Processes of conceptualization in language production: language-specific perspectives and event construal[1]

CHRISTIANE VON STUTTERHEIM and RALF NÜSE

Abstract

Differences in how verbal tasks are solved by speakers of different languages are often attributed to specific cultural and stylistic traditions. In contrast to this position we argue that differences in the organization of information in texts are rooted in structural contrasts between languages. This claim is based on production data from speakers of English, German, and Algerian Arabic which include retellings of a silent film and verbalizations of short video clips. The domains of analysis are patterns of event formation and temporal viewpoint. The data obtained from speakers of the three languages show contrasts with respect to three aspects of information organization: (a) speakers segment the flow of information as presented in the visual input on the basis of different criteria, (b) they systematically select different components when representing a given event in language and (c) they establish different temporal perspectives when anchoring and linking events in discourse. In this paper, we argue that the contrasts observed in selecting and structuring information can be generalized by assuming language contrasts at the level of abstract principles of perspective taking. These principles are in turn rooted in language-specific patterns of grammaticization.

1. The problem

There is general agreement as to the fundamental components which are constitutive for the language production apparatus. Conceptualization, formulation and articulation, this three-partitive model — Levelt's blueprint for the speaker (1989), has been the referential frame for most of the work on language production. Following several years of intensive research in the field of language production, insights have been gained into the two components that involve linguistic structures: the formulator

Linguistics 41–5 (2003), 851–881 0024–3949/03/0041–0851

and the articulator. The conceptualizer, however, is still more or less terra incognita. It is not only that we are unable to specify the nature of the processes involved, we do not know what the format and units are in which to model a LANGUAGE OF THOUGHT.

One of the central questions related to this problem concerns the role of language in this component, and here we mean the role of the specific language used in the production process. Current models of language production are obviously not very explicit about the nature of the conceptualizer, and there are a number of diverging hypotheses on the point in question, the role of language. Three positions can be distinguished:

The radical position 1: processes in the conceptualizer are language-free, universal, and operate on the basis of conceptual primitives. This is the view put forward, for instance, in the approach to language and concepts by Jackendoff (1990) and related work on production by Bierwisch and Schreuder (1992).

The radical position 2: processes in the conceptualizer are language-based in nature (Whorfian view). This position is advocated in recent work by Lucy (1992, 1996) and Levinson and his group (Levinson 1997).

The moderate position: there is an interdependence between conceptualization and linguistic knowledge with two possible assumptions:

a. The preverbal message, as the result of the conceptual planning process, is language-oriented. This means that a reorganization of the conceptual material takes place at the end of the planning process, thereby shaping the message according to language-specific requirements. Levelt (1989, 1996) assumes that these processes do not affect the WHAT or content of the message. This belongs to macrostructural planning. According to his model, language specificity concerns only HOW content is packaged for verbalization. Language-specific planning occurs therefore only at the microstructural level.

b. The THINKING-FOR-SPEAKING hypothesis, which argues that conceptualization as a component of language production is always based on language-specific principles. The level of specificity of these principles remains an open question (Talmy 1988; Slobin 1996a, 1996b). Important for the line of argumentation in the present paper is that under this view language-dependent conceptualization is relevant at the global (and local) level of message generation.

In order to shed some light on the phenomena at issue, ways of looking into the conceptualizer have to be found, since the three positions can only be judged in their explanatory value on the basis of empirical findings. In the following we will first discuss some basic issues related to the functions of the conceptualizer and then present empirical results on processes of conceptualization that are based on different methodolo-

gies. The results will be interpreted with respect to their implications for the discussion on language specificity at the level of conceptualization.

2. The window on conceptualization

What do we know about the nature and function of the conceptualizer in language production? Very generally speaking, the conceptualizer comes into play when a communicative intention is formed by the speaker. It then has to draw upon a knowledge base represented in whatever format (Engelkamp 1990) and to transform parts of it into a structure with which linguistic knowledge at different levels (syntax, lexicon, morphology, phonology) can be accessed. This structure has been called *preverbal message* (Levelt 1989), *input for encoding* (Herrmann and Grabowski 1994) or *temporary conceptual structure* (Carroll and von Stutterheim 1993). The range of terminology found with respect to this level of planning reflects the lack of clarity on the substance matter.

On the path from knowledge activation to an expressible format at the level of the preverbal message, conceptual material is built up that has been subject to a number of planning processes (cf. Habel and Tappe 1999):

1. SEGMENTATION. Units have to be extracted from a knowledge base that is neither structured with respect to sequence nor organized at a specific level into what can be seen as a hierarchically structured body of knowledge about a given substance matter, which is more or less specific. Complex static situations, for instance, have to be broken down into a number of states or property predications, or complex dynamic situations have to be divided into events or processes (cf. the notion of atomic formula in semantics [Parsons 1990]).
2. SELECTION. The speaker has to select those units which he wants to verbalize as well as the components by which these units can be represented. By components we mean the conceptual building blocks, such as entities, spaces, times, properties, or actions, from which propositional units can be formed (cf. the notion of subatomic structure [Parsons 1990]).
3. STRUCTURING. The components selected have to be structured under several aspects that cover options with respect to predicate types and argument roles (e.g. *sell*/*buy* alternative), anchoring within a referential frame (e.g. spatial and temporal anchoring), and attribution of informational status (topic/focus assignment). All these steps in the planning process are perspective-driven (cf. von Stutterheim and Klein 2002).

4. Linearization. The units selected for verbal representation have to be linearized in order to be transformed into the one-dimensional medium language (cf. Levelt 1982).

When planning and producing a complex piece of discourse, these processes operate on different levels, whereby two levels are distinguished in the literature on language production: macrostructural or global planning and microstructural or local planning, where the latter is embedded in and dependent on the first. What Hopper (1995) has pointed out very clearly for structural descriptions of language is just as relevant for research in language production, where most of the work has been concerned with linguistic units at and below sentence level (cf. the survey in Rickheit and Strohner 1993):

> To start with discourse is not simply to use discourse data for the same projects that were previously carried out with invented sentences, but to face the prospect of beginning anew and finding new generalisations about language, based on entire texts and contexts rather than on convenient selections (Hopper 1995: 140).

This view underlies the study at hand. It means that contextually embedded language (text) production has to be analyzed in order to get a hold on processes of conceptualization. On this basis, hypotheses can be developed that can be tested in carrying out more fine-grained experimental work. This is the path taken in the present study.

The properties of the conceptualizer as mentioned so far are too general to be subjected to empirical research. In order to gain access to the processes mentioned, we have to find a window that opens the view onto a specific describable substance. It has to be a window that allows us to pin down the objective perceivable input and its linguistic correlate at both ends of the conceptualization process.

We selected the domain of events as one of the possible windows on the interrelation between perception, cognition and language. The notion of *event* has been treated extensively in different theoretical frameworks both in and outside linguistics. There is philosophical work on *eventhood* (e.g. Kamp 1981; Parsons 1990), psychological work on event cognition (e.g. Newtson 1976), and then we find the notion of "event" as one of the central categories in different linguistic fields, in particular in (logical) semantics and cognitive linguistics. This is not the place to get involved in an in-depth discussion of the notion "event" (cf. Klein 1994 for a critical discussion), so we will only briefly introduce the central theoretical distinctions necessary in keeping categories for empirical analysis clear-cut.

There are three levels to be distinguished:

a. the external world.

b. a partial conceptual representation of the external world. Note that these conceptual representations are dynamic in nature, which means that they can be subjected to processes of reorganization.

c. linguistic representations (predicate–argument structures).

Most work within the field of (logical) semantics and philosophy of language only takes the two levels (a) and (c) into consideration. We quote Parsons as a representative of this view: "I seek a theory that describes the semantics of sentences of English, that is the relations between words of our language and things in the world" (1990: 11).

Approaches that focus on psycholinguistic aspects of production, comprehension, and acquisition of language are often concerned with the levels (b) and (c) only. This distinction is extremely difficult to operationalize in empirical analysis, a problem that becomes evident from the inconsistent use of the theoretical categories in the literature.

