Representing temporal discourse markers for generation purposes

Brigitte Grote
Otto-von-Guericke Universität Magdeburg
Institut für Wissens- und Sprachverarbeitung
P.O. Box 4120, 39016 Magdeburg, Germany
email: grote@iws.cs.uni-magdeburg.de

Abstract

Discourse markers are an important means to signal the kind of coherence relation holding between adjacent text spans. Research on generating discourse markers has been mainly concerned with causal markers, whereas temporal markers have not received much attention. In this paper, we identify semantic, pragmatic and syntactic features that are required to support a motivated choice of German temporal subordinating conjunctions and prepositions during text production. Information on individual markers is assembled in a discourse marker lexicon, which is used as a declarative resource at the sentence planning stage. We illustrate how this resource can be used to produce alternative verbalizations of the temporal relationship holding between two events.

1 Motivation

In text, discourse markers signal the kind of coherence relation holding between adjacent text spans. For any but the most trivial applications of language generation, motivated marker choice is an important task. Whereas several studies have been concerned with causal markers and their interactions with other linguistic means, for instance, Vander Linden and Martin (1995), Rösner and Stede (1992), Delin et al. (1996), temporal markers as signals of the temporal relation holding between two events have not received much attention, with the exception being Dorr and Gaasterland (1995). However, quite often, it is only by means of explicit temporal markers that the correct interpretation of a text can be ensured, as the following examples illustrate:

(1) *Der Abstand ist nachzumessen, nachdem das Band mindestens einen Umlauf ausgeführt hat.* (Check the distance after the belt has completed at least one round.)

(2) *Während Sie den Toaster betreiben, die Brotschlitze nicht abdecken.* (While operating the toaster, do not cover the bread slots.)

In both cases, the order of events as recounted in the text does not correspond to their order of occurence: In example (1), *nachdem* (after) marks the event denoted in the second clause as temporally anterior to the one denoted in the first clause. In example (2), *während* (while) marks both events as cotemporal. Without a temporal marker, the order of events would not be obvious.

Selecting an appropriate discourse marker for a given temporal relation is by no means a straightforward matter, though. First, one and the same temporal relation is verbalized differently depending on the syntactic and lexical properties of the clauses it conjoins. In German, the language we consider, such properties are, among others, tense, aspect, and syntactic structure. Interdependencies occur when two temporal events are realized in the same sentence; we will therefore restrict the study of German temporal markers to subordinating conjunctions and prepositions. Second, temporal discourse markers can even overwrite the temporal relations indicated by other linguistic means like tense and aspect, as noted by researchers who work in the analysis of temporal markers (e.g. Hitzeman et al. 1995).

In a nutshell, when aiming at selecting an appropriate temporal marker in text generation one needs a representation of temporal markers that enables marker choice and other sentence level decisions (such as tense and aspect selection) to mutually constrain each other. In this paper, we take up the idea suggested in Grote and Stede (1998) of a discourse marker lexicon as a declarative resource at the sentence planning stage. We demonstrate how such a lexicon for temporal markers can be employed in text generation to produce different verbalizations of the same underlying temporal organization depending on other generation decisions.

The paper is organized as follows: Section 2 reviews related work on generating temporal markers. Section 3 describes the major semantic, pragmatic and syntactic properties of German temporal prepositions and subordinating conjunctions. Section 4 presents the generation perspective: It briefly discusses the shape of the discourse marker lexicon, introduces the features used in the lexicon, and presents sample lexicon entries and their application in the generation process.

2 Related work

Work on discourse marker generation in general has focussed on marker selection, mainly for causal relations (Elhadad and McKeown 1990; Vander Linden and Martin 1995), and on the realization of RST's subject-matter relations (Rösner and Stede 1992; Delin et al. 1996). As for temporal markers, Dorr and Gaasterland (1995) examine the generation of English temporal subordinating conjunctions. Gagnon and Lapalme (1993), on the other hand, describe the generation of French temporal adverbs based on a DRT representation of the discourse.

While Gagnon and Lapalme (1993) only briefly address conjunctions and prepositions, Dorr and Gaasterland (1995) present a detailed study of temporal connectives, but they consider English markers only. The only account on automatically producing German temporal expressions that we know of is Ehrich (1987); however, she discusses the interaction of tense and aspect in simple sentences only.

Most studies that deal with discourse markers regard their production as a mere consequence of other sentence level decisions such as aggregation, lexicalization, syntactic structuring, and—in the case of temporal markers—as determined by tense and aspectual choices. We believe, however, that one needs a more flexible control to increase the expressiveness of generation systems. Although there have been quite a few studies on individual aspects of sentence planning, little attention has been paid to the interaction between the various tasks—exceptions are Rambow and Korelsky (1992) and Wanner and Hovy (1996)—and in particular to the role of marker choice in the overall sentence planning process.

