time: the entire story, as well as each of its sentences, is related *to the time of experiencing the film,* reactivated, so to speak, by the speaker. This time must be split up for each utterance and results in a form of presentation in which each utterance is prefixed by 'and now I/we see'. The content unfolds again in front of the mind's eye. Thus, the appropriate tense is the present. The anchor is something one might call the 'experiencer's now'. As with the first strategy, aspect depends on 'window size', i.e., on how large the TTs are chosen: if the full event is visible, then the simple present is chosen. Otherwise, the speaker must either use the present progressive or some other, more complex form, such as *He stops digging*. This strategy

8

resembles the 'historical present' in personal narratives, except that in this case the option of a deictic anchor is readily available; in reporting on fictitious events (film world), a different type of anchoring time is chosen.

A third strategy treats the events of the film as fictitious entities and maintains the time of speech, the deictic anchor, as temporal anchor. The entire text is, so to speak, prefixed by 'these are the things I have seen in the film'. In other words, there is a very big TT*, in which the fictitious events are visible like a string of pearls. Each of these can then be described as a self-contained unit within this big topic time. An essential element of this way of evoking the material is the temporal order in which these entities stand in relation to each other. The tense form is the present, since the temporal anchor is included in TT*; the aspect should normally be perfective, and there should be devices which indicate the temporal relation of the subevents, such as *then, next*, etc. Note that this strategy resembles the traditional notion of 'praesens tabulare', in which facts are listed, except that in this classical *praesens tabulare*, the events are really in the past. But this fact does not matter for the way in which they are embedded in time.

These are three ways of presenting the temporal structure of the film - three perspectives under which it can be retold. There may be others, but we will not speculate on them here. Let us now turn to the question what English and German speakers indeed do when they try to solve the task.

4. Temporal perspectivation in English and German texts: a first experiment (Quest 1)

The silent film *Quest* (duration 7 minutes) shows the adventures of a clay figure who wakes up in a desert-like world and embarks on a search for water. His 'quest' takes him through four different worlds (a desert world, a paper world, a stone world, and finally a world dominated by robots and other industrial machines.

The English and German were for the most part university students in the mid twenties. the data collection proceeded as follows: They first viewed the entire film without interruption and then viewed it again, but this time the film was stopped at the end of each "world". Each speaker was asked to tell what happened in the episode or 'world' they had just seen. The texts obtained vary in length in all groups. The average length for the English group is 173 utterance units, for the German corpus it is 175 utterance units. In what follows, we will summarise the findings for this set of data in as far as they are relevant for the question of temporal perspectivation (for a detailed description see Carroll and von Stutterheim in press).

4.1 English film re-tellings

English speakers adopt a viewpoint-related strategy but this viewpoint is not the moment of speech. The temporal anchor is the *now* of the experiencer, that is, as the film is experienced in unfolding in front of the mind's eye. In other words, T* is the 'film experience', rather than the time at which the film is retold, and the temporal anchors of individual utterances are subintervals of this TA*. This corresponds to the 'and now I see'-strategy mentioned above. The result resembles a series of 'screen shots', each with the 'experiencer's now' and a short time window related to the this 'now', the topic time of a given utterance. Note that the 'experiencer's now' and the TT of two subsequent utterances need not differ: it may well be that two or even more subsequent utterances are used to describe the same 'screen shot'; metaphorically speaking, several situations may be visible through the same window. We have called this a *reportive perspective.*

One characteristic feature of a reportive perspective is that circumstances are viewed as leading in some way to sets of events, as illustrated in the following example. The opposition in English between the simple present (sits) and the progressive (is sitting) is crucial in this context.

Ex. 1: Circumstances: use of progressive (*-ing)* form
(1)   *# he starts trying to scoop the water up*
(2)   *but the water is being absorbed into the surface #*
(3)   *and he cannot get any water from that*

        switch to simple tense

(4)   *so he sits there*
(5)   *and looks up*
(6)   *to where he thinks they are*
(7)   *but there's no more drips*
(6)   *they have stopped dripping*
        (termination of events linked circumstantially*)*

In Ex. 1 the *now* of the experiencer, perceiving the events as they unfold before the mind's eye, provides the topic time. The time of situation of the utterances (1-3) is linked to this TT, and is either simultaneous with TT, as in 1, or it includes TT, as in 1 and 3. This means that TT (as well as the 'experiencer's now') is maintained across the first three utterances: they show, so to speak, the same screen shot. The temporal interval introduced by the situation *water being absorbed into the surface* functions as TT-interval for the following events. The TSIT of (4-7) are situated within this frame time-interval, and nothing further is stated about their temporal location or the way in which they are linked. The present perfect form in (8) requires a point in time as TT as the point at which an event is terminated and the post state of the event holds as the point of reference. This point in time can only be given in relation to the 'experiencer's now'. The pattern of topic time management for this piece of text is as follows:

(1)   TT (point in time) = experiencer's *now;*   TT overlaps TSIT
(2)   TT maintained   TT δ TSIT
(3)   TT maintained   TT δ TSIT
(4)   TT (interval) = TSIT   TSIT δ TT
(5)   TT maintained   TSIT δ TT
(6)   TT maintained   TSIT δ TT
(7)   TT maintained   TSIT δ TT
(8)   TT (point in time) = experiencer *now*   TT after TSIT

Ex. 2 gives another example which exhibits the same pattern. The two utterances a and b are linked to the deictic centre as their TT. They establish a temporal interval by their TSIT, which functions again as TT interval for the events reported in the following utterances.

Ex. 2:
Circumstances: progressive forms
(1)   *# there's no water to be found*
(2)   *and he's walking through this area #*

Embedded actions which happen while walking: simple forms

(3)   *and <u>he hears</u> the drip again*
(4)   *and <u>sees</u> this moist area on this piece of paper*
(5)   *so he <u>gets excited</u>*
(6)   *and he <u>runs</u> over*

Events such as *hears, sees*, and possibly *gets excited* occur within the established time span. The termination of the time span *walking through x* is given by the event *and he runs over*.
This yields the following general pattern of TT management:

$$TA \quad = TT \, \gamma \, TSIT_x$$
$$TSIT \ = TT \, \delta \, TSIT_{x+n}$$

In sum, the pattern followed by the English speakers is temporally complex. The initial TT provides the anchor for the introduction of a circumstantial frame, characterised by an imperfective view on the activities or processes reported. The frame includes a set of events, and these are accorded some implicit portion of time within the given span. The events described may overlap or follow each other in time.