In many psycholinguistic and semantic studies, the distinction of levels (b) and (c) is not really necessary, since categories at the level of conceptualization are not operationalized independently of the units at the semantic level. Addressing the question of the role of language in processes of conceptualization — as in the study at hand — requires the separation of the two levels, however. We will therefore introduce the following working definition for the relevant units at the different levels.

a. SITUATIONS for what takes place in the external world.

b. EVENTS an event is a self-contained segment in a conceptual representation of a network of interrelated situations, conceptualized as a time–substance relation. The substance constituting an event is characterized by the features DYNAMIC and POTENTIALLY BOUNDED (cf. Croft 1991).

c. linguistic representation.

Given a particular situation in the external world, there are options with respect to the cognitive representation of this situation as an event, and there are again options in presenting an event linguistically. These options concern first of all the level of granularity, furthermore the components selected (e.g. bounded/unbounded) for representing this particular situation, as well as the perspective under which it is viewed. In the course of language production, all these decisions have to be taken before the linguistic form is activated. If we assume an "unintelligent formulator" — a position which has not been questioned — then all information that is relevant for the selection of lexical and syntactic form has to be provided by the conceptualizer.

Let us illustrate the requirements for the conceptualizer by an example from the domain of events: a speaker wants to talk about a situation, for example, the delivery of a parcel, and produces the following text: *The postman rang, I opened the door and he gave me a small parcel*. Three

events are reported in temporal sequence. As regards SEGMENTATION, the speaker could also produce a chain such as *the postman came and gave me a small parcel*, or present the overall situation as one event: *the postman delivered a small parcel.* He could also be more specific, for instance, *the door bell rang, I went downstairs and opened the door, the postman had come early, he opened his big bag, took out a small parcel and gave it to me* … The differences shown in these short texts are referred to as differences with respect to the level of granularity. Processes of SELECTION concern the components by which the single events are represented. Instead of *the postman rang* the speaker could have produced *the postman rang at my door at 9 o'clock* providing more information about the situation. The processes we have called STRUCTURING are manifold. We will mention a few aspects which illustrate the impact of perspectivation on message construction. The parcel-delivery event could be presented from the viewing point of the recipient leading to a different lexical choice and to different argument structures: *I received a small parcel from him.*[2] Spatial anchoring could be of a different nature, as shown in the following reformulation of the text: *the postman went over to the house, he rang the bell, a woman opened and he gave her a parcel*. The last aspect to be mentioned in this brief overview concerns the necessity of deciding on a specific information structure. If the speaker chooses the option of constructing a RECEIVE event in contrast to a GIVE event he changes the informational status of the participants at the same time.

Given all these options, what are the criteria that enable the speaker to come up with consistent patterns of information organization? While we find studies on situational factors (e.g. attention [Tomlin 1997]), type of subject matter (e.g. object reference [Herrmann and Deutsch 1976]), and contextual factors (e.g. implicit information [Brown and Dell 1987]), language has not been considered as a factor in structuring processes within the conceptualizer. This is the goal of the present enterprise. The question underlying this study can now be formulated more specifically. TO WHAT EXTENT ARE THE PLANNING PROCESSES REQUIRED FOR THE CONSTRUCTION OF AN INFORMATIONAL NETWORK EXPRESSED IN A PIECE OF DISCOURSE RELATED TO STRUCTURAL PROPERTIES OF THE SPECIFIC LANGUAGE USED? In addressing this question we will proceed in three steps. In the main section of the paper, we will present empirical evidence from a comparative analysis of English and German, taking both texts and single utterances. The results show that language structure can be correlated with specific patterns of event conceptualization. Then possible causes for these differences will be discussed, thereby extending the data base to Algerian Arabic. In the last section, we will touch on implications of the findings for models of language production.

3. Empirical studies

3.1. *The segmentation of event sequences*

Background. The first study concerns what we have called the segmentation problem, that is, the question whether English and German speakers break down a stream of dynamic situations according to the same principles or not. A film was selected as stimulus material in order to elicit the verbalization of event sequences within a global frame of reference. The film used was *Quest*, a silent film, seven minutes long, which is structured very clearly into five episodes.[3] In previous studies, retellings from subjects with different languages (including Spanish, French, and Italian) were recorded under conditions such as offline retelling (data type: Quest 1) (cf. Carroll and von Stutterheim 2003; von Stutterheim and Lambert i.p.), past time framing (cf. von Stutterheim et al. 2003). For the data underlying the present study on segmentation issues, however, subjects had to describe the movie ONLINE, and the languages compared are English and German (data type: Quest 3). In particular we looked at the level of granularity with which events are encoded in these two languages. In other words: since the question at issue is whether English and German speakers parse a dynamic stream of situations according to the same principles or not, we looked at whether English and German speakers encode the same number and same type of events when retelling the film.

Method. Seventeen English and seventeen German subjects saw the film *Quest* and were asked to retell the content while watching the film. None of the subjects knew the movie or had seen it before. The interlocutor could not see the film, and they were asked to tell the interlocutor *''what is happening/was passiert?''*

The linguistic material obtained was audiotaped, and the audiotaped data were transcribed and segmented into propositional units corresponding in most cases to the formal unit of a clause. Texts vary considerably in length, ranging from 96 units to 334.

The correlates of the propositional units at the conceptual level are either events — in the sense defined above — or states (including descriptions of the five different landscapes of the movie; see footnote 3). For the quantitative analysis of segmentation, the number of propositional units referring to events was counted. Utterances referring to states, such as *so he's now on top of the rock*, were not included in this analysis, nor were events that are presented as components of another event and thereby lose their status as temporally bounded entities (cf. Parsons 1990:

6). An example of an integrated event of this kind is given in the utterance *he's looking for a way to try and get down off the rock*. This utterance consists of two propositional units, but encodes only one event.

Results. In order to give a first impression of the difference between the English and German film retellings, we will first cite some typical examples of the texts obtained. In the scene selected for illustration, the protagonist — the little clay figure — finds himself on a pillar of rocks, climbs down, and falls to the ground. A typical English and a typical German solution of the task runs as follows:

(1) English
001 and one rock has just come out of the ground
002 and it's lifted him up
so he's now on top of the rock
003 and he's looking around him
004 he's scratching his head
005 and climbs over to the rock
006 he's looking down
he's a good ten feet up in the air
007 he's blinking again
008 he's kneeling on the rock
009 and he's looking just around him
010 he sees the water again
011 and he gets excited
012 so he's looking for a way
to try and get down off the rock
013 okay he's swinging from the rock now
014 he's trying to get down carefully
015 but the rock gives way
016 and he lands on the ground

(2) German
001 jetzt wird er von so einer Steinsäule hochgehoben
'now he is lifted up by some kind of stone pillar'
002 und er wundert sich
'and he wonders'
wie er da wieder runterkommen soll
'how he should get back down (from) there'
003 jetzt hat er auch wieder Wasser entdeckt
'now he has again discovered water'
004 und versucht nun
'and he tries now'
von dieser Steinsäule runterzuklettern
'to climb down from this stone pillar'

005 er hat einige Schwierigkeiten dabei
'he has some difficulties with that'
006 und fällt schließlich auf den Boden
'and finally falls onto the ground'

There are a number of differences between these two kinds of texts: English speakers seem to speak at a certain pace, as if they wanted to report what they see happening every two seconds or so; German speakers, on the other hand, seem to summarize what is happening with regard to certain outcomes or intentions of the protagonist (some of this will be described in more detail below in section 3.3). For the current discussion regarding segmentation issues, however, we want to focus on just one quantitative aspect: English speakers mention more events than German speakers. As can be seen from the numbers given in the quoted texts, the English speaker mentions sixteen events, while the German speaker mentions only six events. Although the two texts have to be seen as extreme examples, they stand for a general tendency that shows up in all texts produced by the 34 subjects: for this scene, German speakers mention on average 9.17 events, while English speakers mention 12.23 events ($F_{(1,33)} = 6.85$, $p = 0.013$). The same difference can be found for other scenes, and of course it also shows up when one counts the number of mentioned events in the whole film (German speakers mention overall 159.24 events, English 212.18; $F_{(1,33)} = 8.53$, $p = 0.006$). We can summarize the findings in saying that in the process of language production, English speakers, when confronted with the same visual input as German speakers, parse the flow of information into more units than German speakers.

