

Klein, W., & Stutterheim, C. von (2002). Quaestio and L-perspectivation. In C. F. Graumann & W. Kallmeyer (eds.), *Perspective and Perspectivation in Discourse* (pp. 59-88). Amsterdam/Philadelphia: John Benjamins.

Quaestio and L-perspectivation

Christiane von Stutterheim and Wolfgang Klein

Professor Dr. Wolfgang Klein
Max Planck Institut für Psycholinguistik
Wundtlaan 1
6525 XD Nijmegen
klein@mpi.nl

Prof. Dr. Christiane von Stutterheim
Institut für Deutsch als Fremdsprachenphilologie
Plöck 55
D-69117 Heidelberg
stutterheim@mail.idf.uni-heidelberg.de

1. Choices

Whatever a language allows its speakers to express - it usually provides them with more than one way to express it. Thus, one and the same historical fact may be stated in an active and a passive variant:

- (1a) In 1453, the Turks conquered Constantinopel.
- (1b) In 1453, Constantinopel was conquered by the Turks.

The fact reported by these two sentences is the same; still, there is a difference in the way in which it is 'presented'. Another familiar variation concerns word order:

- (2a) In 1453, the Turks conquered Constantinopel.
- (2b) The Turks conquered Constantinopel in 1453.

Again, the historical fact is the same. But it seems that these two sentences

answer two different questions; whereas (2a) is perfectly appropriate in answer to ‘What happened in 1453?’, this is not the case for (2b); it is not fully excluded, but somewhat odd, whereas it makes a perfect answer to ‘When did the Turks conquer Constantinople?’. Somehow, the adverbial *in 1453* plays a different role in these alternative sentences: in (2a) it is somehow presupposed, and the real answer to the question is the remainder of the clause, the action of the Turks. In (2b), it is the direct answer to the question - what was asked for is the date; it is more ‘in focus’ than in (2a). Another pair of variants is this:

- (3a) In 1453, the Turks conquered Constantinople.
- (3b) In 1453, the Turks were conquering Constantinople.

In this case, the formal difference is not so very much in syntax but in morphology, and this variation, again, is accompanied by a slightly different way in which the same event is presented: in (3b), we are somehow placed within this event - they are just doing it -, whereas in (3a), it is presented more as a completed fact.

In all of these cases, the expression is a finite clause. English also offers possibilities to express the very fact in an entirely different construction, as in (4):

- (4a) Constantinople’s conquest by the Turks in 1453
- (4b) The Turks’ conquest of Constantinople in 1453
- (4c) Constantinople, which was conquered by the Turks in 1453

In these three cases, the construction is no longer a full sentence but rather a complex nounphrase. It is hard to tell what precisely the difference between them is - but there is clearly a difference.

It is easy to come up with many other variants in which the same historical event can be put into words. But the point to be made here should be clear: in one and the same situation, and with respect to one and the same fact, there are numerous ways of expressing it. These variants are in a way equivalent; but they are not entirely equivalent; there are sometimes subtle, sometimes substantial differences, and it is at the speaker’s discretion to choose between them.

In all examples discussed so far, the speaker’s choice is between various *structural* options. There are also *lexical* options: depending on the lexical wealth of the language, the speaker may select between various words when describing the same fact; he may replace *Turks* by *Osmans*, or *Constantinople* by *former Byzantium* or by *Istanbul*, or *conquered* by more prosaic *took*. Again, these options are equivalent in one way, but not in another, and it is left to the speaker to opt for the one or the other.

Lexical choice and structural choice reflect elementary properties of natural language. Whatever precise form a language may have, it inevitably

has a repertoire of elementary meaning-bearing expressions (words or morphemes), on the one hand, and rules which allow the construction of complex expressions from elementary ones, on the other: a language consists of a lexicon and a morphosyntax. There is a third type of choice which is perhaps less obvious but no less deeply rooted in the nature of human language. In general, the interpretation of an utterance is fed by two sources of information - by linguistic information proper, or 'expression information', and by contextual information. The former stems from what is in the words and in the way in which these are combined. The latter may be of various types. It may come from what has been said before, it may stem from the perceptual situation, or it may be part of the interlocutor's world knowledge. All of this is familiar, and the reason why it is mentioned here is that this permanent interplay between expression information and context information provides the speaker, when talking about something, with a third choice: he must decide what is made explicit by his expression and what is left to contextual information. In all sentences above, for example, the temporal adverbial *in 1453* may be replaced with *then*, *at that time*, *soon afterwards*, *four hundred and forty-seven years ago* and other phrases, which (in a given context!) express the same information. Hence, we have a third type of choice, to be labeled here *contextual choice*.

Very often, these three forms of choice are not independent. Often, a particular form of contextual choice must be expressed by a particular word; deictical and anaphorical constructions illustrate the point. Similarly, the selection of a particular lexical item may constrain the syntactic constructions between which a speaker may choose; some verbs, for example, have a passive, whereas others do not; some pronouns ('clitic elements') are restricted to particular positions and hence exclude certain word order patterns, etc. But this regular interaction does not affect the general point to be made here. Whenever a speaker wants to produce an utterance in a particular language, he has to decide between various options with which this language provides him. In particular, he has a lexical choice, a structural choice, and a contextual choice. In each of these, the alternatives between which he has to choose are equivalent in one way, but not equivalent in another way. His eventual decision, therefore, reflects a particular way of presenting what he wants to say - it reflects a particular 'perspective' on the facts stated. If we want to understand the phenomenon of 'perspective taking' in language, we must analyse how these three types of choice function in language production.

This is perhaps not the most common way to approach the problem of perspective and perspectivation in language. We have chosen it here for three closely interrelated reasons. First, there is a number of fields in which the notion of perspective has a relatively clear definition, for example in mathematics, in visual perception, and maybe in the history of art. In its application to language, however, it has a strong metaphorical character, with all advantages and disadvantages of metaphorical extension: it deals

with an important phenomenon, it captures crucial insights, but it is not clearly defined, and it misses what is specific to language. Take, for example, a core notion such as ‘viewpoint’. In visual perception, this notion is comparatively well-defined - it is the spatial position at which the observer is, or imagines to be. In language, this notion of viewpoint also plays a role, for example in the use of deictic terms such as *left, right, here*, etc. This, however, is only a special case. It is much less clear what ‘viewpoint’ means, for example, in the comparatively well-studied case of (grammaticalised) aspect. Verbal aspect is a category found in many, if not most, languages. The difference between simple form and expanded form, as in *The Turks conquered Constantinople vs The Turks were conquering Constantinople* (cf. example 3 above) illustrates this category. Linguists normally characterise different aspects, as ‘various ways of viewing the situation’ and similar characterisations (Comrie 1976). In fact, the linguistic term ‘aspect’ is a translation of Russian *vid* ‘view’. But clearly, this notion of viewpoint cannot mean two different concrete places from which something is seen. In the case of ‘viewpoint aspect’, the concrete position of speaker, listener, or any other person, plays no role at all, nor is anything ‘seen’ differently, if this word is understood in its literal sense. In example 3a, b nothing is seen at all. Even more problematic are cases in which the difference between the use of active and passive, or varying subjective evaluations of what is expressed, are accounted for in terms of ‘viewpoint’. This is a very suggestive and intuitively appealing way to give the ‘flavour’ of the difference; but it can only be a starting point for an explanation proper of what is going on in these cases. Such an explanation must be specific to the particular cognitive domain in which the phenomenon of perspective and perspective taking is observed. It would be more appropriate, therefore, to talk about ‘V-perspective’ and ‘V-perspectivation’ (V for ‘visual’), on the one hand, and ‘L-perspective’ and ‘L-perspectivation’ (L for language), on the other. It is the latter we are dealing with here. This does not preclude that there are many similarities; but simply talking about ‘perspective’ obliterates the differences - and these are not minor. This already has brought us to the second point, discussed below.