Frames of this kind are marked by a general absence of temporal adverbs such as *then*. The temporal relations between the various TSITs remain implicit. The link between the frame and its associated events is often of a causal nature (there is circumstance x : so this leads to y). Causal links can be more or less prominent, however, as shown in the two examples above.

As to event construal, no specific temporal property is required. All types of situation can be hooked up to the temporal anchor, there is no constraint with respect to the composition of the event. The speakers adjust the specific type of situation in its relation to the temporal anchor by means of phasal decomposition. We observe a high level of aspectual marking (inchoative with *to start* and ongoingness) which serves the purpose of expressing the flow of the events in relation to the viewing point. Occasionally, speakers select another option in text construal, similar to that found in German, but these are not the preferred principles typical of the majority of film re-tellings (see, in detail, Carroll and von Stutterheim in press). The existence of these options, however, demonstrates that the strategy described above is not forced upon speakers. The structure of the language allows for different options (exploit the principles of construal allowed by grammaticised aspect, or not), but the data provide evidence of what can be described as a preferred form of temporal organisation which correlate with grammaticised meanings (ongoingness) and the type of referential frame which they allow.

4.2 German film-retellings

In contrast to speakers of English, the German speakers prefer to view a dynamic situation as a well-delimited entity which can fill a slot on a structured time line. The pattern exhibited in the re-tellings follows the strategy of linking the current topic time TT to the preceding time of situation. This goes with the explicit marking of a relation between different temporal intervals, usually of the type "posteriority". Their strategy evokes the notion of a big window in which the individual events are presented like a well-ordered string (cf. section 3). The following example illustrates this type of temporal organisation which is followed by all German speakers in the sample.

Ex. 3:
(1)   *und wacht **dann** so langsam auf*
      'and then he slowly wakes up'
(2)   *und schaut sich um*
      'and looks around'
(3)   *und sieht **dann** die flasche vor sich liegen*
      'and sees then the bottle lying in front of him'
(4)   *nimmt die flasche*
      'takes the bottle'
(5)   *und guckt*
      'and looks'
(6)   *ob da wasser drin ist*
      'if there is water in it'
(7)   ***dann** steht er so langsam auf*
      'then he gets slowly up'
(8)   *und kniet so*
      'and kneels down'
(9)   *und guckt sich um*
      'and looks around'
(10) *und **dann** donnert es plötzlich*
      'and then there is suddenly thunder'
(11) *und er hört*
      'and he hears'
(12) *wie wassertropfen in den sand fallen*
      'how water drops fall in the sand'

The organisation principle in the German re-tellings is characterised by two main features:

(a) Explicit marking of temporal relations by temporal adverbials in the context of the sequence of events. Adverbials are used to establish TT and to link them explicitly to the preceding TSIT.

(b) A holistic view on the events, including points of completion or results of an event. This provides the anchor point for the shift-relation of the TT interval.

4.3 Summing up

The following two patterns were observed in organising temporal coherence on the basis of the unfolding of topic time:

$\text{TU} = \text{TT } \gamma \text{ TSIT}_x$ $\text{TSIT}_x \geq \text{TT } \delta \text{ TSIT}_{x+n}$ (main structure)
$\text{TSIT} = \text{TT } \delta \text{ TSIT}_{x+n}$ $\text{TSIT}_y \delta \text{ TT}$ (side structure)

(dominant pattern in the English texts)   (exclusive pattern in the German texts)

Given their systematicity the contrasts observed require an explanation. We assume (see Carroll and von Stutterheim in press) that the basis for these different patterns of referential framing can be found in the structure of the languages involved, more specifically, in the different forms of aspectual marking, as indicated above. Note, however, that the differences cannot be reduced to

the fact that English has a systematic aspectual opposition between an imperfective and some unmarked form. Even where German has traces of aspect, the opposition encoded differs. Taking the linguistic means at the level of verbal morphology, the marked form in German is the perfective. These differences have dramatic consequences for the relational potential of sentences.

In order to explore this further, and also in order to exclude the possibility that the differences observed are only a stylistic convention in renarrating films (more reportive in English, see Smith 2000, more narrative in German), we started a second study and checked what happens if the task is changed.

5. Perspective shift and its structural roots

The design of the second study (called *Quest 7* – the seventh experiment with this film) was essentially the same, except that the speakers were now asked to renarrate the film in the past. Our expectations in selecting this design run as follows: the English speakers – bound to their 'and now I see-anchor' – should encounter problems in transferring the content into a frame of reference located in the past. For them, a high degree of transformation is required with respect to the selection of event components, on the one hand, and with respect to the type of relation required between utterances. Broadly speaking, they have to switch from a viewpoint anchored strategy to an anaphorically-based strategy. The latter implies the provision of boundaries for the event times by either verbalising the endpoints of situations or by closing the event time intervals by explicit temporal information (adverbials), in order to establish an intrinsic temporal relation which is independent of viewpoint.

German speakers have an easier task to solve. Their 'here are the events, in this order'-strategy is not dependent on an external viewing point. Therefore, it should not be particularly difficult for them to switch to a past referential frame. Since the general type of linkage between the situations reported is that of 'shift-in-time' with states related to the event times, nothing with respect to the construal of events nor to the type of relation would have to be changed. We therefore expect English speakers to run into difficulties, and the German speakers to have few, if any, in solving the task.