Moreover, English speakers mention a number of "small" events that German speakers do not refer to at all. In the text quoted above, for example, the English speaker refers to events like "he is scratching his head," "he is looking down," or "he's swinging from the rock." These are not mentioned by the German speaker, maybe because they are assumed to be implicitly given by reference to complex events such as "he wonders how he should get back down from there," or because German speakers simply do not find these events worth mentioning. The same is true in other scenes and with other "small" events like "look around," "get up," or "be walking." Here again, these events are either covered by macro events or simply not mentioned in the German texts. So another way to pinpoint the differences between English and German texts is to count how often these "small events," which are quite frequent in the movie, are mentioned by the English and German subjects. Table 1 gives the results:

Table 1. *Mean number of "small events" mentioned (of the type listed in the first column) per subject (N = 34)*

	German	English	F	p
look up/down/around (*nach oben/unten gucken, umschauen, umgucken*)	6.6	13.1	$F_{(1,33)} = 5.89$	$p = 0.021$
get/sit up (*aufstehen, hochrappeln etc.*)	1.9	3.9	$F_{(1,33)} = 7.08$	$p = 0.012$
walking (*[los] laufen, gehen etc.*)	2.7	4.5	$F_{(1,33)} = 4.38$	$p = 0.044$

Table 1 reads as follows: events of the type "look up/down/around" (which, for our purposes, correspond roughly to the occurrences of the verbs *look up*, *look down*, and *look around* — or that sample of German verbs mentioned in the table) were on average mentioned 6.6 times by German subjects and 13.1 times by English subjects. This difference is statistically significant at the 2%-level. So the overall result shown by the table is this: in all types of events counted (namely "look up/down/around," "get up," and "is walking"), English speakers mention these events twice as often as (or, for that matter: more often than) German speakers.

Both comparisons — the total number of events selected for verbalization and the number of mentioned "small events" — allow the following conclusion: English speakers verbalize more events than German speakers and, thus, parse the same event sequences into smaller units than German speakers do. This means that English and German speakers encode different events given identical visual input. This result has to be borne in mind when it comes to the discussion of the different positions in the literature mentioned in the first section (see below).

3.2. *The selection of event components*

Background. Given the fact that English and German speakers segment the flow of visual input according to different principles in language production we next wanted to see whether there are also differences to be observed when speakers of the two languages verbalize the same single event. This concerns what we have called the selection problem (see section 2), that is, the question as to which components are selected in representing a specific event verbally.

What are the possible options in this respect? A first impression is given by the film retellings described in the preceding paragraph. In these texts, there was, for example, a difference in the way FALLING and JUMPING events (which also are quite frequent in the film) were linguistically packaged. In the encoding of these events, English speakers mainly use bare verbs, as in cases like *he runs and jumps* or *he slips and falls*. Only rarely (in 30% of the cases) do they mention the endpoint or the direction of "the falling" or "the jumping" (as, e.g., in *he falls down/he falls onto the ground* or *he jumps down/he jumps into the hole*). This is exactly the opposite in German. In German one rarely finds bare verb constructions like *er fällt* or *er springt*, and speakers almost always specify the direction or the endpoint of the falling event (*er fällt runter/fällt hin/fällt auf den Boden*). This difference is not confined to motion events. Activities where some object is affected or created are also represented in different components by speakers of the two languages. A situation in which the protagonist makes a hole in the sand, for instance, is verbalized as an open-ended activity by speakers of English (e.g. *he is digging in the sand*); German speakers prefer to "close" the event by referring to the result, in this case an effected object, such as in *er gräbt ein Loch im Sand* 'he digs a hole in the sand'.

Given these findings, we wanted to find out whether there is a systematic difference between speakers of German and English in the selection of event components for coding. The results from the film retelling revealed two options with respect to the perspectives underlying and structuring the selection processes. Under one perspective, events are viewed as ongoing, and the endpoint or outcome of an action does not form a necessary component of each event. This would constitute the preferred option for speakers of English. The other perspective looks at the action holistically, a view which includes some kind of closure to an activity. This is the viewpoint underlying the German verbalization.

In order to investigate this hypothesis more systematically, a series of computer animations as stimuli for a production experiment were designed showing single situations. The aim was to see whether the elicited verbalization differed significantly for speakers of German and English with respect to the components selected.

Method. Twenty English and twenty German subjects saw sixteen computer animations that depicted a variety of situations such as "someone eating a banana," "someone jumping off a cliff into the water," "a boat sinking to the bottom of the sea." Subjects were asked to verbalize *what is happening* as soon as they could identify the type of situation. Their utterances were tape-recorded and categorized by three independent

observers as to whether the event was closed or not. The category "closed event" included those which contained reference to an endpoint of a motion event as well as those where products of activities were mentioned.

Eight of the sixteen items were expected to elicit differences with regard to this contrast, while the other eight were either fillers or items where no differences with regard to the mentioning of endpoints were expected. These were items where both English and German subjects were expected to select the same components for verbalization of the situation. In one case, it was an ongoing activity like "someone is sleeping," in the other case there was a conceptually required endpoint such as in "someone putting a book on a shelf," where both English and German subjects presumably do mention the endpoint.

Results. With respect to items of the last group there were no differences to be found, that is, items like "someone is sleeping," "someone is eating a banana," and "someone is putting a book on a shelf." The eight test items proper are shown in Table 2.

As Table 2 shows, there are different distributions for different items. Some items elicit more closures across both groups (up to 50% of the English speakers), as in items like "someone painting vs. someone painting a picture" or "someone jumping off a cliff vs. someone jumping into the water," while other items overall tend to elicit less closures (only 50% of the German speakers doing so), as in cases like "someone walking along the street vs. going to the bus stop" or "someone hitting/chiseling a stone vs. someone making a sculpture." In each case, however, German speakers mention more endpoints than English speakers. If only 50% of the German speakers "close" the event, English speakers do so even less frequently; and if 50% of the English speakers mention the endpoint of the event in question, the amount of German speakers who do so is 90%. Correspondingly, if the number of mentioned endpoints per subject over all eight test items is counted, German speakers mention the endpoint 5.75 times on average, while English speakers mention the endpoint 1.80 times on average. This difference is highly significant ($F_{(1,39)} = 104.11$, $p < 0.0001$).

Discussion. The results show that German and English speakers indeed select different event components for encoding: English speakers prefer to mention a phasal segment of an event, in the given case mainly the ongoing activity, while German speakers (additionally) mention the endpoint of an event. It should be noted that the computer animations actually did not show the endpoint of the situation in question, that is, it was not shown that the protagonist actually GETS to the endpoint at

Table 2. *Distributions and Chi-square values for the eight test items in the description of single situations (computer animations)*

Item	English		German		χ^2	df	p
	endpoint mentioned	endpoint not mentioned	endpoint mentioned	endpoint not mentioned			
the boat is sinking vs. the boat is sinking to the bottom (… *auf den Grund*)	4	15	16	4	$\chi^2 = 13.55$	df = 1	$p < 0.0003$[4]
someone sawing vs. sawing a log in half (*jemand sägt … durch/zersägt einen …*)	4	16	16	4	$\chi^2 = 14.40$	df = 1	$p < 0.0002$
the paper is being blown off the table vs. falling onto the ground	3	17	13	7	$\chi^2 = 10.42$	df = 1	$p < 0.002$
someone is digging in the sand vs. digging a hole/building a sandcastle	5	15	13	7	$\chi^2 = 6.47$	df = 1	$p < 0.011$
someone jumping off a cliff vs. jumping into the water	8	12	16	4	$\chi^2 = 6.66$	df = 1	$p < 0.010$
someone is painting vs. is painting a picture	10	10	17	2	$\chi^2 = 7.13$	df = 1	$p < 0.008$
someone is hitting/chiseling a stone vs. tries to break a stone/is making a sculpture	2	18	12	8	$\chi^2 = 10.99$	df = 1	$p < 0.0001$
someone is walking along the street vs. is going to the bus stop	4	16	11	9	$\chi^2 = 5.23$	df = 1	$p < 0.023$

issue. This is especially true in those cases where the situation in question involved a change in place. In these cases, the computer animation stopped before the moving object actually reached the endpoint of the change in place. So the diver did not reach the water in the case of the jumping event, for example, nor did the sinking boat and the floating paper actually get to the ground. Together with those cases where a goal-directed event like ''building a sand castle'' is at issue, this means that German speakers in a sense construct and supplement the endpoint in question, while English speakers only verbalize what is actually shown. For the moment, it remains an open question whether English speakers also would refrain from mentioning the endpoint of a motion event if the stimulus showed that this endpoint is reached. Summarizing the findings from the studies on event segmentation and information selection in language-production, one can say that English and German speakers not only encode different events, but also encode events differently.