The *principles* of perspective-taking are in many ways different when language comes into play. Someone may see a tree from various perspectives, he may draw a tree from various perspectives, and he may describe a tree from various perspectives. In the two former cases, the difference in perspective is essentially determined by the nature of our perception. In the latter case, it is determined by highly abstract cognitive principles which determine the choice of words and constructions against the background of shared assumptions among the interlocutors. Now, this example is relatively simple because the common element, the tree, is a physical object. But perspective-taking in language goes far beyond this case. In examples 1-4 above, the common element which is presented from

various perspectives, is a historical event; hence, it is a much more abstract entity than, for example, a tree which can indeed be seen, drawn, and described from different angles. And there are still more abstract cases in which the intuitive notion of perspective-taking in language makes sense, for example arguments or instructions. In all of these cases, however, the principles of perspective-taking are much the same; they are constraints on choices, in particular the ones mentioned above.

Third, L-perspectivation is a highly complex, active process which involves a series of interrelated decisions on many levels. This is already obvious for the elementary constructions 1-4 above. But in actual fact, the production of such a sentence is normally part of a longer communicative task, in which the speaker transforms selected elements of his knowledge into concrete sound waves. In each phase of this process, the speaker is permanently faced with a multitude of choices. In our example, we have assumed that it is already clear what 'the fact' to be talked about is - here the conquest of Constantinople in the year of 1453. But the decision to talk about something specific is, or at least maybe, one of the many choices which the speaker has to take. This becomes clearer if we look at typical communicative situations in which a speaker is challenged to produce a specific text. Suppose, for example, someone is asked by a passer-by: 'How do I get from here to the station?'. If the interlocutor takes on this task and sets out to fulfil it by verbal means, then he may be able to do so by producing a single sentence. But as a rule, he will produce an ensemble of well-organised sentences, that is, a text. An essential part of this process is the decision about which information is to be selected for verbalisation, and which part of this selected information is to be packed into a single sentence. Only then does it make sense to ask what structural form this sentence should have and which lexical items it is to contain. Hence, the question of perspectivisation crops on very different levels, and on each of these levels, it takes a slightly different form. What is constant, though, is the idea that there is always a set of alternatives, which are equivalent in one way and not equivalent in another. The question which we will address in the main part of this chapter is now: Are there overarching principles which determine, or at least influence, the speaker's choice and hence the particular perspective which he takes in his production?

The fact that the choice, on whichever level, is always a task of the speaker does not mean that the listener has no role in this process. In fact, he may even 'set the stage' by posing a particular question, which the speaker then has to answer. The example of the route directions question above illustrates the point. Even if there is no explicit question of this sort, the speaker may behave as if there were such a question and a particular listener, or group of listeners, who asks it. As we shall see later, the particular - implicit or explicit - question which the speaker sets out to answer is a core constraint on possible perspectives. It should be clear, however, that it is still left to the speaker how he deals with such a question.

He may not accept it at all, he may accept it but ignore some of the constraints which it imposes on his production, and even if he accepts it in its entirety, the question still leaves him with many degrees of freedom as regards the various choices he has to make. This will be discussed in sections 3-6 of this chapter. It will be useful to begin, however, with a brief look at the various stages of the production process and the way in which the speaker can go the one or the other way.

2. Levels in production

There is considerable theoretical and empirical work on human language production. In this section, we shall not try to review this work (see, for example, Levelt 1989 and Herrmann and Grabowski 1994), but briefly discuss some elementary facts which every theory must deal with. It will be helpful to use a simple example. Suppose someone was involved in a traffic accident and is then asked on some occasion to speak about what happened during this accident. Then, at least the following four cognitive levels may play a role in L-perspectivation.

a. Intake

The underlying event itself, the traffic accident in this example, is a complex agglomeration of individual facts and incidents, of persons and objects involved, of temporal, spatial and causal relations. Only some of these are perceived and stored by the individual participants. The process which leads from the 'objective situation' to the particular mental representation which an individual forms of this situation is in many ways 'perspective-driven'. The most obvious source of selection is given by the fact that the perceiver has only restricted sensory access to this situation; he has a particular 'viewpoint' in the literal sense of the word:

In correspondance of the bodily nature of the perceiving and acting subject, the objects of experience are only present in those aspects that are seen from the spatiotemporal point of view taken by the subject. To be present in *aspects* with respect to a given viewpoint is the basic meaning of *perspective*. (Graumann 1989: 96)

But there are many other factors which govern the way in which a mental representation is built up. They range from very general principles that determine the cognitive processing of visual and other sensory input to highly subjective preferences for one or the other aspect of the situation. Some features of a situation are perceptually more salient than others and hence more likely to be noticed and stored. There are social conventions on what is more and what is less relevant in such a situation. These conventions may be culture-specific, they may be specific to a particular

speech community or social group. And finally, there are subjective preferences which depend on the individual's personal history, her interests, his momentary emotional disposition, etc.

The perspectival nature of the intake is not necessarily related to how this situation is later presented in a verbal account. At that point, language plays a role only insofar as particular properties of a language may influence the observer's attention. Thus, a deictic system which does not discriminate between 'left' and 'right', 'front' and 'back' may lead the observer to ignore certain aspects of the spatial arrangement, or to give them less weight than a system of spatial reference which depends on the position of the speaker (see Levinson 1997). But the existence and the weight of such factors is arguable, and to the extent to which they exist, they are surely weak. In general, the intake is not an active process based on linguistic knowledge; it should rather be seen as a language-independent filter which rules out certain features of the situation and lets others pass. But obviously, this particular filtering has strong consequences for what can later be selected in text generation. It creates the knowledge basis on which the speaker can draw. The clearest case is surely the 'initial vantage point' with which it provides the participant: a time and a position, which then later may surface in the participant's account of this situation, for example in form of a particular 'topic time' (see section 4.3 below).

b. Update

Normally, some time elapses between intake and text generation. In the course of this time, the original representation is in many ways transformed in the observer's memory. Certain features are dropped, others are added, some features may change their relative weight, they might gain or lose certain emotional connotations, moral evaluations are added that were absent from the original representation, certain observations may be re-interpreted in the light of later experiences, or due to what other parties say about the situation etc etc. This steady transformation is not arbitrary. It depends on a wealth of factors, ranging from very general properties of human memory to highly specific and subjective preferences of the individual. Hence, this permanent update may also be considered to be 'perspective-driven'. But then, the original, non-metaphorical meaning of this notion is largely lost and replaced with an understanding of perspective, that is highly suggestive but much less well-defined; we might speak here of 'M-perspectivation' (M for memory), in contrast to V-perspectivation and L-perspectivation. Again, there is no reason to assume that language plays a substantial role in this updating. But its eventual result restricts the available knowledge on which the speaker can base his text generation.

c. Forming a Discourse Representation

In a specific communicative situation, the speaker sets out to verbalise part of what he has stored in memory. This is the point where language, and

hence L-perspectivation, comes in. Very often, text generation is initiated by a particular question which the speaker is supposed to answer, for example: 'What did you see?' or 'What did the truck that came from the left side look like?' or 'What did you do yourself?' or 'Have you ever been in danger of life?'. There need not be such an explicit question; the speaker may simply want to speak, for one reason or another, about what he experienced some time ago; in a way, he is posing such a question to himself. In each case, however, there is a particular communicative goal that causes the speaker to activate his updated mental representation and to create a new, temporary conceptual structure, a *discourse representation* (other labels are *conceptual structure*, *preverbal message*, and similar ones). It is this discourse representation which underlies the concrete text to be produced by the speaker. The discourse representation differs in four crucial respects from the final representation in memory:

Selection. The discourse representation contains only selected parts of the entire representation. This selection is determined, or at least constrained, by the particular communicative goal. Selection not only concerns which bits and pieces of the original representation are chosen for verbalisation at all but also at which level of 'granularity' these are to be put into words. Thus, the speaker may decide to give only a very global account of some subevents but go into much more detail for other subevents, or to begin with a global account and then refine it, etc.

Addition. Typically, the material contained in the discourse representation includes bits and pieces of information which stem from other sources than from what the speaker has stored about the original situation. He may add comparisons, try to give explanations, relate the particular subevents or participants to other events or persons, include moral and other evaluations, etc.