The analyses of the data proceed in the following steps. First the English texts *Quest 1* and *Quest 7* will be compared with respect to the referential frame selected and the patterns of event construal. Use of tense and aspectual forms will be analysed quantitatively (frequency of occurrence and tense switch) as well as qualitatively in relation to situations which trigger their use. The same analysis will then be carried out for the two German data sets. In a third step, English and German *Quest 7* texts will be compared with respect to patterns of temporal organisation. In conclusion we will relate structural properties of the languages involved to principles of information organisation guiding the production of complex texts.

5.1 Comparing English Quest 1 and Quest 7

The basic strategy selected with *Quest 1* cannot be maintained by speakers under the second condition given with *Quest 7,* where speakers were asked to take a 'past' perspective on the situation, as mentioned above. One solution is to switch to the framework of a past time narrative, but this is typically found in English in narratives of personal experience (cf. the classical study of Labov 1972)[4]. The other possibility is to produce some kind of hybrid between a viewpoint anchored strategy ('the experiencer's now') and the requirements of a temporal structure which is anaphorically linked. In this case we expect texts to be inconsistent with respect to the frame of reference and with respect to the formal devices used.

The *Quest 1* texts – with two exceptions – are anchored with respect to the *now of the experiencer*. The present tense is used for both aspectually unmarked and marked forms. The second set (*Quest 7*) is characterised by the dominant use of the past tense form, showing, however, a considerable number of tense switches, corrections and inconsistencies in topic time (TT) management. The following quantitative and qualitative analyses substantiate these observations. We take the relative frequency of the tense forms used (table 1), switches into present tense and corrections from present to past within one utterance (table 2), contexts which trigger switches (table 3). The data analysed covers the first three 'worlds' of the film, which amounts to 2/3 of the material.

present		past		total number narrative utterances
simple	progressive	simple	progressive	
165	35	1187	204	**1591**
10.4%	2.2%	74.6%	12.8%	

Table 1: English Quest 7: Relative frequency of tense forms

present tense forms	past tense forms
200 / 1591	1391 / 1591
12.6%	87.4%

Table 2: English Quest 7: Switch into present tense

The figures show that the perspective selected in constructing the texts is relatively unstable. In 12,6% of the narrative utterances speakers switch to a present tense form. This type of tense shift forms the majority of tense shifts in the texts.

In order to see if these switches are motivated by the strategy of the 'true' *historical present,* or whether they result from an interference with the *reportive perspective,* the contexts which trigger a switch in tense were analysed. If the tense switches are based on clash in perspective, then switches to the present tense should occur in contexts which attract a viewpoint anchoring. Given the possible pattern of TT management outlined above, a state or an ongoing action requires an externally established temporal reference point (*now*). These are informational units (a) which establish the circumstantial frame at the beginning of the scenes and (b) which present events as ongoing within sequences involving phasal segmentation (*he started pushing, he is/was pushing*). As presented in table 3, there are five types of contexts in which switches into the present systematically occur; clearly indicating that the contexts which trigger tense switches are not random.

narrator viewpoint	phasal segmentation	states	situation reoccurs	activity protagonist/environmental force
15,8%	15,8%	28,5%	4,7%	34,9%

Table 3: English Quest 7: Contexts for switches and corrections

Switches where the temporal perspective selected requires an externally anchored TT (*narrator viewpoint, phasal segmentation, states*) add up to 64.8%. To use a term introduced by Carlota

Smith, these are situations which need 'capture' (Smith 2000). Let us briefly illustrate the five cases.

(a)   Narrator viewpoint, i.e., explicit references to the speakers activity of watching and interpreting the film (e.g. predicates of seeing, hearing, interpreting), as well as comments and qualifications on the events: *suddenly these boulders began to move and <u>we see</u> it is actually a human like figure;* or, *the clay man <u>seems</u> to wake up.*

(b)   Phasal segmentation, e.g.: *the figure <u>started falling</u> and <u>he's pushing</u> sand around*

(c)   State type predicates relating to circumstances, e.g.: *and it is windy*

(d)   Reoccurrence of a situation, e.g.: *and again the creature hears water dripping*

The fifth category 'activity' is more problematic. It covers cases with no obvious contextual motivation. These switches to the present may be linked to inattention on the part of the speaker to the task requirements, since they tend to occur at the beginning of a new scene where speakers are more likely to fall back into what may be called the preferred or 'natural perspective' (reportive) when recounting the content of a film. These switches in referential continuity show that speakers do not establish an intrinsically organised time-event-structure. There is no TT link between event times and states related to these events. Local planning leads the speakers to go the 'natural' way, that is, relating those predicates which 'need capture' to the viewpoint *now*. The tense switches suggest that speakers have superimposed the *past strategy* onto an underlying *reportive perspective* which is still conceptually dominant.

Further evidence for the maintenance of at least basic aspects of a reportive perspective comes from parallels in the processes of information selection: the level of granularity. When we look at the level of granularity we have to consider the degree of resolution or level of segmentation of the situations shown in the film into subevents and event phases. Looking at the data (see the examples below), we find that the level of granularity chosen by the Quest 7 speakers corresponds to a large degree to that of the reportive perspective. This is supported by the average number of utterances produced per text: 173 in *Quest 1* and 182 in *Quest 7* (20 speakers in each group) which do not differ significantly. Events reported in the utterances are frequently segmented into phases, and the inchoative and imperfective aspect is selected for events in the story line. This high degree of resolution is very similar to the retellings in the present.

A further indicator of the maintenance of a basic reportive perspective is the comparison with the referential frames found in the narratives of Quest 1 speakers (two in all), who spontaneously opted to re-tell the events in the past tense, and not in the present. In contrast to the narratives of speakers who were asked to adopt a past tense perspective (Quest 7), there are markedly fewer events which relate to an external viewpoint (the *now* of the experiencer); in other words there are fewer events which are segmented into phases, and ongoing events (*water was dripping*) are not integrated into the flow of narrative events, i.e anaphorically related events, but are marked as secondary by means of devices such as subordination, attribution, etc. (cf. Ex.12. below).

In comparing *Quest 1* and *Quest 7* texts qualitatively in the next section we will begin with evidence for the maintenance of some basic features of a reportive perspective in the English Quest 7 narratives. The indicators found will be illustrated by examples.