3.3. *Structuring information*

The role of perspective. The results show that speakers of English and German differ both in the construal of events as well as in the components selected for expression. In the retellings, speakers of English select a higher degree of resolution when segmenting dynamic events, compared to German speakers, and focus in their verbalizations on different components of individual events. The events are often coded in the progressive, the form which encodes ongoingness. However, this form in itself does not prevent a speaker from focusing on endpoints, as shown in the following example: *the boat is sinking* TO THE BOTTOM. Similarly, speakers of German need not specify events as bounded entities. The syntactic frame of a verb such as *fallen* does not strictly require an adjunct in the form of a locational phrase.

Explanations for the observed differences have to go beyond the level of form as such and answer what form means in *perspective taking* (cf. von Stutterheim and Klein 2002; von Stutterheim and Lambert i.p.). In the following, we will argue that the specific patterns found in the verbal output are created by the speakers on the basis of language-specific principles of information organization. They are not just alternative manifestations of one and the same underlying conceptual structure, but rather construals of worlds in which universal conceptual categories are interrelated and weighed along different dimensions. We claim that the relevant knowledge consists of language-specific principles of information organization that constrain the options given along the lines of the

planning processes described above, and allow for the construction of a coherent sequence of propositional units.

The systematicity of the differences found for the two languages and the internal logic of the different modes of perspective taking can be better understood if we look in detail at verbal material which has undergone not only micro- but also macrostructural planning. Producing a text requires — besides the conceptualization of sequentially ordered information units — the conceptualization of global, consistent frames of reference in different conceptual domains (space, time, person) in which the individual units are embedded and by which they are made coherent.

One of the "tools" used to establish criteria for consistent principles of organization across these conceptual domains lies in the choice of a specific perspective — which can be pinned down as the selection of a specific referential frame (time, space and person). In order to be able to talk about dynamic situations in a way that a hearer can reconstruct the communicated model of the reality in question, in more or less adequate terms, the speaker has to be consistent in how he linearizes and relates the individual pieces of information. It is well documented in the literature that the relation between temporal categories and events provides the crucial criterion for coherence creation in narrative type texts. And this is where the differences found in the encoding of single events — as analyzed above — should be mirrored at the level of macrostructural planning.

Before we report on the different patterns of text organization in these kinds of texts, let us first take a brief look at the conceptual components involved in the conceptualization of complex time–event relations. What kind of choices does the speaker have to make? The EVENT in its substance is represented by a predicate and its arguments. The predication is dynamic in nature and the situation referred to has a beginning, a course, and an endpoint. We use the notion of PHASAL DECOMPOSITION to refer to this aspect of event construal. Then there is the concept of a TIME LINE that is structured as an abstract sequence of intervals. And, third, there is the subject who conceptualizes, the observer/speaker. In producing coherent discourse, the speaker has to decide how to link substance to time structure, which requires the selection of a specific anchor point at the level of the single event, and he has to decide how the events should be related to each other. A referential frame, anchored with respect to a given perspective, ensures that the type of information that is mapped into the different components making up a propositional unit can be linked to form a coherent pattern. In other words, perspective taking leads to a specific form of interrelation between principles that guide the

selection of components at the microstructural level and principles of information flow at the macrostructural level.

In order to investigate these macrostructural planning processes we looked at another corpus of *Quest* retellings which was elicited under an offline condition (data type: Quest 1; see the introduction to section 3.1). In this mode of presentation, speakers first viewed the film *Quest* as a whole (desert, paper, stone, and machine world). They then saw it again, world by world, and after each transition from one world to the next (which are natural breaks in the film) the film was stopped and subjects were asked to tell an interlocutor *what happened/was ist passiert?* in the world they just saw.[5] In this case, the subjects knew what they were going to talk about, and in fact it could be shown in the analysis that texts of this type are based on macrostructural planning processes (see von Stutterheim i.p.; Carroll 2000). They differ from the texts obtained in the online retellings, which did not exhibit indicators of macrostructural planning to a comparable degree.

German. As was shown in the experiments described above, German speakers prefer to conceptualize a dynamic situation as an entity with boundaries. The internal structure of the events is not taken into account. In German, speakers map the event structure (events as entities) and the time line (as an abstract, interval structure) onto each other. Metaphorically speaking they fill a series of time slots with substance. In order to fit this pattern, events have to be bounded entities (for a more detailed analysis, see von Stutterheim and Lambert i.p., Carroll and von Stutterheim 2003). In principle, left as well as right boundaries can serve the purpose of providing an anchor point for temporal shifts, and in the case of German, right boundaries are crucial for organizing temporal structure in texts. The following typical example of a German retelling illustrates this pattern:

(3)

001 und sucht nach dem Wasser
'and he searches for water'
(unbounded)

002 und er gräbt mhm erst ein kleines bisschen
'and he digs a little'
(bounded)

003 und dann gräbt er immer stärker
'and then he digs even more intensively'
(unbounded)

004 bis ein Trichter entsteht
'until a funnel appears'
(boundary for 3, bounded)
005 und dann fällt er selbst in den Trichter rein
'and then he falls into the funnel himself'
(bounded)
006 und versucht sich allerdings noch freizuschaufeln
'and tries to dig himself out'
(unbounded)
007 was er nicht schafft
'which he doesn't succeed in doing'
(unbounded)
008 und wird dann von dem Trichter verschluckt
'and is then swallowed up by the funnel'
(boundary for 6/7, bounded)

The pattern exhibited in this text follows the strategy of linking the reference time of an utterance[6] to the preceding time of situation. The following features are characteristic of texts that are structured according to this principle:

a. Explicit marking of temporal relations by temporal adverbials for the sequence of events. Adverbials are used to establish the link between the time of situation and reference time, typically by the temporal shifter *then*.

b. A holistic view is taken on the events, including points of completion, or results of an event, phasal decomposition occurs very rarely.

c. The perspective chosen follows the event line from the inside, as a participant so to speak. Situation times are hooked up to each other. The deictically anchored point of view of the speaker does not play a role in constituting temporal structure at the macrostructural level.

If it is the case that German speakers take a perspective where they "fill" time slots with substance, what is the criterion in selecting the size of these slots? In deciding on the structure of the temporal slots, German speakers tend to draw upon the logic of the substance matter. They segment the flow of events in relation to the protagonist's intentions. The perspective is anchored at the protagonist — as an experiencer or actor — which leads to a goal-oriented segmentation.

Boundedness is a crucial notion with this perspective, as has been demonstrated in experiments for production as well as comprehension (cf. von Stutterheim et al. 2002). The viewing point or deictic origo provided by the speaker generally remains in the background in this system of information organization.

English. English speakers prefer a perspective under which events are hooked up to an external viewing point resulting in phasally decomposed event representations. This perspective is also the one underlying the construction of temporal relations at discourse level. In the English texts the central categories "event structure," "time line," and "reference point or origo of the observer" are linked as follows: event times are hooked up to a deictically anchored reference point, leaving the temporal relation between the events themselves implicit. The perspective underlying the English film retelling could be paraphrased by the question *what is happening now?* This perspective implies a view on the events as ongoing actions. The events are presented as unbounded in most cases without explicit information about the temporal relations between the event times. The hearer has to draw upon situational and world knowledge in order to bridge the lack of information with respect to the exact nature of the relevant event times.