Linearisation. Somehow, the various conceptual units which belong to the discourse representation must be sequentially ordered. In some cases, this is comparatively easy; narratives often reflect a simple temporal order of subevents. But as a rule, linearisation is much more problematic. Already in the case of a car accident, subevents may overlap or be completely simultaneous, and then, the speaker has to make a choice what to represent first. It may also be that the level of granularity is not kept constant, for example if the speaker chooses to give first a rough account and then go into more detail with some of the subevents. In this case, there is still a temporal ordering, which provides a natural base for linearisation. Other cognitive representations do not involve such a linear order, for example in the case of a picture description or a logical argument. In these cases, the speakers may choose very different 'perspectives', under which the entire information to be verbalised is put into sequential order (see, e.g., Levelt

1982).

Function assignment. There is good reason to assume that already before the decision on the final language-specific form is taken, various bits of information are marked for special linguistic features. This includes, for example, the decision whether a certain entity - person or object - should be encoded as the subject or as the object, whether an active or passive way of presentation is chosen, what belongs to the topic component and what belongs to the focus component of the final utterance, etc. Form and amount of this marking are much at dispute, and in order to keep it different from the eventual language-specific marking of, for example, the grammatical subject or the passive form of the verb, we called it here 'function assignment'.

All four properties are the result of continuous 'perspectival' choices. This is obvious for both aspects of *selection*: the speaker weighs the various elements from his stored knowledge representation and decides which ones to include and on which level of granularity. Note, however, that this decision does not reflect what has been called 'contextual choice' in section 1. Contextual choice is the speaker's decision about *what is explicitly stated and what is left implicit because the speaker assumes that the listener can infer it from context*. Hence, it presupposes that the speaker wants that the listener eventually has this information. *Selection* concerns the speaker's decision which information in which detail the listener should have due to his, the speaker's, efforts. It is an adaptation to the listener - but not in the sense that the speaker plans his text such that it fits the listener's contextual knowledge.

Analogous considerations apply for *addition*. The extent to which the speaker wants to provide the listener with subjective comments and evaluations on the event, for example, is surely a matter of his 'subjective perspective' on this event. They may be very different, for example, if he is the victim of the traffic accident, a policeman or simply a bystander, they may be different if he is a passionate bike-driver, if he has a driving-license or not, if he had a car accident himself before, etc etc. But all of this has nothing directly to do with L-perspectivation.

The speaker may also choose different linearisation strategies; he may prefer, in the case of the traffic accident, to follow a strictly temporal order wherever this is possible. But he may also choose to tell the events backwards, as seen from the result, and then explain how this result came about. Similarly, he is free to choose where to plug in additional information, such as comments or background statements. Again, one may call this choice perspectival, but it should be clear that this, again, is a different notion of perspective.

The situation is much more difficult with respect to *function assignment*, because the speaker's decision here immediately affects the choice of what to express in which concrete way. The decision, for

example, to present a particular participant of the situation as an agent, rather than as a patient, is a 'structural choice', in the sense of section 1. Similarly, the decision to assign certain subparts of the information to the topic component and other subparts to the focus component of the utterance immediately affects the structure of this utterance. Another decision taken at this point is the fixation of a particular 'origo', i.e., vantage point to which times and spaces are related and which is then reflected in the utterance by the choice of particular words, such as the deictic terms *left*, *right*, *here* etc. or by particular tense forms. The problem is here that, whatever the speaker's decisions are, they can be implemented in various ways in the final utterance. The speaker may be able, for example, to transform the same - or essentially the same - discourse representation into an English, a German or a French sentence. Or if he chooses his own position as the spatial origo, he is still free to say *here* or *where I was* or *at my position*. This has already brought us to the last level of text generation - the level in which eventual perspectival decisions are transposed to linguistic form.

d. Constructing a linguistic form

Whatever the discourse representation is - any language provides their speakers with very specific possibilities to implement it. The discourse representation may fix that a certain element belongs to the topic information; still, there are various ways in which this is marked - by intonation, by word order, by the choice of a particular particle which indicates the information status of some element. Similarly the discourse representation may have fixed that some subevent is presented as on-going, rather than as completed. Metaphorically speaking, the speaker puts himself, and wants to put the listener, 'into the situation', rather than seeing it from the outside. Then, he has still an option between, for example, a particular verbal aspect, an adverbial, or a combination of these verbal means. As was already said in the preceding section, it is very difficult to say to which extent the decision about the eventual form are determined by the formation of the discourse representation itself or by its translation in a concrete linguistic form, and opinions vary considerably on this point. Bierwisch and Schreuder (1992), for example, assume that the discourse representation - their 'conceptual structure' - is neutral with respect to the final linguistic form; as a consequence, it is the same no matter in which language the speaker wants to express it. Under such an assumption, there is no 'function assignment' in the discourse representation, as has been assumed here. We shall not discuss this point in detail (see von Steutterheim 1997, chapter 9). It seems clear, however, that there are intermediate stages between a language-neutral cognitive representation of what is to be expressed, on the one hand, and the final linguistic form, on the other. The precise number and nature of these interim stages is a matter of dispute. It may well be that there is no fixed set of levels at all but that actual

production is much more flexible, depending on the nature of the particular communicative task and maybe even the speaker's competence. We shall not try to sort this out here. The crucial point in the present context is this: It is exactly this passage within the entire text generation at which L-perspectivising comes in.

So far, we have given numerous examples which illustrate structural choice, lexical choice and contextual choice; others can easily be added. In fact, a great deal of the literature on perspective-taking in language consists of very striking examples and anecdotal evidence for this. But are there any *principles* which determine this part of the process - in other words, are there any general constraints that which are characteristic of L-perspectivation? This is the question which we will address in the next section,

3. The Quaestio

3.1 Questions and answers

What causes a speaker to activate parts of this stored knowledge, build a discourse representation and transform it into sound waves? In the easiest case, this is done by an explicit question on the part of some interlocutor. In the case of a traffic accident, such a question might be, for example 'Which car came from the left?' This question defines, within limits, the communicative task to be solved. The speaker may choose to answer it simply with a single sentence or even a part of such a sentence, for example

- (5) A BMW 730 came from the left.
- (6) A BMW 730.

What he is asked for by the question, is the specification of some entity. The question raises a set of alternatives - all those x's that could have come from the left on that occasion, and the speaker is challenged to select one of those - the one which really came from the left on that occasion. In other words, the question defines a choice, and it is up to the speaker to deal with this choice. In doing so, he is left with a certain amount of freedom. It is this freedom which allows the speaker, if he takes on the task at all, to set a particular perspective: the question imposes constraints on a possible answer - but it does not determine the answer, of course. In particular, the speaker may choose to go into more detail with respect to the predicates he is asked to specify, and may answer the question by a whole series of interconnected utterances - i.e., with a text:

- (7) From my position, I could not see it very well. Everything went so fast. But it was a big car, a limousine; blue, dark

blue. One of my neighbours had such a car. I guess it must be very expensive, one of these old-fashioned fossils...

Not all utterances in such a sequence are directly 'to the point', which is, here, to contribute to the specification of the object. The first utterance in 7, for example, highlights the role of the speaker's position in the intake phase and thus explicitly introduces a perspectival component. At the same time, it qualifies the reliability of the description: it has not only a spatial but also a modalising component. Clearly, such information can be relevant and important; but it does not directly serve to answer the question. We shall call those utterances which directly contribute to answering the question, its 'main structure', and those which give additional - and communicatively often important - material, its 'side structures'. Side structures can be of various type - comments, evaluations, background information, etc. Thus, they are a rich source of perspectivation in one sense of the word (cf. Sandig 1996). For the speaker, they may even be more important than main structure utterances, because they allow him to express his own subjective attitude. But note that this is one way in which the text reflects a particular perspective. What counts as main structure, and what as side structure, directly depends on the specific question which the text as a whole intends to answer. In the example, the speaker is asked to specify a particular object. Had the question been 'What happened next with the car that came from the left?', then the categorization as main or side structure would have been the reverse. In that case, the speaker's task would have been to specify a sequence of small events, which in its entirety constitutes a subpart of the entire traffic accident; and this sequence of subevents would then comprise the main structure of the text, its 'foreground'. Any other material - for example a more detailed description of a car, or of some other vehicle involved in the events - would belong to the side structures, or 'background', of the text. The concepts 'foreground' and 'background', are normally not introduced in this way but rather in terms of whether or not they are a 'narrative sequence'. It should be clear, however, that the background-foreground partitioning in narration is only a special case of a much more general phenomenon: they reflect various reactions to the underlying question.