1.   The first indicator concerns processes of information selection: Example 6 from the *Quest*

1 corpus gives an extended example of the *reportive perspective* as found in narratives told in the present tense. Situations with a high degree of resolution or event segmentation were selected in order to illustrate and compare the relevant principles. The temporal aspects of the events relate some subinterval of TSIT to TT and therefore do not need reference to an endpoint or boundary of TSIT. The event time intervals and their intrinsic relation remains implicit:

Ex. 6
Quest 1 (speaker 21)
(1)	*so he starts digging in the sand*	(ongoing)
(2)	*towards where he thinks he hears the water*	
(3)	*and he digs*	
(4)	*and digs*	(ongoing)
(5)	*and and he starts sliding into*	(ongoing)
(6)	*what looks like dry quicksand*	
(7)	*and he starts falling into the hole of the sand*	(ongoing)
(8)	*okay he's falling*	(ongoing)

There is a marked contrast to *Quest 7* texts in that the narratives in the past tense, with a similar degree of resolution, consistenly present ongoing events as reaching a point of closure. This can be achieved at different levels of the macrostructure and by different means. The comparison between the Quest 1 and Quest 7 narrations shows that in the latter case speakers attempt to express closure not just for the coda or final phase (*ended up falling*) but very often for events which form the onset and nucleus (*began digging but after a very short time ...; kept digging and eventually ...*; etc). This difference is illustrated by the following set of examples.

In example 7 speakers unfold a scene which corresponds in perspective to that in Quest 1, presenting the events as ongoing, unbounded. At the end of the scene, however, they close the sequence by explicitly binding the last event either as a temporal endpoint (Ex. 7, 7) or a causal result to the preceding discourse (Ex. 8, 3,4). Note that events which form part of the story line are reported on in an subordinate clause in this function (*until he sunk*), a packaging phenomenon which is due to the hybrid character of the perspective selected and which is not found in self-selected past time narratives.

Ex: 7
Quest 7 (speaker 18)
(1)	*he looked up at the sky*	
(2)	*to see if there were rain clouds*	
(3)	*but didn't see any*	
(4)	*so then he dug in the sand*	
(5)	*he kept digging*	
(6)	*and kept digging*	
(7)	**until he sunk**	(Closure*)*

Ex.: 8
Quest 7 (speaker 07)
(1)	*he started to dig around*	
(2)	*and like a cat kind of eh throwing up the sand behind him*	
(3)	*and he dug **so hard***	(qualification of activity – cause)
(4)	***that he fell through into a different kind of equally desolate world***	(Closure)

2.	Another trace of the pattern of adaptation can be observed at the local level, where utterances which are aspectually presented as ongoing are closed by aspectual termination (*to end up*) or temporal quantification, either in the utterance itself or by the following utterance.

Ex.: 9
Quest 7 (speaker 11)
(1)	*and he started to dig*
(2)	*but he didn't find anything*
(3)	*and he kept digging*
(4)	*and **eventually** found out*
(5)	*that it was too late*
(6)	*there was quick sand*
(7)	*and he fell through*
(8)	*and went somewhere eh under the desert*

Ex.:10
Quest7 (speaker 02)
(1)	*and / then he heard the sound of water*
(2)	*dripping somewhere around him*
(3)	*so he got up*
(4)	*and went in search of the source of this sound*
(5)	*and he began digging in the sand*
(6)	*and **after a very short time** of digging*
(7)	*he fell through some sort of hole in the sand*
(8)	*and **moments later** found himself on another level of consciousness / or another plain / or another planet*
(9)	*and as he was lying on the ground*
(10)	*he again heard the sound of water dripping*
(11)	***then** he found the spot*
(12)	*where in fact water was dripping*
(13)	*and found a small puddle on the paper*
(14)	*and began trying*
(15)	*to scoop it up*
(16)	*and as he did so*
(17)	*he created a hole in this piece of paper*
(18)	*and again fell through the hole*
(19)	*and went walking around this planet*
(20)	*and **ended up** on top of a rather large pile of rocks*
(21)	*so he began to climb down from the rocks*
(22)	*and it was quite a high pile*
(23)	*so he had difficulty*
(24)	*getting down*
(25)	*and **ended up** falling the last bit of the way*

3.	The problem of providing an anchorpoint for a shifted Topic time can also be solved by adding an explicit temporal shifter such as *then* or *a moment later* or by introducing some indefinite anchor point through a time adverb such as *suddenly*. The number of adverbials used in this function increases in *Quest 7*, measured with respect to the total number of narratives

utterances.

Quest 1	Quest 7
89 / 1577	149/1591
5.6%	9.4%

Table 4: English Quest 1 and Quest 7: Temporal Adverbs

Let us now summarise the findings for the English *Quest 7* texts, where speakers are forced into a perspective which they normally would not take. They try to comply with the instruction and use past tense forms. They realise that coherence cannot be achieved by using the viewpoint *now* as constant anchor for TT. This leads to attempts to establish anaphoric links. The perspective under which events are construed for verbalisation and weighted relative to one another within the hierarchical structure of the text is not fully shifted to that of a narrative anchored in the past in anaphoric terms. Information organisation follows a hybrid pattern in which features of a frame anchored in the 'experiencer's now' are marginally reconciled with the requirements of a past time perspective which has no place for this anchor.

In order to give an indication of the contrast with past time narratives which follow an anaphoric temporal perspective, we will close this section by looking at an example of this type. There are not enough examples of self-selected past time narratives in the corpus to carry out a representative analysis (two out of a corpus of 37 *Quest 1* texts show this perspective, as mentioned earlier), so its function can only be illustrative.