(4)
001 he's on his knees
(unbounded)
002 and he's starting to dig
(inception phase, unbounded)
003 and starts digging faster and faster
(inception phase, unbounded)
004 and the hole starts to multiply
(inception phase, unbounded)
005 and gets bigger and bigger
(unbounded)
006 and he starts to slide into this hole
(inception phase, unbounded)
007 and he slides down
(unbounded)
008 and he's falling in what seems like through the earth
(unbounded)
009 and he just disappears
(bounded)

The time line with its interval structure remains in the background since the observer provides an external anchor point or origo for locating the events in time. The speaker presents the situations as coming into being or ongoing. The stream of dynamic situations passes, metaphorically speaking, in front of the observer's eye.[7] This corresponds to Talmy's (1988) windowing of attention in that with a fixed origo as reference point — in the given case the deictic "now" of the observer — attention

is directed to one slot or window at a time. In contrast to the German texts, the relation between event times is not central for establishing temporal coherence in the English texts.

So in the preferred pattern of information organization in English, the observer as anchor point of a deictic referential frame is the source for creating coherence between events. The time line with its abstract interval structure does not take over a linking function in this pattern of conceptualization of event–time relations. Explicit links between event times can be made by activating the time line, where inherent information is not explicit enough, but this is carried out only in marked cases in the data analyzed. This implies that unlike in German, the entity (protagonist) involved in an event is not accorded a structuring role in the segmentation or construal of what counts as an event.

Summary. Summarizing and interpreting the findings, we can say that, in the German texts, speakers adopt an event-based perspective (with the observer/narrator remaining in the background) from which events are related to each other by linking event times. The typical anaphoric relation established between the utterances in the texts is a shift-in-time relation. This requires the construal of events that are bounded.

In the English texts, the observer functions as temporal anchor point, opening a window on the flow of the events. This perspective does not require endpoints in order to relate events. Each event is hooked up to the deictic "now," and the relation between the events has to be inferred, or remains unspecific. Under this perspective, events are segmented into phases dependent on their real-time duration. Very short situations appear completely within one frame, and they can be presented holistically (*he jumps into the hole*). Other events having longer duration are presented as ongoing or coming into being (cf. Carroll 2000; Carroll and von Stutterheim 2003).

This difference in perspective taking leads to very different patterns of information organization at the macrostructural level (von Stutterheim and Klein 2002). In earlier studies, we have shown that these different perspectives are not easily mapped onto each other. Studies on translations from English-German and vice versa support the claim that the perspective underlying a piece of text in one language cannot be transported into the other language in a comparable way (cf. Carroll 1993; von Stutterheim 1997a, 1997b). Selecting information according to the target language's perspective means that a content preserving translation is not possible.

4. The roots of the contrast

4.1. *Alternative hypotheses*

In the preceding section we have argued that the contrasts observed in selecting and structuring information when talking about events can be generalized by assuming contrasts at the level of abstract principles of perspective taking. The choice of a specific referential frame sets constraints on how the "substance matter" represented in some knowledge base is selectively reconstructed in the process of conceptualizing for speaking. These constraints guarantee the creation of a consistent and coherent information structure as observed in the texts.

Why should speakers of different languages rely on different principles of perspective taking in solving the same communicative task? Although it is clear that differences hold between two groups of subjects that are defined by their native language, it is not at all clear how language as an abstract system could be responsible for the differences at issue. After all, there is nothing in the English GRAMMAR that prohibits the mentioning of endpoints, nor is it the case that German does not provide means to refer to the notion of ongoingness. Besides influences of linguistic structure there may be a variety of possible nonlinguistic causes for the differences found, ranging from individual stylistic differences or different learning traditions on how to construct a text, to deeply rooted cultural differences of various sorts.

There is, of course, one structural feature of English that is central to the verbalization of events and that differs from German: verbal aspect. Aspectual distinctions are grammaticized in English and the predicate is obligatorily marked for aspect, with the PROGRESSIVE ASPECT as the marked form. In Standard German[8] the aspectual distinction PROGRESSIVE–NONPROGRESSIVE can only be expressed by lexical expressions, as mentioned earlier. This means that the notional category of ongoingness is less salient in planning an utterance in German compared to English.

In line with current work on grammaticalization (cf. Traugott and Heine 1991; Hopper 1995), we will argue that grammaticalized conceptual categories play a predominant role in deciding how conceptual material is organized for expression. In our case we postulate that the structural feature [+/− aspect] induces a specific pattern of event construal. With only two languages as the empirical basis for comparison, however, we cannot exclude other language external factors as a driving force behind the contrasts observed.

4.2. *The case of Algerian Arabic*

In order to test these different hypotheses, we had to look at speakers of a language that is culturally different from and structurally similar to one of the two languages studied. For this reason, Arabic (in the variety of Algerian Arabic) was selected. Algerian Arabic belongs to the Semitic languages, and the cultural tradition is certainly not closely related to middle European culture. What is important in our case, however, is that Algerian Arabic shares the crucial feature of verbal aspect with English.[9] This constellation of languages — English, (Algerian) Arabic, and German — allows us to control the two variables involved, namely cultural and linguistic ones. In other words: if the construal of events in Algerian Arabic follows other patterns than the ones observed for the English speakers, this would favor an explanation that related the contrasts to cultural differences; if, however, events are verbally represented in a way that is similar to English patterns, this would exclude an explanation on a cultural basis and thus favor a linguistically based line of argumentation.

In order to obtain comparable data, the same verbal tasks were run with Algerian Arabic subjects, namely the offline retelling of *Quest* (Quest 1), online retelling of *Quest* (Quest 3), and the description of single situations (computer animations).[10] In a first step, the same analyses as carried out for German and English (cf. sections 3.1 and 3.2) were also applied to the Arabic data; that is, for the online retellings, the number of encoded events for a particular scene and the number of mentioned "small events" (of the type mentioned in Table 1) in the whole film were counted, while the computer animations were analyzed with regard to the number of endpoints mentioned. Table 3 shows the results:

Table 3. *The results for the Arabic texts compared to those for German and English texts reported in sections 3.1 and 3.2. In the first row, the so-called "paper world" was chosen for the comparison, because the number of encoded events was not counted for the whole film in Arabic; note that the special result for this scene was not reported in section 3.1*

	German	English	Arabic
number of events encoded (paper scene of movie *Quest*)	12.2	17.4	17.1
number of "small events" encoded (in whole movie)	11.2	21.5	19.7
number of endpoints mentioned (in descriptions of single events)	5.75	1.80	1.63

The results are clear-cut: With regard to the numbers at issue, Algerian Arabic speakers behave like English speakers. Arabic and English speakers mention more or less the same number of events in their retellings of *Quest*, and there is also no difference with regard to the mentioning of endpoints. (These results yield the same statistically significant differences between German and Arabic as those found between German and English [see sections 3.1 and 3.2].)

Turning to macrostructural temporal properties in the Algerian-Arabic[11] Quest-1 texts (offline retellings), the picture again is similar to that observed for English. As already shown by the experiment using single situations (computer animations), the construal of events corresponds to the English patterns both with respect to the selection of components as well as phasal decomposition. Speakers present the events mainly under an imperfective or inchoative perspective, leaving endpoints or results unspecified.