There is a second important feature of 'textual answers' to a question. If the entire information is distributed over a series of several utterances, then certain meaning components within each utterance are maintained from the preceding utterance (or utterances), while other meaning components are freshly introduced. In this example, there is only one crucial entity referred to, the car; it is introduced in the question itself and then merely maintained in the subsequent utterances of the main structure. Had there been more than one vehicle, the related question, 'Where did the cars come from?' would have forced the speaker into a much more complicated pattern of referent introduction and maintenance.

Objects and persons, in brief, the referents of noun phrases, are not the only meaning elements that must be introduced or maintained across utterances. Temporal intervals and subspaces are others. Again, this crucially depends on the nature of the initial question. A question such as ‘What did the car look like?’ leaves time and space constant across the text - more precisely, across the main structure utterances of the text. A question such as ‘What did you observe?’, by contrast, normally invites regular switches in the spatial and temporal domain. Such a question provides the speaker with more degrees of freedom how to organise his discourse representation and how to put it into words.

Thus, a coherent text that is produced as an answer to a question involves a ‘referential movement’ within various semantic domains (or, as we shall say here ‘referential domains’), such as persons, place, time, and others. This referential movement is reflected in the use of specific linguistic means - it immediately influences the speaker’s lexical, structural and contextual choice.

Summing up, the structure of a text, is systematically constrained by the nature of the question which the text in its entirety is produced to answer. These constraints include:

- a. the partitioning of the text into main structure and side structures;
- b. the assignment of specific meaning elements to the topic component or to the focus of a main structure utterance;
- c. the ‘filling’ of various possible domains of reference within each main structure utterance;
- d. the referential movement within the domains from one main structure utterance to the next.

They narrow down the choices of the speaker, if he takes on the communicative task - but they leave him certain degrees of freedom, and thus various ways to ‘present his case’. It is this fact which allows the speaker to impose an L-perspective on what he sets out to say. We shall explain this in some more detail in sections 3.2 - 3.4 below.

Not all texts are solicited by an explicit question. People sometimes take the liberty to speak without being asked, or, to put it alternatively, they define their own questions. In other words, the question which underlies the production of an utterance or a longer text may be explicitly asked, or it may be implicitly given - either by the speaker himself, or just because the situation is such that it suggests a particular question to be answered. We shall use the term *quaestio* for both cases - a real question in a real dialogue, or an implicit question with a similar function. This idea is not new; in fact, it is already found in ancient rhetoric, although its structural consequences on the concrete form of texts and individual utterances have hardly ever been a concrete object of investigation in the rhetorical tradition (for a discussion, see von Stutterheim 1997, chapter two).

As was said above, the quaestio imposes certain constraints on the speaker's answer, providing him at the same time with a limited set of options among which he can choose. In no case is the speaker bound to obey these constraints. If he does not, then this either leads to side structures - for the entire text - or to particular rhetorical effects. In what follows, we shall discuss constraints b. - d. in some more detail (the constraint which leads to a partitioning in 'foreground' and 'background' was already discussed above).

3.2 *Topic component and focus component*

In general, an utterance can be used to answer very different questions. Consider, for example 7, which could be part of an account of the traffic accident:

- (7) The bike came from the left.

Such an utterance could be made in answer to:

- (8) What came from the left?
(9) Where did the bike come from?
(10) What happened?
(11) What happened next?

In each of these cases, sentence 7 settles an open alternative raised by the question - it specifies one out of a set of candidates at issue. In 8, the alternative to be settled includes those entities - vehicles, as the context suggests - that could have come from the left at that time; and the one is selected among the various possible candidates is 'the bike'. In 9, the possibilities include all the places where the bike could have come from, and what is selected is a the particular place described by *from the left*. In 10, the choice is between all those events that could have happened, and what is selected is the particular event of the bike coming from the left. In 11, the alternative raised is very similar to 10, but the time is explicitly restricted: what is asked for are the events that could have happened at some time after the time of a previous event, and one of those is selected and specified in the answer.

In all of these cases, we have an alternative set by the question, and an element from this alternative that is specified in the answer. Such an alternative at issue will be called the 'topic' of the utterance, and the element from that alternative which is chosen and specified, the 'focus' of the utterance. Topic and focus, as these terms are used here, are components of the 'meaning' of an utterance - a person, an action, a time span. They are not the verbal means - a word or an entire construction - which express this meaning. When talking about a speaker's choice in the production process,

we must carefully distinguish between the choice of, for example, a person he wants to talk about, and the choice between various ways to do this. Both decisions may reflect a particular form of ‘perpectivation’; but only the latter is what is understood here to be L-perspectivation, i.e., the combination of contextual, structural and lexical choice.

It is also important to distinguish between the expression of a topic or focus, on the one hand, and the marking that that entity ‘is’ topic or focus, on the other. In 7, when said in answer to 8, the focus is the bike, and this focus is referred to by the expression *the bike*. With a different contextual choice, it could also have been expressed by *that bike* or by *it*. Its focal status is marked by intonation. Intonation is not the only device available to make clear what the focus (or the topic) is; word order or specific particles serve also as devices for this, at least in some languages. Very often, however, it is not explicitly marked at all, or is ambiguous.

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

- (12) There came a car from the left and a car from the right.
Which car hit the bike? - The one from the left.

The alternative raised by the question is between the two cars, and the focus of the answer is ‘the one from the left’. But obviously, this car has already been introduced, just as everything else in the answer: it is ‘maintained information’. This is clearly reflected in the form in which it is referred to in the utterance. The rest of the answer is a repetition of the expression of the topic.

The quaestio answered by an individual utterance may also be derived from the ‘higher’ quaestio of a whole text, to which the utterance belongs and which the text in its entirety is intended to answer. In this case, there may be global constraints what belongs to the topic component and to the focus component of the individual utterances. Take a question such as ‘What happened to you last week?’ which elicits a narrative text. It asks for the specification of some complex event, which the speaker may subdivide into a series of sub-events, each of them happening during some time interval t_i within the time of the total event. Thus, the quaestio of the whole text can be broken down into a temporal sequence of quaestiones answered by all of those utterances which specify one of the subevents, roughly

- 1: What happened to you at t_1 ?
- 2: What happened to you at t_2 ?
- 3: What happened to you at t_3 ?

n: What happened to you at t_n ?

For each individual utterance, we have a time span which belongs to the topic component of this utterance. These time spans need not be individually specified, they follow from a general principle - the 'topic condition' of narrative texts. This condition gives us the 'backbone' of the narrative, its main structure. It may be interrupted at any point by utterances which do not answer the general quaestio but rather subquaestiones such as 'What did you think of this?', 'What's the point?', 'Would you do this yourself? Are you sure?', etc., which all lead to side structures of different type. The specification of a particular side structure may also extend over several utterances. For example, a narrative sequence can be interrupted by a descriptive sequence, or vice versa.