Ex. 12
Quest 1 (speaker 30, past time)
(1)   *eh this sort of sand man woke up*
(2)   *ehm half buried in sand in the middle of the desert*
(3)   *and he wanted some water*
(4)   *so he checked his bottle*
(5)   *there was none there*
(6)   *and he started digging ehm / digging around*
(7)   *where the bottle was lying*
(8)   *to see if maybe he could gather some*
(9)   *and instead he dug a hole*
(10)   *and fell through*
(11)   *ehm our sand man friend fell through into a sort of ehm paper world*
(12)   *where the ground seems to be made of pieces of paper*
(13)   *ehm he had heard some water*
(14)   *dripping*
(15)   *and went to find it*
(16)   *he found a little puddle*

What we can see in this excerpt is that events are bounded and where states have to be integrated this is carried out by means of causal relations (4/8). In general, the degree of phasal segmentation is lower compared to the other texts. But speakers use the means available in the language to express temporal contours. Significant in this respect is the fact that the ratio between situations reported under the aspect of ongoingness and situations which are viewed holistically is inversed, compared to the 'present tense' texts. The contexts triggering subordination (ongoing

events-*water dripping*, states) provide a further indicator of this difference (see also Carroll and Lambert in press).

Although the data base is very limited, there are narratives in which speakers talk about past events using a 'past time' perspective. Hence, English speakers can indeed select a perspective which essentially corresponds to the only option German speakers have: the presentation of a sequence of events with an intrinsic temporal link between the single units, whereby phasal segmentation is integrated accordingly.

4.2. Comparing German Quest 1 and Quest 7

Let us look first at tense switches in the *Quest 7* data. All 20 speakers select past tense forms as their dominant tense.

present	past	total number of narrative utterances
94	909	1003
9,4%	90,6%	

Table 5: German Quest 7: Relative frequency of tense forms

Switches to the present are less frequent in the German texts. However, there is a significant degree of variation, however, with respect to the two tense options selected in re-telling the film - simple past (Präteritum 16 speakers) and Perfekt (Perfekt 4 speakers).

Looking at the distribution of the switches to the present tense across subjects, we see that most of the occurrences are produced by 4 speakers, and the dominant tense for those speakers is the Perfekt. But we cannot speak of unstable patterns of TT management in this case, unlike in the English texts. Rather two options which belong to a 'past tense' referential frame (anaphoric) are realised in the German texts. One is a past time anchored pattern with the Präteritum as the main tense in which speakers follow what the underlying quaestio suggests: "Tell me, what happened in the film?". Example 13 and 14 illustrate this pattern.

Ex.: 13
Quest 7 (speaker 01)
(1)   *also ein kleiner mensch aus knete erwachte in der wüste*
         'okay a small man made out of plasticene woke up in the desert'
(2)   *und sah neben sich eine wassserflasche*
         'and saw beside him a bottle of water'
(3)   *er hob sie auf*
         'he picked it up'
(4)   *schüttelte*

         'shook it'
(5)   *und nichts kam raus*
         'and nothing came out'
(6)   *er setzte sich auf*

         'he sat up'
(7)   *und äh schaute um sich*
         'and looked around'
(8)   *und sah nur wüste*

'and saw only desert'

(9) *und fing an zu graben*
     'and began to dig'

(10) *und dabei wurde das loch unter ihm so groß*
     'and thereby the hole under him got so big'

(11) *daß er selbst hineinfiel*
     'that he himself fell in'

Ex.: 14
Quest 7 (speaker 12)

(1) *die szene spielte in einer wüste*
     'the scene takes part in the desert'

(2) *das sandmännchen lag da bewegungslos*
     'the sandman lay there without moving'

(3) *dann wachte es auf*
     'then it woke up'

(4) *und hatte eine flasche neben sich*
     'and had a bottle beside him'

(5) *und ähm suchte wasser*
     'and searched for water'

(6) *das sandmännchen lag ähm ja im sand*
     'the sandman lay in the sand'

(7) *und hat nach wasser gegraben*
     'and (has) dug for water'   (Perfekt)

(8) *und hat immer / immer tiefer gegraben*
     'and (has) dug always deeper'   (Perfekt)

(9) *hat sich ein loch gegraben*
     'and (has) succeeded in digging a hole'   (Perfekt)

(10) *bis ähm / bis das sandmännchen selbst in das loch einsank*
     'until the sandman himself into the hole sank'

(11) *und verschwunden ist*
     'and (has) disappeared'   (Perfekt)

(12) *das sandmännchen ähm befand sich nun in einer papierwüste*
     'the sandman found himself in a desert'

(13) *es lag wieder da*
     'it lay there'

(14) *und wachte langsam auf*
     'and woke up slowly'

(15) *und stand auf*
     'and stood up'

(16) *es herrschte ein ziemlich starker wind in der papierwüste*
     'there were strong winds in the paper desert'

(17) *und es flogen ihm ähm sehr viele blätter wind an den körper*
     'and to him flew many sheets on his body'

(18) *und das sandmännchen ging immer weiter durch die wüste*
     'and the sandman continued on through the desert'

(19) *bis es irgendwo ähm eine pfütze fand*
     'until it found a puddle somewhere'

(20) *und ähm da hat sich das sandmännchen umgeschaut*
     'and there the sandman (has) looked around' (Perfekt)

(21) *woher das wasser kam*
     'from where the water came'

(22) *am anfang hat es gedacht*

'at the beginning he (has) thought' (Perfekt)

(23) *es käme von oben*
 'it might come from above'

(24) *und dann / ja es / es g / und hat nach oben geschaut*
 'and (has) looked up' (Perfekt)

(25) *und dann ähm hat er nach der pfütze gegraben*
 'and then he (has) dug for the pool of water' (Perfekt)

(26) *und er hat weiter gegraben*
 'and he dug on' (Perfekt)

(27) *bis er wieder äh runterfiel in ein loch unter der papier äh pfütze*
 'until he again fell down in a hole under the paper eh pool'

Example 14 illustrates the systematicity of tense switches between Präteritum and Perfekt. Before discussing the motivation for these tense switches, we will give an example for the other basic option selected by the German speakers. Some use the Perfekt as the narrative tense with occasional switches to the present.