(5)

001 gɛad yħawwes ɛla l-ma (inceptive phase, unbounded)
begin PFV.3SG.M search IPFV.3SG.M on the-water
'he began searching for water'

002 gɛad yšūf (inceptive phase, unbounded)
begin PFV.3SG.M look IPFV.3SG.M
'he began looking'

003 gɛad yeħfar b-lidi-h (inceptive phase, unbounded)
begin PFV.3SG.M dig IPFV.3SG.M with-hands-his
'he began digging with his hands'

004 yeħfar (unbounded)
dig IPFV.3SG.M
'he is digging'

005 yeħfar (unbounded)
dig IPFV.3SG.M
'he is digging'

006 sartāte-h r-ramla (bounded)
swallow PFV.3SG.FEM-him the-sand
'the sand swallowed him up'

Although there are important differences in the aspectual systems of Arabic and English that have consequences for the way dynamic situations are segmented in an online retelling, correspondences remain with respect to one crucial feature: the global perspective that provides the basis for coherence creation at text level. The speaker establishes an event-external deictic ''now,'' and event times are hooked up to this point of reference. This perspective implies phasal decomposition and temporal

relations based on event times are not exploited for the creation of coherence.[12]

In conclusion, the results in Table 3 and the examination of macro-structural properties show that English and Arabic are similar, while both languages differ from German.

4.3. *Roots of the contrast*

Given the different hypotheses outlined above, what do the results tell us about their relative explanatory value? An explanation which attributes the contrast to individual preferences can be rejected on the basis of the experimental findings. Although the differences found are not 100% one way or the other, but should rather be described as preferences, they have proven to be stable and significant across a large number of subjects, tasks, and modes of speaking (see also Carroll and von Stutterheim 2003).

The different-culture hypothesis can be rejected on the basis of the comparative study based on the languages English, Algerian Arabic, and German. The parallel patterns found for speakers of English and Algerian Arabic with respect to event construal, and in contrast to German speakers, excludes the explanation based on cultural differences and leaves the structural differences in the languages involved. English and German differ with respect to verbal aspect, an essential category in the present context, and this is where English and Algerian Arabic meet: in both languages the predicate is marked for aspect. If a language marks the predicate for aspect, then speakers will parse event sequences into comparably small units, select different components for encoding (phasal structure), and prefer a global perspective under which event times are hooked up to a deictic center.

So far we have extended the line of argumentation at several points from a mere correlation between linguistic structure and patterns of event construal to postulating language-specific patterns at the level of conceptualization. There is, however, no relation of determination between formal properties of a language and options at the level of conceptualization. As was noted earlier, there are many ways of expressing a given content, drawing upon the full range of lexical, morphological, syntactic, and prosodic means. How can differences in form be responsible for patterns of perspectivation in discourse planning rooted at the conceptual level?

One possibility is that the causal link from linguistic structure to conceptualization runs via the wording of the instructions. Remember that the English (and the Arabic) instructions in both the online film

retelling and the computer animations contained a progressive form (''tell what is happening''), while the German instructions, simply because of the lack of a progressive form in German, were stated in the simple present (*erzähl was passiert* 'tell what happens'). So it might be the case that the use of the progressive in the instructions somehow ''primed'' the English subjects to structure events the way they did. What speaks against this possibility, however, is the fact that although (cf. section 3.3) no progressive form was used in the instructions for the offline retellings (''tell what happened''), the segmentation of the event sequence was still different for English and German subjects, that is, English speakers still selected smaller units, compared to German speakers. A control study with the computer animations in which the instruction for English speakers is ''what happens'' (that is, the instruction is stated in the simple present), indicates that this does not significantly affect the extent to which endpoints are mentioned. So all in all, differences between English and German speakers are not affected by the way the instruction is formulated; rather the conceptualization of events in English and German can be attributed to factors that are independent of the specific wording of the instruction.

In order to find an explanation for these preferences, we have to go back to the roots, to the place where linguistic knowledge and principles of conceptualization develop. The child who acquires an aspect language is led to distinguish different phases of an event right from the beginning. For English, it is a well-documented fact that children acquire the *-ing* as the first marking on the verb (cf. Brown 1973). Tied to the deictic ''here and now'' in the earlier stages of acquisition, the English child focuses on ongoing activities. The German child, by contrast, does not get a form in his or her input for referring to WHAT IS HAPPENING RIGHT NOW. The first functionally relevant opposition marked in the verbal system acquired by German children is the one between some unmarked form and the past participle. The participle, as well as verbal particles, are very productive in German and often influence the AKTIONSART of the lexical stem, many of them with a resultative reading. These forms set the focus on the result of an action, implying some endpoint of the action expressed by the lexical content of the verb (*umefallt* 'has fallen') (cf. Behrens 1993). So German and English children are confronted with an elementary difference in perspective taking, what we have called a HOLISTIC ENDPOINT-ORIENTED perspective versus a perspective of PHASAL DECOMPOSITION; thus, they are ''trained'' to conceptualize events differently just by learning their languages. Accordingly, habitual differences in event conceptualization develop along with the acquisition of different types of languages (i.e. aspect languages vs. nonaspect languages).

The hypothesis that in the course of language acquisition the child is oriented to particular event dimensions is in line with current work on language acquisition. This claims that learning a specific target language directs the child's attention toward distinctions encoded in that target language (cf. Bowerman 1996a, 1996b; Bowerman and Choi 2001; Slobin 1991, 1996). The empirical facts in language acquisition "provide a more conclusive support for the FORMAL PACESETTING HYPOTHESIS, which predicts that children are sensitive to the semantic distinctions encoded in their target language instead of going through a universal conceptual-semantic phase first" (Behrens 1993: 178).

Patterns of perspectivation induced by grammatical categories in one domain, in our case the temporal domain, may also have consequences in organizing conceptual material in domains where there is no immediate grammatical counterpart, such as the mentioning of an endpoint in the two languages in question. These implications may result in systematic constraints given the need to select and structure information from different conceptual domains according to overreaching principles in discourse production. One manifestation of these constraints could be seen in patterns of perspective taking — the ones observed in the data. Verhaar has formulated the interrelation between grammar and content looking at it from the other end, the fully-fledged language: "What is 'structural' in language is a grammaticalized/syntacticized form of discourse-pragmatic strategies. 'Structures' cannot be adequately understood without recourse to discourse-pragmatic research (…) (Verhaar 1983: 21).

5. Conclusion

In conclusion, we will relate the findings to the questions raised at the beginning. What do the results tell us about the nature of the conceptualizer, in particular the role of the specific language in processes of conceptualization?

We have seen that speakers of different languages — in our case English, Algerian Arabic and German — organize information systematically under different perspectives. The differences observed are statistically significant for both language pairs: English–German and Algerian Arabic–German. The contrasts are manifested in the choice of a referential frame, selection of aspects of a situation for explicit representation (components of the factual knowledge left for inferencing, components left unspecified/open) and structuring of the selected information. It is

further reflected in the choice of particular patterns of information structure (e.g. deictic versus anaphoric temporal linkage).

If we do not want to attribute decision-making powers to the formulator, then the selection of the content components for verbalization, and the choice of a particular perspective, have to be located within the process of conceptualization. There is no grammatical constraint that restricts the verbalization of an endpoint in English, nor is it excluded in the German system to talk about ongoing activities such as in *er gräbt gerade* 'he digs, just now'. Therefore the differences found in the two languages cannot be explained as a mapping phenomenon, where the same preverbal message is forced into different linguistic clothes.

On the basis of the empirical findings we come to the conclusion that processes of conceptualization follow language-specific principles, to some extent at least. As soon as a communicative intention is formed linguistic knowledge is triggered and in this context we speak of language-specific PRINCIPLES OF INFORMATION ORGANIZATION. So far we can only say that these principles are closely linked to perspective taking, which is basic for planning at the macrostructural and microstructural level.[13]

What does this imply for the three views on the language specificity of conceptualization mentioned at the beginning? If one accepts the conclusion that our findings indeed tell us something about conceptualization, then the first of the positions outlined above, the view that conceptualization is universal and language-free, has to be rejected. Conceptualization in language production must, or at least in certain respects, be based on language-specific principles.

The moderate position advocated by Levelt (1996), who argues for language specificity at the level of microplanning, has to be rejected on the basis of our results, too. For if macroplanning is DECIDING WHAT TO SAY (Levelt 1996: 77 and 103), the studies described in section 3 show that English and German speakers do encode (and, thus, ''decide to say'') different things. In particular, the film retellings revealed that English and German speakers encode different events, thus making different decisions on what to say: that is, they differ in macroplanning.