Suppose now you happen to be the victim of the traffic accident, and the quaestio is 'What happened to you after you saw the car?' It specifies a time span, which is after the time span of the previously mentioned event. It also specifies a person, the addressee in this case ('to you'), and moreover it indicates that what is asked for is an event, not a state, as would be the case with the question 'What was the situation at that time?' Hence, the quaestio narrows down the focus of the answer to include only those events which could have happened to the addressee during the intervening time since the event reported before. The answer, which settles the alternative *can* repeat the topic (fully or in part); it can also elaborate on it. But it *must* contain a part which expresses the focus, that is, which specifies an event that meets the conditions mentioned above. This has many consequences for the structure of the utterance. If, for instance, the topic time - the temporal interval about which something is said - is expressed by an adverbial, the protagonist by a noun phrase, and the event in the narrower sense by a verb, and if furthermore the language in question has a rule 'topic expression before focus expression', then a word order such as 'noun phrase - verb - adverbial': the verb must be last. This may conflict with purely syntactic constraints on word order, such as 'the verb is in second position', and different languages have found different ways to solve competing requirements of this type. Intonation, special particles, cleft constructions, passives, etc., or by simply not requiring consistent marking of what is focus and what is topic in these cases. Thus, the speaker's structural choice is narrowed down by two factors: first, by the structural potential of the language, and second, by the quaestio. In section 4.3, we shall come back to this point.

3.3 Domains of reference within an utterance

Any proposition - that is, the content of a full sentence - is a web of meaning components or 'referents', as we shall say here. These are of various kinds - spatial, temporal, personal, modal etc. Ever since the days of

the Greek philosophers, there have been innumerable proposals to categorize various domains of referents. It is useful to distinguish at least the following five:

- a. time spans
- b. places
- c. persons and objects
- d. states, events, properties
- e. modalities, such as the 'possible', 'real', 'necessary', 'fictitious', etc.; this should also include other characterisations by which the speaker indicates a particular subjective attitude to what he expresses.

An utterance selects referents from these domains and integrates them into a whole, the proposition. This is not done arbitrarily; it follows certain principles. Most typically, an element from the domain 'persons and objects' is combined with an element from the domain 'states, events, properties' to form the 'inner core' of the proposition; this inner core is then located in time and space. The resulting combination of referents is then provided with a modality which, in one way or the other, fixes its reality status. This simple picture can be complicated in various ways. Not all domains of reference must be represented; it does not make much sense to associate a proposition such as the one *seventeen is a prime number* with a place. On the other hand, a particular domain of reference may be repeatedly represented (for example, there may be several protagonists of the action; similarly, a sentence may involve more than a single time span and a single place). Therefore, the result may be a very complex structure of referents from these five domains, brought about by the speaker's contextual, structural and lexical choice. What the quaestio does, is to impose constraints on these choices, and thus, on L-perspectivation. It defines, for example, a time frame, about which the speaker is asked to say something; but it is up to the speaker to deal with this time frame: he may choose to speak about the entire time frame, he may split into subintervals, he can leave these implicit or mark them by tense or adverbials, etc. Very often, the quaestio also introduces a vantage point in relation to which the speaker is supposed to describe the various referents. Take, for example, a 'court question', such as 'What did you observe from where you were?'. It defines the speaker, sets a temporal frame and a place (time and position of the speaker during the intake phase). Moreover, it also imposes a certain 'modality' on what the speaker is to express. In this example, he is supposed to make assertive statements. But the modality is not 'real with respect to the event' - it is 'real with respect to the speaker's visual perception of the event during the intake'. This would be different, if the quaestio were 'What happened at that time?'. Here, the constraints on the speaker's possible choices are much weaker; it would be up to him to indicate crucial features of his perspective. He might say, for example:

From where I stood, I had the impression that..., thus defining himself his place and the subjective commitment on the reality status of what is described.

3.4 Referential movement

Referential movement is the way in which information from the five domains of reference shifts from one utterance to the next. Suppose that there are two subsequent utterances A and B, in which the proposition consists of just one referent from each of the five domains. This gives the following picture of possible referential movements.

- (13) A: MOD_a TIME_a SPACE_a PERSON_a PREDICATE_a
B: MOD_b TIME_b SPACE_b PERSON_b PREDICATE_b

In principle, each referent in B can be maintained from A or freshly introduced in B. In reality, the possibilities are much finer; but in the present context, we will only illustrate the way in which the quaestio constrains referential movement (for a detailed discussion, see von Stutterheim 1997; Kohlmann 1997). If the quaestio is ‘What happened at that time?’, then MOD in general is defined as ‘real’; it remains constant throughout the text - unless the speaker explicitly chooses to deviate from it (see below section 3.5). As for TIME, the quaestio sets a temporal frame which the speaker is expected to deal with. Typically, he would subdivide this time frame into smaller time spans, which are sequentially ordered. Hence, TIME_b is *after* TIME_a. This is automatically given by the ‘topic condition’ on narratives mentioned above. Again, the speaker may deviate from this condition, thus giving rise to side structures or to special rhetorical effects. But if he remains within the frame defined by the quaestio, he is still free to choose the structural and lexical means by which to express this temporal shift. The quaestio does not impose any constraint on what the protagonists are. Hence, the speaker is free with respect to all three types of choice. He can introduce what he wants to, he can maintain what he wants to, in whichever form he wants to. *Space* is again not explicitly fixed by the quaestio: Note however, that in this case, a spatial frame may be implicitly included: the speaker is not supposed to speak about anything that might have happened at that time, but about what happened at that time at some particular place, and his freedom in introducing and maintaining spatial referents is limited to that place. As for the final domain *predicate*, all that is predetermined by the quaestio, is that it must be ‘happenings’, rather than, for example, visual properties or states. Otherwise, the speaker is completely free in what he chooses to introduce and to maintain in a particular utterance.

3.5 Deviating from the quaestio

In section 1, we introduced the idea that L-perspectivising can be described in terms of various choices which the speaker has to make in a given communicative situation - structural choice, lexical choice, contextual choice. In this section, it was discussed how the quaestio limits these choices and hence narrows down the ways in which the speaker can say what he intends to say. Hence, the quaestio is a principled constraint on possible perspectives; it is not the only one, and others will be discussed in section 4. It is up to the speaker to which extent he accepts the constraints defined by the quaestio. He may, to begin with, reject the entire communicative task, at the risk of more or less severe social consequences. He may also take on the task but redefine it in his own sense, for example by telling a long-winded story instead of giving an argument, when an argument was asked for. These are radical deviations from the quaestio. What is more interesting are „local deviations“; they occur when the speaker accepts the quaestio and its constraints in principle, but violates them from time to time. There are two such cases. First, the speaker might include a full utterance, or even a sequence of utterances, which is not an answer to the quaestio. This leads to what has been called ‘side-structures’. From a communicative point of view, they may be no less important than the main-structure utterances. In particular, they allow the speaker to express his personal views on what is told in answer to the question, and in this sense, they are important for L-perspectivation. The other case are small deviations within a main-structure utterance. The speaker may, for example, initially accept the position imposed on him by the question, but then present the entire story from a different vantage point. Normally, such changes must be explicitly marked. Suppose, for example, the question was: ‘What did you observe from your position?’ Then, the speaker is actually bound to a particular position, to a particular time and to a particular type of information he is allowed to report: the incidents he has observed. He then may start with a series of utterances which do exactly this, and then switch to something he has heard, or something he has observed later, or he may say something that he has not observed but simply inferred. Deviations of this sort give a particular flavour to the way in which the task is solved; they are, in other words, one of the speaker’s means to indicate a particular perspective. This does not devalidate the constraints as imposed by the quaestio: It is precisely the violation of these constraints which leads to the specific effect.

4. Other constraints on L-perspectivation

L-perspectivation is analysed here in terms of constraints on lexical, structural and contextual choices. Many of these constraints stem from the quaestio. But these are not the only ones. In the following we will discuss the potential sources for L-perspectivation which interact with the

constraints set up by the quaestio. These are (a) the hearer model, and (b) the linguistic system used.

4.1 *The hearer model*

In planning the amount and flow of information provided as an answer to a particular question, the speaker has to construct a mental model of the hearer (Graumann and Herrmann 1989, Herrmann 1989, Herrmann and Grabowski 1994, Levelt 1989). This model encompasses information and assumptions about the hearer's knowledge base with respect to the subject matter, his factual viewing point, his subjective perspective and expectations concerning the potential communicative content. The hearer model effects crucially influences the way in which the speaker constructs his text, with perspectivation coming into play at different levels. In the construction of a discourse representation, it crucially influences the two types of *selection*; it furthermore is largely responsible for the right balance between what is explicitly said and what is left to contextual information.