Ex.: 15
Quest 7 (speaker 18)

(1) *das sandmännchen ist wieder in eine welt unter der zweiten welt gefallen*
 'the sandman has again fallen under the second world' (Perfekt compatible with present perfect)

(2) *ähm die diesmal nicht aus papier besteht*
 'which this time is not made of paper'

(3) *sondern aus / also / quasi eine art steinwüste ist*
 'but is almost a stone desert'

(4) *und ähm muß aufpassen*
 'and he has to be careful'

(5) *dass nicht steine von oben runterfallen*
 'that stones do not fall down from above'

(6) *oder mußte aufpassen entschuldigung*
 'or he had to be careful, sorry'

(7) *und ähm ja manchmal sind auch steine von unten nach oben gedrückt worden*
 'and sometimes stones were pushed up from underneath'

(8) *und beim zweiten mal hat dann auch so ein steinberg ihn erwischt*
 'and the second time he was caught up by a column of stone'

(9) *der von unten aus dem boden geschossen ist*
 'which had shot up out of the ground'

Ex.: 16
Quest 7 (speaker 17)

(1) *und dann blickt sich das sandmännchen*
 'and then the sand man looks'

(2) *hat sich suchend umgeblickt*
 'has looked around searchingly' (i.e. has been looking around..)

(3) *nichts gefunden*
 'found nothing'

(4) *oder / und dann hat es angefangen im sand zu graben*
 'and then it (has) begun to dig in the sand' (Perfekt)

(5) *wahrscheinlich auf der suche nach wasser*
 'probably searching for water'

(6) *und hat angefangen zu graben*
 'and has begun to dig' (Perfekt)

(7) *und zu graben*
  'and to dig'
(8) *und immer tiefer*
  'and continually deeper'
(9) *und irgendwann ist es mit hineingerutscht*
  'and at some point he (has) slipped down'   (Perfekt)
(10) *und die flasche mit ihm*
  'and the bottle with him

Most of the German texts are consistent in their use of tense forms in that those which exhibit tense switches remain within a constant frame of reference. This contrasts clearly with the English data. We will briefly look at the two options selected by the German speakers and then come to an interpretation of the Perfekt in its aspectual properties.

(a) tense switch Präteritum – Perfekt (cf. example 14)

The two forms clearly differ in their contribution to the temporal interpretation of a discourse.[5] In the Präteritum, TT overlaps and, as a default, includes TSIT; in the Perfekt, TT follows TSIT. This is illustrated in the following excerpt:

er hat sich ein loch gegraben   TT after $TSIT_1$
'he (has) dug a hole'   (Perfekt)
bis er einsank   TT maintained and includes $TSIT_2$)
'until he sunk in'

Intuitively the use of the Perfekt underlines the eventhood of a situation, since boundedness is one crucial, though not necessary property of dynamic situations.
        In texts where the Präteritum is the dominant tense, TT management is based on an anaphoric strategy. There is no shift to an 'experiencer-now-frame'. Speakers remain in the past, but they change perspective in a significant way. The shift to the Perfekt occurs either together with specific predicate types or at particular points in the narrative macrostructure. Predicate types which trigger the shift are 2-state predicates. They move the story line and refer to the events which form the culmination of the respective scene.

(b) tense switch Perfekt – present

As is well-known there are fundmental differences between the German Perfekt and the English English present perfect. Very often, it has to be rendered by the English simple past, as shown in the glosses for examples 13/14 above, where the temporal frame is located somewhere in the past or rather independently of any temporal anchor (the fairy tale type of indefinite temporal reference). The relationship between predicates in the Perfect is therefore not mediated through *now* – as the case in English. Rather the propositions are anaphorically related according to the strategy of a TT shift described above. Frequent use of *dann* (then) as a temporal connector underlines this relationship. We explain the shift to the present tense by the possibility of linking the TT for a Perfekt form to the time of utterance TU, although this is not a necessary concept for temporal linkage. The compatibility between the Perfekt and a TT = TA seems to provide the basis for a switch into the preferred perspective of a TA anchored TT, as shown in the choice of present tense in *Quest 1*.

What we rarely get in German is a tense shift from the Präteritum (simple past) to the present tense. Texts are more consistent compared to the English data. Looking at the other features analysed for the English texts we do not find evidence for conflicting perspectives either. The number and type of temporal adverbs does not differ, the pattern of event construal remains the same and information packaging by subordination reflects the same event-state-hierarchy (Carroll and Lambert in press, Carroll and von Stutterheim in press, von Stutterheim 1997).

4.3 Comparing German and English for perspective shift (Quest 7)

The comparison within the two language groups revealed a contrast in that the English *Quest 7* texts differed substantially from Quest1; in the German data, there is much less of a contrast. In the following we will give a brief overview of the findings for the two languages and then come to the discussion of the results.

4.3.1 Survey of the results

The two groups considerably differ in the level of granularity for Quest 7, as shown in table 6.

	Quest 1	**Quest 7**
English	173.3	182.2
German	174.5	112.4

Table 6: length of narrative (average 20 speakers)

The English texts maintain the same level of resolution under both conditions, which includes basically the same proportion of subevent segmentations. This is taken as an indicator that the English speakers, although using the past tense, do not actually shift the underlying reportive perspective, as discussed earlier. Unlike in English, the German *Quest 7* texts are shorter than the *Quest 1* texts. Since there are no other significant indicators which point to a change in perspective, and there are no indicators in the domain of temporal structure, the reason why the texts are shorter has to be sought elsewhere. The next table compares the use of adverbials such as (*then, a moment later, suddenly*) which mark a temporal shift, showing that although there an overall increase in their use in Quest 7, the rate of occurrence is still lower in English.

	Quest 1	**Quest 7**
English	89/1577	149/1591
	5.6%	9.4%
German	247/1943	228 / 1003
	12.7%	22.7%

Table 7: frequency of adverbials English German

The next survey brings together the results for tense frequency and the relative frequency of shifts into present tense (tables 8 and 9).

	present tense forms	**past tense forms**
English	12.6%	87.4%
German	2.7%	97.2%

Table 8 : Relative frequency of tense forms

In table 8 the results for German speakers are split with respect to the dominant tense used, and the table shows the rate of occurrence of the present tense in texts in which the Präteritum is the predominant tense form. This provides an indication of the limited role of the present in German, since speakers who use both the Präteritum and Perfekt (4/20) account for the majority of present tense forms found in the German narratives, as shown in table 9.