Does this mean that the second radical position is correct, that is, the Whorfian position, which claims that cognition in general is language-dependent? Not necessarily. For although this position is compatible with the results, it is not implied by these results. Our studies so far only deal with conceptualization WHILE the subjects are speaking, that is, they deal with the conceptualization of events for purposes of speaking. Nothing is said in these studies about event cognition in general. Indeed, a study on nonverbal event segmentation shows that there is no difference between English and German subjects when they distinguish events by

nonverbal means (cf. Nüse i.p.). So these results, as they now stand, do not provide support for this position.

This leaves us with the second moderate position, namely the thinking-for-speaking hypothesis put forward by Slobin (1991, 1996a). We believe that the present findings are compatible with this position, with the important claim that language-specific knowledge is already crucial at the global, macrostructural level of planning. This means that there is language-specific knowledge that is not part of the lexicon or grammar, namely knowledge of language-specific PRINCIPLES OF INFORMATION ORGANIZATION. The function of these language-specific principles — which have been described as modes of perspective taking — can be seen as an interface that provides guidelines for the selection and structuring of knowledge that is stored in different formats but not in a language-like linear structure.

As can be concluded from our brief discussion of the language-acquisition process, speaking of language specificity at the level of conceptualization does not imply a reversal of sequentiality for the production process, that is, an influence of the effect (the produced language-specific formulation) on the cause (the conceptualization and planning process). Rather, language-specific patterns of conceptualization develop along with the knowledge of the first language one acquires. So in the language-production system of the adult speaker the conceptualizer is already "tuned" to the specific language by drawing on this knowledge before the formulator is addressed. At this stage we cannot say anything about the details of this form of language specificity; for example, it is not known whether it works from the outset of the conceptualization process or whether it constitutes a later phase in which a possibly universal cognitive representation is reshaped. Likewise, it is not clear how "deep down" language specificity reaches, whether, for example, visual attention is already affected. The question could be whether English and German speakers focus their attention (for purposes of speaking!) on different aspects of the same situation. This will be the topic of further experimental investigations, including, among others, an eye-tracking study on what is scanned by English and German subjects when describing events. For the present the findings reveal that conceptualization in language production is not as language-free and universal as most current models of language production assume.

Received 14 June 2001
Revised version received
20 March 2002

University of Heidelberg

Notes

1. We would like to thank the Deutsche Forschungsgemeinschaft for financial support and two anonymous reviewers for valuable comments on earlier versions of the paper. Correspondence address: Professor Christiane von Stutterheim, Institut für Deutsch als Fremdsprachenphilologie, Universität Heidelberg, Plöck 55, 69117 Heidelberg, Germany. E-mail: stutterheim@mail.idf.uni-heidelberg.de.
2. An in-depth study of argument alternations on the basis of conceptual characteristics of the event types involved is presented in Härtl, this issue.
3. The film *Quest* was produced by T. Stellmach in 1997. It tells the story of a little clay figure, which in its search for water passes through different worlds: a sand, a paper, a stone, an iron, and a machine world.
4. Chi-square values were calculated by using ESTIMATED expected frequencies (with the usual formula that involves the product of the marginal totals divided by the number of subjects; see e.g. Hayes 1973: 730). That is, the H_0 is not an equal distribution of subjects who mention and subjects who do not mention the endpoint within one language, but rather, no difference of distributions between languages, or, in other words, that there is the same percentage of German and English speakers who mention the endpoint with a given item.
5. There is a general methodological problem related to the function of the question used in the instructions with respect to the perspective selected by the speakers in the texts. The fact that questions have to be posed differently in German and English if formulated in the present tense, might induce, at least to some extent, the differences observed in the texts (*what is happening/was passiert?* 'what happens'). However, the Quest-1 corpus shows that speakers do not follow the perspective suggested by the introductory question (*what happened/was ist passiert?*) unless it matches the preferred pattern for film renarrations. The past-tense perspective was not adopted by the speakers, since they all retell the film in the present tense.
6. For a more fine-grained analysis of temporal categories on the basis of the notion of *topic time* see Klein (1994); von Stutterheim and Lambert (i.p.).
7. This interpretation is supported by Hopper's claim that narratives in English do not consist of a sequence of reported events but "rather, at every point the narrator is concerned not to produce events 'wie sie eigentlich gewesen', but instead to present a past situation from a particular perspective, naturally her own" (Hopper 1995: 145). Hopper continues: "The narrations combine an account of what is being reported with a set of epistemic attitudes and perspectives, in other words to construct an account of 'what happened' with a view to coopting the listener/reader into the narrator's perspective" (1995: 146). Hopper attributes this feature to spoken narratives contrasting it to a written German text which presents a sequence of events "objectively." On the basis of our findings we would argue that this is not a difference between spoken and written language but rather between the preferred perspective in English in contrast to German.
8. There are varieties of German which have developed a grammatical construction which expresses ongoingness. *Er ist am/beim Schreiben* (copula + prep. + nominalized verb). This, however, is not a grammatically obligatory aspectual opposition. In Standard German use of this form is very constrained.
9. English and Algerian Arabic are examples of different types of aspect languages (for a description of the North African dialects of Arabic, see Simeone-Senelle 1986; Boucherit 1987; Cohen 1989; Caubet 1993). Algerian Arabic has a perfective/

imperfective opposition expressed by verbal morphology; English has a progressive/nonprogressive opposition. These differences, however, are not relevant for the present question. Here we are concerned with languages which structurally oppose different ways of viewing the temporal properties of a situation and therefore encode phasal decomposition of events in their verb morphology, as opposed to languages which do not.

10. Subjects were matched with the others groups for age and social background.
11. Recorded in north-western Algeria, Oran.
12. Text structure of this type provides evidence that a complementary relation — as discussed in Bohnemeyer (1999) — does not hold in all cases: "It is the specific aim of this study to explore, in a case study on Yukatek Maya, to what extent boundary information and event order information are complementary, such that event order may be inferred from information about the boundaries of events in discourse representation" (Bohnemeyer 1999: 45).
13. This type of knowledge, which we describe as principles of information organization, appears to be very deeply rooted. Data from very advanced L2 speakers show that these levels of linguistic knowledge are very difficult to restructure (cf. Carroll and von Stutterheim 1997, 2003; Carroll et al. 2000; von Stutterheim and Lambert i.p.).

References

Behrens, Heike (1993). Temporal reference in German child language. Unpublished doctoral dissertation, University of Amsterdam.

Bierwisch, Manfred; and Schreuder, Robert (1992). From concepts to lexical items. *Cognition* 42, 23–40.

Bohnemeyer, Jürgen (1999). Time relations in discourse. Unpublished doctoral dissertation, University of Tilburg.

Boucherit, Aziza (1987). Note sur l'expression de la concomitance dans. le système verbal de l'arabe parlé à Alger. MAS-GELLAS (Association Gellas — Université Paris III). 11–20.

Bowerman, Melissa (1996a). Learning how to structure space for language: a cross-linguistic perspective. In *Language and Space*, Paul Bloom, Mary A. Peterson, Lynn Nadel, and Merritt F. Garrett (eds.), 385–436. Cambridge, MA: MIT-Press.

—(1996b). The origins of children's spatial semantic categories: cognitive versus linguistic determinants. In *Rethinking Linguistic Relativity*, John J. Gumperz and Steven C. Levinson (eds.), 145–176. Cambridge: Cambridge University Press.

—; and Choi, Soonja (2001). Shaping meanings for language: universal and language-specific in the acquisition of spatial semantic categories. In *Language Acquisition and Conceptual Development*, Melissa Bowerman and Steven C. Levinson (eds.), 475–511. Cambridge: Cambridge University Press.

Brown, Roger (1973). A first language: the early stages. London: Allen & Unwin.

Brown, Paula A.; and Dell, Gary S. (1987). Adapting production to comprehension: the explicit mention of instruments. *Cognitive Psychology* 19, 441–472.

Carroll, Mary (1993). Deictic and intrinsic orientation in spatial descriptions: a comparison between English and German. In *Cognition and Culture: A Cross-Cultural Approach to Psychology*, Jeanette Altarriba (ed.), 23–44. Amsterdam: North-Holland.