There exists a large body of research in NLU on analysing the temporal structure of texts, including the role of temporal markers, though again restricted to English (Moens and Steedman 1988; Lascarides and Oberlander 1993; Hitzeman et al. 1995). We turn to these studies when it comes to identifying the information that needs to be assembled for representing temporal markers.

3 Linguistic perspective: Describing temporal markers

Selecting an appropriate German temporal marker given two events in a temporal relationship requires detailed knowledge of the semantic, pragmatic and syntactic properties that characterize temporal markers. This section introduces the major properties and explores the correlations between temporal markers and other linguistic means that indicate temporal organization. We base our account on two sources: descriptive linguistic studies, mainly by Helbig and Buscha (1991), Bäuerle (1995), Buscha (1989) and Steube (1980); and our analysis of temporal marker usage in the German LIMAS corpus (Glas 1975).

3.1 The 'meaning' of German temporal markers

Temporal subordinating conjunctions and temporal prepositions conjoin two events where the event in the subordinate clause (or the PP) provides the temporal framework for interpreting the event in the main clause: *Bevor Sie den Toaster reinigen, den Netzstecker ziehen.* (Before you clean the toaster, unplug the device.) and the corresponding 'shorthand' form *Vor dem Reinigen des Toasters den Netzstecker ziehen* (Unplug before cleaning the toaster).

Semantic properties German grammars such as Helbig and Buscha (1991) list about 20 temporal subordinating conjunctions and 20 temporal prepositions. Their semantics is usually described by the kind of temporal relation they establish between two events, see for instance, Steube (1980) and Helbig and Buscha (1991): The event in the main clause can either overlap with (**simultaneity**), succeed (**anteriority**), or precede (**posteriority**) the event depicted in the subordinate clause or the prepositional phrase. In table 1 we provide a synthesis of the classifications of the most frequent German temporal markers by Helbig and Buscha (1991), Buscha (1989) and Bäuerle (1995). The markers listed in the table reflect the scope of the marker study in this paper.

Two aspects are especially prominent: First, each of the three temporal relations can be realized by a number of temporal markers. Alternatives within a class differ in that they realize some additional meaning aspect. Consider the markers of simultaneity: *Solange*, for instance, conveys the idea of a strict simultaneity where two events have the same start and end time, and is more specific than *während*; *sooft*, to give another example, highlights the concurrence of two events.

Second, table 1 shows that some markers are ambiguous: *Als* and *wenn* occur in all three classes, *seitdem*, *sobald* and *sooft* in two. Apparently, neither of them has any special temporal implicature on its own; instead, these markers depend on syntactic and lexical contexts to receive an umambiguous temporal meaning. We will return to this issue in section 3.2.

Pragmatic properties The choice of a particular marker to express a temporal relation between two events interacts with the focus structure as in:

(3) (a) *Bevor ihr Mann das Haus verließ, ging sie zur Arbeit.* (Before her husband left the house, she went to work.)
(b) *Nachdem sie zur Arbeit gegangen war, verließ ihr Mann das Haus.* (After she had gone to work, her husband left the house.)

Alternatives (3a) and (3b) both express that the event of 'going to work' precedes the event of 'leaving the house'. They differ in that they focus on

temporal relation	temporal markers	
simultaneity	subc:	*als (as), indes(sen) (meanwhile), seitdem (since), sobald (as soon as), solange (as long as), sooft (whenever), sowie (as soon as), während (while), wenn (when)*
	prep:	*an (at), auf (on), bei (during), binnen (within), durch (for), in (in), über (over), während (during)*
anteriority	subc:	*als (when), kaum daß (no sooner), nachdem (after), seit(dem) (since), sobald (as soon as), sooft (whenever), sowie (as soon as), wenn (when)*
	prep:	*ab (from), nach (after), seit (since)*
posteriority	subc:	*als (when), bevor (before), bis (until), ehe (before), wenn (when)*
	prep:	*bis (until), vor (before)*

Table 1: *German temporal subordinating conjunctions (subc) and prepositions (prep) classified by temporal relations. Note that the corresponding English markers are only approximate translations.*

different aspects of the situation: In (3a) the earlier event is in the centre of attention, in (3b) the later one (assuming that the matrix sentence is more prominent). This phenomenon interacts with other discourse phenomena, for instance, given and new information, and—when placed in a larger discourse context—with presuppositions and their accommodation (Lascarides and Oberlander 1993). However, the treatment of the discourse behaviour of temporal markers is beyond the scope of this paper.

Pragmatic issues further concern style. Regarding temporal markers, stylistic features are of minor importance: We only observe variation between archaic and neutral (*da* vs. *als*), and formal and neutral (*kaum daß* vs. *sobald*) markers.