German	**present tense forms**	**past tense forms**
"Präteritum"-speakers	27.4%	72.6%
"Perfekt"-speaker	72.3%	27.7%

Table 9: "Präteritum"-speakers and "Perfekt"-speaker

The comparison, both inter- and intralinguistically, clearly illustrates the cross-linguistic differences. English speakers tend to shift into the present tense much more frequently than the German speakers (with the exception of the small group of speakers using the Perfekt, as mentioned). German texts are more consistent, exhibiting an anaphoric temporal structure throughout. The occasional shifts in tense forms are not motivated by a switch to an external viewpoint and do not occur in contexts which are relevant in this respect (see table 10).

	narrator viewpoint	**phasal segmentation**	**state**	**situation reoccurs**	**activity protagonist/environmental force**
English	15.8%	15.8%	28.5%	4.7%	34.9%
German	–	–	26.0%	–	73.9%

Table 10: Contexts for switches and corrections

4.3.2 Discussion

The figures presented above illustrate the cross-linguistic differences, and we now come to the factors which determine this overall picture. In earlier studies on narratives we sought evidence for the different principles of construal in the linguistic means which are grammaticised in the languages in question (see Carroll and von Stutterheim in press, von Stutterheim and Klein in press, von Stutterheim and Nuese in prep). The relevant structural property in the case of narratives lies in the categories grammaticised in the verbal system, as mentioned earlier. The main differences between the German and English verbal systems are not given by tenses but by aspectual categories. These categories are bound to viewpoint (see the definition in Comrie 1976, Klein 1994) and can be expected to have consequences for the construal of events in actual language production.

English has a marked grammatical category for ongoingness which is not constrained with

respect to temporal location, and a perfective aspect which is constrained with respect to the deictic TA. The imperfective aspect requires an external reference point, a topic time in relation to which a situation is viewed as ongoing. This can be the *now of an experiencer,* as outlined in section 4. But there is no basic constraint in English which excludes the choice of another reference point, as evidenced in the spontaneous past time narratives. The decisive factor which determines the selection of one or the other option may lie in the nature of the dynamic situation and whether it includes actual personal involvement on the part of the narrator or not. In other words, if the discourse situation suggests a reportive perspective (recount the content of a film), and the language provides grammaticised means which can easily anchor a reportive or external perspective – speakers are more likely to avail of this option in the construal of events. From this follows the need for an external TT, independent of TSIT, anchored in terms of the *now* frame of the observer or experiencer.

German speakers do not have a grammaticised imperfective aspect. The means available for aspectual marking predominantly encode the perfective viewpoint (verbal system and morphological means)[6]. The focus on the endpoint and post state of a situation provides the grounds for an anaphoric strategy in TT management. The interference of TA is secondary in that a post state perspective allows for a convergence of TA and TT. There are occasional switches to the deictic anchor *now* which, however, do not affect the status of the overall temporal frame.

These considerations lead to the conclusion that given different grammatical systems speakers develop specific principles in the construction of referential frames and with this the types of event/time relations which they incorporate. We put forward the hypothesis that speakers of languages with the imperfective as a grammaticised category have a preference for locating situations within an externally anchored frame. Speakers of languages with no systematic aspectual system encode situations holistically, and this provides the relevant informational components for anaphoric referential progression in the temporal domain. As illustrated in the analysis, speakers of German do not set up a global referential frame (i.e. which holds for the narrative as a whole) that is anchored in respect of an external viewpoint. The grammaticised means which facilitate this on a systematic scale are absent. This cross-linguistic contrast is not trivial since speakers are, on the whole, bound by the logic of a consistent frame in order to ensure the basic requirement of text structure – its coherence. The cross-linguistic differences thus lie in the overall referential frame which speakers establish, and there is a logical interrelation between the type of temporal relations selected and specific patterns of event construal.

The language-specific grammatical means relate not only to what Levelt (1996) has termed *microplanning*, they also determine planning at the *macrolevel* where speakers "consider what to say", since there are also significant cross-linguistic differences in the type of information which is selected for mention in the narratives (cf. Carroll and Lambert in press; von Stutterheim and Nuese in prep.). These wide ranging implications of grammatical structure for information organisation in language production have to be recognised in order to explain the complex but consistent picture revealed by the empirical facts.

6. Conclusion

In our analysis we have brought together levels which are traditionally the object of different linguistic subdisciplines. The mediating *locus* is the speaker's conceptualiser. Planning processes at this level require the adaptation of extra-linguistic content to linguistic structure, in particular to grammatical structure.

Each grammatical system provides a specific format which the speaker, when trying to

solve a complex verbal task, has to comply with. Just as some formal languages are better qualified for solving certain types of tasks, given their specific syntactic properties, natural languages will also reflect similar constraints. The general logical principles behind all artificial languages are probably universal, just as the reservoir of possible categories and relations in natural language syntax can be formulated at an abstract universal level. A specific grammar, however, manifests a weighted selection of this elementary system and thus focusses certain categories and relations while others are doomed to irrelevance. It is important to stress that there are probably no absolute constraints at a representational level, but it is much more effective to use a format which is tuned to the specific task.

Our findings with English and German reveal a similar type of implication. Grammaticised categories point the way to an optimal format for processing conceptual substance. Thus the optimal format for organising information differs according to the grammatical system. This is what the data have shown. On the side of the speaker the *optimal format* for encoding time-event structures is reflected in a clear preference for one referential frame over another.

Notes

[1] Note that these features hold for texts in everyday communication. In poetic or literary texts they can be violated (poetic license) in order to achieve effects which go beyond informativeness.