—(2000). Representing path in language production in English and German: alternative perspectives on figure and ground. In *Räumliche Konzepte und sprachliche Strukturen*, Christopher Habel and Christiane von Stutterheim (eds.), 97–119. Tübingen: Niemeyer.

—; and Stutterheim, Christiane von (1993). The representation of spatial configurations in English and German and the grammatical structure of locative and anaphoric expressions. *Linguistics* 31(6), 1011–1041.

—; and Stutterheim, Christiane von (1997). Relations entre grammaticalisation et conceptualisation et implications sur l'acquisition d'une langue étrangère. *AILE (Acquisition et interaction en langue étrangère)* 9, 83–115.

—; and Stutterheim, Christiane von (2003). Typology and information organisation: perspective taking and language-specific effects in the construction of events. In *Typology and Second Language Acquisition*, Anna Giacalone Ramat (ed.), 365–402. Empirical Approaches to Language Typology 26. Berlin: Mouton de Gruyter.

—; Murcia Serra, Jorge; Watorek, Marzena; and Bendiscioli, Alessandra (2000). The relevance of information organisation to second language acquisition studies: the descriptive discourse of advanced adult learners of German. *Studies in Second Language Acquisition* (Special issue ed. by Clive Perdue), 441–466.

Caubet, Dominiques (1993). *L'arabe marocain. Tome II. Syntaxe et catégories grammaticales.* Paris-Louvain: Editions Peeters.

Cohen, David (1989). *L'aspect verbal.* Paris: Presses Universitaires de France.

Croft, William (1991). *Syntactic Categories and Grammatical Relations: The Cognitive Organisation of Information.* Chicago, IL: University of Chicago Press.

Engelkamp, Johannes (1990). *Das menschliche Gedächtnis: Das Erinnern von Sprache, Bildern und Handlungen.* Göttingen: Hogrefe.

Habel, Christopher; and Tappe, Heike (1999). Process of segmentation and linearization in describing events. In *Conceptual and Semantic Knowledge in Language Production*, Ralf Klabunde and Christiane von Stutterheim (eds.), 117–152. Opladen: Westdeutscher Verlag.

Hayes, William L. (1973). *Statistics for the Social Sciences*, 2nd ed. London: Holt, Rinehart & Winston.

Herrmann, Theo; and Deutsch, Werner (1976). *Psychologie der Objektbenennung.* Bern: Huber.

—; and Grabowski, Joachim (1994). *Sprechen: Psychologie der Sprachproduktion.* Heidelberg: Spektrum.

Hopper, Paul (1995). The category 'event' in natural discourse and logic. In *Discourse Grammar and Typology: Papers in Honor of John W. M. Verhaar*, Werner Abraham, Talmy Givón, and Sandra Thompson (eds.), 139–150. Amsterdam: Benjamins.

Jackendoff, Ray (1990). *Semantic Structure.* Cambridge, MA: MIT Press.

Kamp, Hans (1981). A theory of truth and semantic representation. In *Formal Methods in the Study of Language*, Jeroen Groenedijk (ed.). Amsterdam: Mathematisch Centrum.

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

Levelt, Willem J. M. (1982). Linearization in describing spatial networks. In *Processes, Beliefs, and Questions*, Stanley Peters and Esa Saarinen (eds.), 199–220. Dordrecht: Reidel.

—(1989). *Speaking.* Cambridge, MA: MIT Press.

—(1996). Perspective taking and ellipsis in spatial descriptions. In *Language and Space*, Paul Bloom, Mary A. Peterson, Lynn Nadel, and Merritt F. Garrett (eds.), 77–107. Cambridge, MA: MIT Press.

Levinson, Steven C. (1997). From outer to inner space: linguistic categories and nonlinguistic thinking. In *Language and Conceptualization*, Jan Nuyts and Eric Pederson (eds.), 13–45. Cambridge: Cambridge University Press.

Lucy, John A. (1992). *Grammatical Categories and Thought.* Cambridge: Cambridge University Press.

—(1996). The scope of linguistic relativity: an analysis and review of empirical research. In *Rethinking Linguistic Relativity*, John J. Gumperz and Steven C. Levinson (eds.), 37–69. Cambridge: Cambridge University Press.

Newtson, Darren (1976). Foundations of attribution: The perception of ongoing behavior. In *New Directions in Attribution Research*, John H. Harvey, William Ickes, and Robert F. Kidd (eds.), 223–247. Amsterdam: Erlbaum.

Nüse, Ralf (i.p.). Segmenting event sequences for speaking. In *The Syntax-Semantics-Interface: Linguistic Structures and Processes*, Holden Härtl, Susan Olsen, and Heike Tappe (eds.). Berlin: Mouton de Gruyter.

Parsons, Terence (1990). *Events in the Semantics of English: A Study in Subatomic Semantics.* Cambridge, MA: MIT Press.

Rickheit, Gerd; and Strohner, Hans (1993). *Grundlagen der kognitiven Sprachverarbeitung.* Tübingen: Francke.

Simeone-Senelle, Marie-Claude (1986). Système aspecto-temporel en Arabe Maghrébin. *MAS-GELLAS*, 57–61.

Slobin, Dan I. (1991). Learning to think for speaking: native language, cognition and rhetorical style. *Pragmatics* 1, 7–26.

—(1996a). From 'thought and language' to 'thinking for speaking'. In *Rethinking Linguistic Relativity*, John J. Gumperz and Steven C. Levinson (eds.), 70–96. Cambridge: Cambridge University Press.

—(1996b). Two ways to travel: verbs of motion in English and Spanish. In *Grammatical Constructions: Their Form and Meaning*, Masayoshi Shibatani and Sandra A. Thompson (eds.), 195–217. Oxford: Oxford University Press.

Stutterheim, Christiane von (1997a). *Einige Prinzipien des Textaufbaus.* Tübingen: Niemeyer.

—(1997b). Zum Ausdruck von Raum- und Zeitkonzepten in deutschen und englischen Texten. *Zeitschrift für Germanistische Linguistik* 25, 147–166.

—(i.p.). Zur Konzeptualisierung von Ereignissequenzen. *Sprachproduktion*, Christopher Habel (ed.). Opladen: Westdeutscher Verlag.

—; and Klein, Wolfgang (2002). Quaestio and L-perspectivation. In *Perspective and Perspectivation in Discourse*, Carl F. Graumann and Werner Kallmeyer (eds.), 59–88. Amsterdam: Benjamins.

—; and Lambert, Monique (i.p.). Crosslinguistic analysis of temporal perspectives in text production. In *The Structure of Learner Varieties*, Henriette Hendricks (ed.). Berlin: Mouton de Gruyter.

—; Nüse, Ralf; and Murcia Serra, Jorge (2002). Cross-linguistic differences in the conceptualisation of events. In *Information Structure in a Cross-Linguistic Perspective.* Hilde Hasselgård, Stig Johansson, Bergljot Behrens, and Cathrine Fabricius-Hansen (eds.), 179–198. Amsterdam: Rodopi.

—; Carroll, Mary; and Klein, Wolfgang (2003). Two ways of construing complex temporal structures. In *Deictic Conceptualization of Space, Time and Person*, Friedrich Lenz (ed.), 97–133. Berlin: Mouton de Gruyter.

Talmy, Leonard (1988). The relation of grammar to cognition. In *Topics in Cognitive Linguistics*, Brygida Rudzka-Ostyn (ed.), 165–205. Amsterdam: Benjamins.

Tomlin, Russel (1997). Mapping conceptual representations into linguistic representations: the role of attention in grammar. In *Language and Conceptualization*, Jan Nuyts and Eric Pederson (eds.), 162–188. Cambridge: Cambridge University Press.

Traugott, Elisabeth; and Heine, Bernd (1991). *Approaches to Grammaticalization.* Vol. 2. Amsterdam: Benjamins.

Verhaar, John (1983). Ergativity, accusativity, and hierarchy. *Sophia Linguistica Working Papers in Linguistics* 11, 1–23.