The most obvious reflection of the hearer model is the choice of a particular temporal or spatial referential frame. All languages have specific constructions which relate the information expressed to the time of speaking or hearing (*now*, specific tense forms), to the participants of the speech situation (*I, you, he*), and to the position of these participants within the speech situation (*here, there*, or, in a more complex way, *left, right, front, back*). It is the speaker who defines what, in a given situation, should count as the vantage point, the origo in Bühler's (1934) famous analysis of deixis. In the most elementary case, the speaker takes himself as the origo - in particular, his present position and the moment of his own speaking: he defines the perspective. But there are many complications. First, in written communication, it is not at all clear what should count as the moment of speech, or the position

of the speaker. These two vantage points must therefore often be explicitly introduced (say by giving date and place at the beginning of a letter). Second, even in spoken communication, the speaker is free, for example, to choose the position of the hearer as the spatial vantage point. This is often observed in instructions, where terms such as *turn left* or *put it on the blue block behind the red block* are normally seen from the hearer's, rather than the speaker's, viewpoint. And finally, the speaker may choose an entirely different 'origo' - for example, some time in the past and the position where he *was*, or where some other person was, at that time. Such a frame can be set by the quaestio. If the quaestio is 'What did you observe next?', then *left* and *right* relate not to the present position of speaker or hearer but to the position of the speaker at that time in the past. It is a 'Deixis am Phantasma', in Bühler's terms, and this quaestio which introduces it. Other quaestiones are even more complicated in this regard, for example 'How do I get from here to the station?'. In answer to this familiar question, virtually

all speakers choose the perspective of an ‘imaginary wanderer’, i.e., deictic terms in the route directions relate to the permanently changing fictitious position of a person who moves through the streets (cf. Klein 1982). In other cases, the quaestio itself fixes no origo at all (‘Are physical exercises healthy?’), and even if it does, the speaker may occasionally deviate from it. The flexible choice of vantage point is one of the crucial factors in L-perspectivation.

A second relevant factor in contextual choice is given by the assumed amount of shared knowledge. Here again, we can look back at a rather comprehensive research tradition (see the survey in Stutterheim and Kohlmann 1998). Studies on communication between interlocutors with different states of knowledge, such as experts and novices or adults and children, have shown that speakers construct different texts depending on what they take to be shared knowledge. Knowing more or less about a particular subject matter means that the potential maximal amount of knowledge is only presented in parts. As we have argued above, selective knowledge representation is always based on processes of perspectivation. An adaptation to the hearer’s perspective in this respect surfaces in different forms. It can be the determining factor for what is taken as the global level of granularity in a text. Usually, the quaestio does not predetermine at what level of specificity a speaker should provide information. In order to produce a consistent globally structured text the speaker *has to* take a choice with respect to the degree of granularity. One factor that governs this choice is the adaptation to the hearer. The scaffolding force of this globally established perspective can be seen in cases where the speaker deviates intentionally from this perspective. He is then led to use specific linguistic devices to mark the local nature of this change. Further evidence comes from cases where changes are required in the course of text production. If, for instance, the hearer requires a change of this parameter, the speaker cannot easily restructure his text without losing track altogether (cf. v. Stutterheim/Kohlmann 1998).

A third domain in which perspectivation is determined by the hearer model are constellations in which the quaestio leaves certain options unspecified, and the choice is led by what, in the speaker’s mind could make understanding easier for the hearer. Consider, for example, a quaestio which asks for the description of some highly complex configuration of entities. One of the speaker’s tasks is then to find an appropriate linearisation strategy. The speaker has different options how to connect the entities involved spatially: he can introduce an imaginary person, who walks from entity to entity or he can anchor the entities within some abstract referential frame by means of deictic or intrinsic relations. Which option a speaker will choose is – besides other factors such as the nature of the configuration described – dependent on the hearer model.

So far we have only looked at the role of the hearer model for one component of L-perspectivation - contextual choice. There are also

effects on lexical choice which can be traced back to this factor. The speakers may select a particular lexical item among equivalent ones, because they want to impress the hearer, but also because they feel that it is particularly appropriate for the hearer: they base their lexical choice not only on their own lexical repertoire but also on the alleged lexical repertoire of the hearer (E. Clark 1997). A particularly striking case is the adaptation to non-native hearers. Similar adaptations are possible but less often observed for structural choice - except again in communication with non-native interlocutors.

Summarizing we can say that the factor *hearer model* implies perspectivation along different dimensions some of which operate at one level with the globally established quest-constraints, some specifying options opened by the quaestio, some concerning local decisions such as lexical choice.

4.2 Linguistic system

When confronted with a particular subject matter and quaestio, speakers of different languages show different preferences for perspectivation of the communicative content (Carroll 1993, Slobin 1991, Talmy 1988). This is not surprising given that languages encode perspective at all levels of the linguistic system. Some of the linguistic means which require perspectivation for their use have been mentioned already. There are phonological means such as stress and intonation patterns which are applied on the basis of processes of perspectivation; morphological devices which encode perspectivity in relation to temporal, spatial and modal categories. Then there are syntactic structures which require perspectivation for their use such as word order in some languages, subordination or phrasal structures (prepositional phrases versus adverbials with respect to the explicitness of the reference object). Languages differ in what type of perspective is frozen in the system and how it is distributed across the range of the different devices. For processes of conceptualisation a relevant distinction is given by the fact that some of these devices are obligatory in nature (e.g. verbal morphology, word order) others can be chosen depending on the specific communicative intention.

This leads us to the following question. Is it the case that the respective linguistic system induces specific perspectives on the informational structure to be communicated? In order to pursue this question we have to look at texts which have been produced by speakers of different languages under conditions which otherwise remain constant. Studies of this type have been carried out for several languages. We will focus on the language pair English – German, the domain of interest lying on the differences between the two verbal systems.

A crucial contrast between the English and the German verb is related to the notion of aspect. English requires obligatory marking of aspect

whereas German has no such device. In English, this goes along with a rich repertoire of phasal verbs, such as to start, to keep on, to get to, which is extensively used. German has forms of this type in its lexicon, but the field is less differentiated, and phase markers are not high in frequency. Generally one can say that English provides a rich and structurally diverse repertoire for expressing the phasal structure of events. German, on the other hand, has no obligatory grammatical devices for this semantic category nor is the lexicon in this domain as rich as in English.

Let us now look at complex language productions by English and German speakers with respect to the temporal perspective under which events are presented and the function of the respective devices in context. The texts used as data base are film retellings elicited by a very general quaestio: 'Please tell me, what happened in the film?' As regards the level and type of event segmentation there are no constraints set by the quaestio. The speaker has to decide which event to select for verbalisation and how to present this event. Here we are only interested in temporal perspectivation. In order to describe temporal structures in texts we will draw upon the notion of *topic time* as developed in Klein (1994). Topic time is defined as the time interval for which an assertion is made. Whereas the event time, that is the time for which an event holds, is objectively fixed, the topic time selected in a given utterance is a matter of choice. We can therefore say that one central type of temporal perspectivation lies in the selection of a specific relation between topic time and event time.¹

This relation always has to be decided upon by the speaker which means that this type of perspectivation is part of what has been described as obligatory choices in constructing a complex informational structure. Given that a particular event is selected for verbalisation, what are the temporal perspectives under which the speaker can present it? He can either look at an event holistically with the topic time including the time of the event: *dann gräbt er ein Loch (e1) und da fällt er hinein (e2) (then he digs a hole and there he falls into)*. In e1 the topic time interval is given as the post time of some preceding event, marked by the temporal shifter *dann*, the event time lies completely within this time interval, the same pattern shows in the second utterance. Here the topic time is implicitly given as post time of e1 and the time of e2 lies completely within this interval. Typically this temporal perspective goes along with a certain aspect of completeness with respect to the event described. As in the example given above, both events are specified for a point of completion, in this case in the spatial domain. Another possibility for referring to completion of an event lies in referring to effected objects as in *er malt ein Haus*.