[2] In Klein (2000a), it is argued that we actually should speak of an 'argument-time structure', since various time spans can be relative to certain arguments. The transitive form *to close*, for example, has two time spans for the direct object and one time span for the subject. Since this plays no role in the present context, we shall stick to the simpler notion of 'temporal structure'. Note, incidentally, that the 'two-state verbs' in Klein (1994) represent a special – and actually a simple – case of such a temporal structure.

[3] A distinction should be drawn between the temporal structure of what the film shows and a particular speaker's representation of it, since the latter constitutes the actual speaker's text.

[4] Hopper 1995 notes that English speakers often talk about factive events in the past under the perspective of the current *now*.

[5] Opinions vary considerably on how to analyse the German Perfekt and how it related to the Präteritum. The problems are eminent, and we shall not try to examine them here but just look at how speakers indeed use these forms in the given task (but see Klein 2000a for a discussion and re-analysis).

[6] On the role of aspectual categories in L1 acquisition (English/German) see the discussion in von Stutterheim and Nuese in prep.

References

Bäuerle, Rainer   1979 *Temporale Deixis – temporale Frage: Zum propositionalen Gebrauch deklarativer und interrogativer Sätze*. Tübingen: Narr.

Brown, G. and Yule, C.   1983 *Discourse Analysis*. Cambridge: Cambridge University Press.

Carroll, Mary 2000   Representing path in language production in English and German: Alternative perspectives on Figure and Ground. In: Christopher Habel and Christiane von Stutterheim (eds.), *Räumliche Konzepte und sprachliche Strukturen*, 97–118. Tübingen: Niemeyer.

Carroll, Mary and Lambert, Monique   in press   Information structure in narratives and the role of grammaticised meanings: A study of adult French and German learners of English. In: Christine Dimroth and Marianne Starren (eds.), *Information structure, linguistic structure, and the dynamics of acquisition.* Amsterdam: Benjamins.

Carroll, Mary and Stutterheim, Christiane von   in press   Typology and information organisation: perspective taking and language-specific effects in the construction of events. In: Anna Ramat (ed.), *Typology and Second Language Acquisition*. Berlin: de Gruyter.

Comrie, Bernhard   1976 *Aspect: An introduction to the study of verbal aspect and related problems*. Cambridge: Cambridge University Press.

Fabricius-Hansen, Cathrine   1986   *Tempus fugit: Über die Interpretation temporaler Strukturen im Deutschen*. Düsseldorf: Schwann.

Garrad, Simon C. and Sandford, Anthony J.   1988 Discourse   models   as   interfaces between language and the spatial world. *Journal of Semantics* 6: 147–160.

Goldberg, A. E.   1996 *Conceptual Structure, Discourse and Language*. Standford: CSLI Publications.

Hopper, Paul J.   1995   The Category 'Event' in Natural Discourse and Logic. In: Werner Abraham, Talmy Givon and Sandra Thompson (eds.), *Discourse Grammar and Typology*, 139–153. Amsterdam: Benjamins.

Kamp, Hans and Reyle, Uwe   1993 *From Discourse to Logic. Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Dordrecht: Kluwer Academic Publishers.

Kintsch, Walter and van Dijk, Teun A.   1978   Toward a model of text comprehension and production. *Psychological Review* 95: 163–182.

Klein, Wolfgang   1994   *Time in Language*. London: Routledge.

Klein, Wolfgang   2000a An analysis of the German Perfect. Language 76/2: 358–382.

Klein, Wolfgang   2000b Was uns die Sprache des Rechts über die Sprache sagt. *Zeitschrift für Literaturwissenschaft und Linguistik* 118: 116–148**.**

Klein, Wolfgang and von Stutterheim, Christiane   1987   Quaestio und referentielle Bewegung in Erzählungen. *Linguistische Berichte* 109: 163–183.

Labov, William   1972   The transformation of information in narrative syntax. In: William Labov (ed.), *Language in the inner city,* 354–396. Philadelphia: Philadelphia University Press.

Levelt, Willem J. M. 1996   Perspective Taking and Ellipsis in Spatial Descriptions. In: Paul Bloom, Mary A. Peterson, Lynn Nadel and Merrill F. Garrett (eds.), *Language and Space,* 77–107. Cambridge/Mass.: The MIT Press.

Ochs, E., Taylor, C., Rudolph, D. and Smith, R.   1992   Story-telling as a theory-building activity. *Discourse Processes* 15/1: 37–72.

Partee, Barbara H.   1983   Temporal and Nominal Anaphora. *Linguistics and Philosophy* 7, 243–286.

Paul, Hermann   1882   *Principien der Sprachgeschichte*. Jena. Niemeyer.

Reichenbach, Hans   1947   *Elements of Symbolic Logic*. London: Free Press.

Roßdeutscher, Antje 2000   Lexikalisch gestützte formale Textinterpretation. *Arbeitsberichte des Sonderforschungsbereichs 340, 117.* Tübingen. (Habilitationsschrift Universität Stuttgart).

Smith, Carlota T.   2000   *The Domain of Tense.* International Round Table `The Syntax of Tense and Aspect". Ms. Nov. 2000, Paris.

Stutterheim, Christiane von 1997 *Einige   Prinzipien   des   Textaufbaus: Empirische Untersuchungen zur Produktion mündlicher Texte.* Tübingen: Niemeyer.

Stutterheim, Christiane von and Lambert, Monique   in press   Crosslinguistic analysis of temporal perspective in text production. In: Henriette Hendricks (ed.), *The structure of learner varieties*. Berlin: de Gruyter.

Stutterheim, Christiane von and Klein, Wolfgang 2002 Quaestio und l-perspectivation. In: Carl Friedrich Graumann and Werner Kallmeyer (eds*.), Perspectivity and perspectivation in discourse*. Amsterdam: Benjamins.

Stutterheim, Christiane von and Nuese, Ralf and Murcia-Serra   2002   Crosslinguistic differences in the conceptualisation of events. In: B. Behrens and Catherine Fabricius-Hansen and S. Johansson (eds.), *Information structure in a cross-linguistic perspective*. Amsterdam: Rodopi: p.179-198.