This holistic view of a situation can be contrasted with a view which segments events into phases. Under this perspective the topic time established overlaps with parts of the event time. The speaker can choose an inchoative aspect expressed in utterances such as *he starts to dig a hole* or *he starts digging a hole* with a further subtle distinction in perspective

concerning the degree of overlap between topic time and event time. Another option lies in an imperfective perspectivisation: *he is digging in the sand*. Here the topic time lies within the time of the situation. Note that the choice of this temporal perspective often goes along with presenting the event without a point of completion - in contrast to what has been said about a holistic perspective above. The last systematic option is given by the perfective aspect which places the topic time after the time of situation: *he has dug a hole*.

When we now look at how speakers of English and German speakers proceed in perspectivising events we find a systematic contrast. In the English texts an overall perspective is chosen which establishes a temporal viewing point as a deictic origo to which events are related in a very differentiated manner. Speakers present events segmented in their temporal phases each one hooked up to the origo rather than related intrinsically to each other. German speakers in contrast follow exactly the latter pattern by linking topic times to event times in the flow of the event chain, as demonstrated above. This implies a holistic or perfective view on the events since temporal boundaries of the event times are needed in order to function as a boundary for the topic time of the following event. This contrast in temporal perspective implies a number of further differences in conceptualising the content to be verbalised such as the explicit presentation of components such as spatial reference or effected objects versus their suppression in overt text.

Speakers of German and English clearly prefer one over the other pattern in the way outlined above. How can these differences be accounted for? Obviously there is an interrelation between the linguistic devices available for the speakers, in that the German speaker cannot draw upon an aspectual verbal system. On the other hand the interrelation between linguistic structure and information organisation in text cannot be that simple. German provides means to express phasal structures of events, only not morphological means but lexical forms. The interesting fact, however, is that speakers hardly make use of these forms because the global perspective chosen does not induce phasal segmentation! One possible explanation for this intricate interrelation between linguistic structure and perspectivisation could lie in the eminent role of obligatory linguistic categories for conceptualisation. Whereas the attention of a learner of English is drawn to aspectual properties of events (cf. Berman/Slobin 1994) this is not a prominent category for the German learning child. This in turn could lead to preferences in perspectivising situations for linguistic representation: the patterns we have observed for the adult speakers. These patterns at the level of conceptualisation can again induce language change in that those forms which are used very frequently, e.g. the phasal verbs in English, might in the end turn into new morphological markers.

5. Linguistic reflection of perspective management: the case of subordination

In this section, we will illustrate for one type of constructional device how the principles outlined above are operating in actual text production. This device is subordination. Since paratactic and hypotactic forms can be referentially equivalent, motivation to use a subordinated construction cannot come from the ‚substance‘ encoded, but reflects particular perspectual choices. Based on the results of comprehensive empirical studies (Carroll 1997, v. Stutterheim 1997) it will be shown that subordination serves one very general function: it is the most prominent means to maintain a globally established perspective by excluding information which deviates from this perspective from the main body of the text.

Semantically, subordination is a constructive device to exclude a proposition from assertion (Frege). This is also reflected in formal properties of hypotactic constructions which in many languages are reduced forms of main sentences as regards finiteness, subject realisation, etc. With respect to its function within textual structure one can say that subordinated information is – metaphorically spoken – backstaged a classical domain of perspectivation.

What is behind this metaphor will be shown for one central domain of perspectivation: the topic – focus structure as it is predetermined by the quaestio and specified by further globally functioning factors.

Let us take up what has been said about the global determination of topic-focus-structure above. The constraints set up for topic-focus partitioning of a text can basically be of two different types. On the one hand, they concern the determination of particular referential domains as parts of the topic or focus component including cases in which specific references are established as topic elements (*substantial topic constraints*). On the other hand, they determine specific structural properties of the text, here we speak about *topic conditions*. In what follows we will give examples for both types of constraints and their relevance for selecting subordination as an expressive device (cf. v. Stutterheim 1997).

Substantial topic constraints

One of the perspectives to be chosen in a narrative concerns the viewing point from which the events are presented. A well known choice is the one between narrator’s and protagonist’s perspective. Given that a story involves several important acting characters the speaker may select one as the protagonist or introduce several characters as being potential candidates for that role. Looking at film renarrations in which exactly this choice is opened up we find a very systematic pattern. If the speaker chooses a perspective which is anchored in relation to one protagonist, he will refer to events in which other characters take the role of the agent by means of

subordination, as in 14:

- (14) er beobachtet eine Frau, *die gerade erwischt wird, weil sie ein Brot geklaut hat*
(he observes a woman who is being caught because she has stolen a piece of bread)
- und sagt zu der Polizei, *die sie geschnappt hat...*
(and says to the police man, who has caught her..)

This examples shows that subordination is not only a backgrounding device in the sense of Hopper (1979), but it is also used for foregrounded events in order to keep the hearer's attention with the global topic candidate of the text.

Let us look at another case. Instructions answering a question of the type 'how has x to be done?' show a very clear global topic/focus distribution. In our particular case speakers have to instruct a hearer how to put building blocks together to get a particular object configuration. Here we can say that the objects to be manipulated belong to the topic domain, what has to be specified is the location of these objects. Spatial references therefore form the focus component of the text. A typical utterance in these texts would have the following form:

- (15) du steckst den lila Ring auf die grüne Schraube
(you put the purple ring onto the green screw)

Besides the information about the actions the speaker might have to specify the objects involved in certain cases. This can be required for the object to be manipulated as well as for the object functioning as reference object in the locational phrase. As regards the linguistic packaging we get a very systematic picture. Information specifying the affected object is more often paratactically integrated, whereas information specifying the reference object is *always* subordinated by a relative clause, e.g.:

- (16) du steckst die grüne Schraube durch den roten Würfel, *der über dem gelben sitzt...*
(you put the green screw through the red die, which is located above the yellow one)

We explain this absolutely regular pattern in language use across a large number of speakers again as resulting from global perspectivation. The speaker wants to maintain what has been established as focus component in the text. If he made the object reference expressed in this component topic element in its own rights by referring to it in subject position in a main

clause, he would change the globally established perspective. The use of a hypotactic construction in this case signals the hearer that this shift of perspective is only a local one without implications for the following text.

A parallel phenomenon can be observed in route directions and descriptions of spatial configurations. Here we get somewhat a mirror image of topic/focus distribution and consequently also a mirror image of what is presented in paratactic versus hypotactic form.

- (17) route direction
dann gehst du auf die Kirche zu, *die am Ende der Straße zu sehen ist...*
(then you approach the church, which can be seen at the end of the street)
- description
wenn man weiter runtergeht da steht rechts in der Mitte eine Kirche...
(if you walk down there is a church to the right in the middle)

Depending on the constraints set up by the quaestio the speaker has to provide different types of information in main structure utterances. The route direction require information on paths with spatial reference forming the focus component. Further specification of objects introduced as reference objects implies – just as in the case of the instructions outlined above – changing the topic-focus pattern established globally. Descriptions, on the other hand, require information on relative object positions. If the speaker chooses to draw upon an imaginary wanderer then this introduces a dynamic perspective which forms a contrast to the global perspective established by the quaestio.

The use of hypotactic forms reflects exactly this difference between the two text types. In the route directions, object specification in the form of relative clauses is systematically subordinated, in descriptions reference to the motion events is almost exceptionless subordinated. Again, we see the function of this syntactic device in maintaining the global perspective by excluding information constructed from a different perspective from the chain of asserted propositions.

Structural topic constraints

As has been explained above, the linearisation principle underlying the sequential ordering of information in text production can be interpreted as global topic condition. This, in turn, implies a particular perspective under which the single components of the overall referential structure are related to each other.

Looking again at empirical data, we find that subordination serves

the function to integrate information which violates this global constraint. Taking texts which follow a spatial linearisation principle, e.g. descriptions, subordination is used where the speaker locally deviates from this pattern. If, for instance, he switches to an object-oriented linearisation strategy - often linked to a change of the level of granularity – then this is expressed by use of hypotactic forms.

- (18) daneben ist ein Cafe, *das über einem Schuhladen ist und dicke rote Vorhänge hat...*
(next to it is a cafe, which is above a shoe shop and has thick red curtains)

A parallel pattern can be observed in temporally linearised texts, which follow a chronological order. Violation of the principle of temporal sequencing is also frequently „put aside“ by means of subordination.

Another structural aspect which is subject to perspectivation is the level of granularity chosen at text level. Here the same pattern emerges across different text types. Speaker subordinate information which imply a local change of the level of granularity. Again we cannot say that information of this kind is part of background information, but it is backstaged for reasons of perspective – continuity.

To summarise what has been found about the functional motivation of subordination we come to the following conclusions. In order to construct a consistent information structure as an answer to a quaestio the speaker has to establish and maintain certain parameters which allow him to locally proceed in selecting and contouring the single informational units. As has been argued, taken perspectives is one crucial component in this planning process. It has also been made clear that these globally set perspectives function as default values, which means that the speaker is free to deviate from them. In order to meet both demands – maintainance of global perspectives and integration of alternative perspectives – the speaker can draw upon means which openly mark something like a hierarchy of perspectives. Subordination is one of these devices. It prevents the hearer from being led to global changes in perspective where only a local detour is intended. This function is further supported by the fact that information which completely leaves the referential frame of the text usually is not presented in subordinated form (e.g. comments, explanations). Here there is in a way no competition between the newly introduced perspective and the globally established one. Only if the grounds are set to possibly take another point of view on the subject matter the speaker has to make explicit which line he intends to follow by using the respective devices.

6. Conclusions

Not many concepts in the social sciences are so appropriate for metaphorical extension than 'perspective'. We do not want to belittle the heuristic, and even the explanatory, value of such extensions from visual perception to, for example, the way in which certain facts are evaluated by people of varying social background ('from the perspective of a street worker', 'from the perspective of an unemployed widow in Calcutta'). Quite to the opposite: it is not accidental that we feel these extensions of the term to be so suggestive. But at one point, it must be made clear, what is specific to 'perspective' and 'perspectivation' in a particular field. In the paper, we have tried to do this for L-perspectivation - for the way in which natural language allows its speakers to express a certain perspective. It was argued that this is essentially a matter of three choices which the speaker must permanently make in the course of his production process: lexical choice, structural choice, contextual choice. All three choices, and hence the speaker's L-perspectivation, is constrained by a number of interacting factors. The most important of these is the 'quaestio' which the speaker sets out to answer. Others include adaptation to the hearer and the specific possibilities of the language in question. This list is surely not exhaustive. We only looked in passing at the role of the speaker's subjective attitude, because it may lead to 'side-structures' in his text. But subjective attitudes also influence the lexical choice independent of the particular quaestio, and it may well be that this influence follows certain principles. But their investigation is primarily a matter of social and emotional psychology, rather than of linguistics. It is the linguistic constraints on L-perspectivation that were discussed here; these constraints are often strong; but they do not determine the speaker's choice.

Literature

Bierwisch, M. & R. Schreuder (1992). From concepts to lexical items.

Cognition, 42, 23-60.

Bühler, K. (1934). *Sprachtheorie. Die Darstellungsfunktion der Sprache*.

Jena: Fischer.

Canesius, P. (Eds.) (1987). *Perspektivität in Sprache und Text*. Bochum:

Brockmeyer.

Comrie, B. (1976). *Aspect*. Cambridge: Cambridge University Press.

- Carroll, M. (1993). Deictic and intrinsic orientation in spatial descriptions: A comparison between English and German. In J. Altarriba (Ed.), *Cognition and Culture: A Cross-Cultural Approach to Cognitive Psychology*. Amsterdam: Elsevier.
- Carroll, M. (1997). Information organisation and direction of attention when structuring space in English and German. In J. Nuyts / E. Pederson (Eds.). *Language and conceptualisation*. Cambridge: Cambridge University Press.
- Clark, E. V. (1997). Conceptual perspective and lexical choice in acquisition. *Cognition*, 64, 1-37.
- Graumann, C.F. (1989). Perspective setting and taking in verbal interaction. In R. Dietrich & C. F. Graumann (Eds.). *Language processing in social context* (95-122). Amsterdam: North Holland.
- Graumann, C. F. & T. Herrmann (1989). *Speaker: The role of the listener*. Clevedon: Multilingual Matters.
- Herrmann T. (1989). Sprachpsychologische Beiträge zur Partnerbezogenheit des Sprechens. In H. Scherer (Ed.) *Sprache in Situation. Eine Zwischenbilanz* (179-204). Bonn: Romanistischer Verlag.
- Herrmann, T. & J. Grabowski (1994). *Sprechen. Psychologie der Sprachproduktion*. Heidelberg: Spektrum Akademischer Verlag.
- Hopper, P. (1979). Aspect and foregrounding in discourse. In T. Givón (Ed.), *Discourse and Syntax* (Syntax and Semantics 12) (213-241). London: Academic Press.
- Klein, W. (1982). Local deixis in route directions. In: R. J. Jarvella & W.

- Klein (Eds.), *Speech, Place and Action* (161-182). Chichester: Wiley.
- Klein, W. (1994). *Time in Language*. London: Routledge.
- Klein, W. & C. v. Sutterheim (1992). Textstruktur und referentielle Bewegung. *Zeitschrift für Literaturwissenschaft und Linguistik*, 86, 67-92.
- Kohlmann, U. (1997). *Objektreferenzen im Kontext. Eine Untersuchung zum Zusammenhang von kommunikativer Aufgabe, Textstruktur und Objektreferenzen*. Frankfurt/M.: Lang.
- Levelt, W.J.M. (1982). Cognitive styles in the use of spatial direction terms. In R. J. Jarvella & W.Klein (Eds.) *Speech, place, and action* (251-270). Chichester: Wiley.
- Levelt, W.J.M. (1989). *Speaking. From intention to articulation*. Cambridge: MIT Press.
- Levinson, S. (1997). From outer to inner space: linguistic categories and non-linguistic thinking. In J. Nuyts & E. Pedersen (Eds.), *Language and conceptualisation* (13-45). Cambridge: Cambridge University Press.
- Mikame, H. (1996). Markierte Perspektive, perspektivische Annäherung des Sprechers an das Objekt und direkte Wahrnehmung. *Sprachwissenschaft*, 21, 367-420.
- Sandig, B. (1996). Sprachliche Perspektivierung und perspektivierende Stile. *Zeitschrift für Literaturwissenschaft und Linguistik*, 26, 102, 36-63.
- Slobin, D.I. (1991). Learning to think for speaking. Native language, cognition and rhetorical style. *Pragmatics*, 1, 7-26.
- v. Sutterheim, C. (1997). *Einige Prinzipien des Textaufbaus. Empirische*

Untersuchungen zur Produktion mündlicher Texte. (Reihe Germanistische Linguistik 184). Tübingen: Niemeyer.

v. Sutterheim, C. & W. Klein (1989). Textstructure and referential movement. In R. Dietrich & C.F.Graumann (Eds.). *Language processing in social context* (39-76). Amsterdam: North Holland.

v. Sutterheim, C. & U. Kohlmann (1998). Selective hearer-adaptation. *Linguistics* 36-3, 517-549.

v. Sutterheim, C. & M. Lambert (to appear). Event structure. In H. Hendrix (Ed.). *The Structure of learner varieties.*

Talmy, L. (1988). The relation of grammar to cognition. In B. Rudzka-Ostyn (Ed.), *Topics in Cognitive Linguistics* (165-205). Amsterdam: Benjamins.

Tversky, B., H.A. Taylor & S. Mainwaring (1997). Language et perspective spatiale. In M. Denis (Ed.), *Langage et cognition spatiale* (25-50). Paris: Masson.

Notes

1.

Another possibility for establishing a temporal perspective lies in the choice of an origo for anchoring the referential frame. In the unmarked case this is given by the deictic origo, the time of utterance, but as already described in Bühler (1934) the origo can be shifted along different dimensions (cf. in detail Klein 1994).