3.2 Syntactic and lexical constraints

When expressing several events in the same sentence, marker choice interacts with other linguistic means: Temporal markers impose particular constraints on the syntactic and lexical contexts they can occur in. Conversely, these contexts can influence the meaning of markers.[1]

Markers and Aktionsart/aspect Aspect is traditionally taken to have two components, the non-inherent grammatical features, and the inherent lexical features. Inherent features characterize facets of the situation denoted by a verb, for instance, whether it is an event or a state. We will label these features *Aktionsart* to avoid confusion. According to Bussmann (1990), the major Aktionsarten in German are stative (*wissen/to know*) and dynamic. For the latter, the basic dichotomy is that between durative (*schlafen/to sleep*) and non-durative verbs, which are subdivided into iterative (*flattern/to flap*), semelfactive (*klopfen/to knock*), resultative (*verbrennen/to burn up*) and causative verbs (*tränken/to water*).

Two kinds of interdependencies are generally acknowledged, see Ehrich (1987), Buscha (1989) and Bäuerle (1995). First, temporal markers are sensitive to the Aktionsart of a verb. Consider *während* and *als* which can both express simultaneity:

(4) (a) *Als das Kabel schmolz / riß, war ich nicht im Raum.* (When the cable melted / tore, I wasn't in the room.)
(b) *Während das Kabel schmolz / *riß, war ich nicht im Raum.* (While the cable melted / *tore, I wasn't in the room.)

Während expects a durative verb in the subordinate clause, hence it can occur with *schmelzen/to melt* but not with *reißen/to tear*. *Als*, in contrast, can be used with durative and resultative verbs, as (4a) illustrates. Second, temporal markers may even shift the Aktionsart of a verb, for instance from a semelfactive reading to an iterative one as in:

(5) (a) *Wenn es an der Tür klopft, schreit das Baby.* (When someone knocks at the door, the baby cries.)
(b) *Während es an der Tür klopft, schreit das Baby.* (While someone knocks at the door, the baby cries.)

Grammatical aspect reflects the individual perspective a speaker adopts with respect to an event, such as perfective (temporally closed) or imperfective. In German, this distinction is grammatically realized by choosing a perfective or simple tense.[2] Aktionsart and aspect closely interact, consider example (6) where the anterior reading (6b) is due to the use of a perfective tense with a non-durative verb in the subordinate clause, which indicates that the activity has been concluded:

(6) (a) *Seitdem ich ihn kenne, ist er Nichtraucher.* (Since I know him, he is a non-smoker.)
(b) *Seitdem seine Frau gestorben ist, sehe ich ihn nur selten.* (Since his wife has died, I only rarely see him.)

[1] Traditional grammars, which the present account is based on, usually list aspect, Aktionsart and tense as constraining parameters on marker choice. However, there is no consensus on the role of these parameters; Bäuerle (1995) provides a good overview of the range of positions.

[2] In contrast to English, and especially to slavic languages, German has no elaborate aspect system: Distinctions like progressive and simple cannot be signalled by morphological features of the verb, but require a separate temporal adverb: *He is reading* vs. *Sie liest gerade* (She reads right now).

Here, verb properties determine the reading of the temporal marker. Our study of temporal marker occurrences in the LIMAS corpus suggests that markers belonging to the simultaneity class typically realize imperfective aspect, whereas temporal connecting words that signal anteriority correlate with a perfective aspect in the subordinate clause.

Markers and verbal tense Some markers can only be used with particular tenses, for instance, *als* in its simultaneous reading cannot occur with present tense, whereas *wenn* as signal of simultaneity correlates with present and past tense:

(7) (a) *Als er in Dresden war (*ist), suchte (*sucht) er seine Freundin auf.*
(b) *Wenn er in Dresden ist / war, sucht / suchte er seine Freundin auf.*

However, tempus sensitivity of temporal markers is not a matter of the grammatical tense form (such as simple past, present perfect, etc.) but relates to the temporal structure of the individual events, and to how their temporal structures are related. Assuming the Reichenbachian threefold distinction between Event Time (E), Reference Time (R), and Speaking Time (S) (the Basic Tense Structure, BTS, (Reichenbach 1947)), we observe that the constraints imposed by a marker on verb tense concern the underlying relation between E and S of both clauses: Selecting either *als* or *wenn* to express simultaneous events in the main clause (e_m) and in the subordinate clause (e_s) depends on whether the event times precede S (E(e_m),E(e_s)_S) or concur with S (E(e_m),E(e_s),S).[3] The grammatical tense results from combining the BTS of both clauses and their aspectual features.

Markers and syntactic structure The most straightforward correlation is that between syntactic structure and marker choice: If two events are expressed by a hypotactic structure, a subordinating conjunction is required. When a deverbal realization of an event is possible (e.g. *treffen/das Treffen*; to meet/the meeting), a clause with an adverbial (temporal) prepositional phrase is realized.

Markers and temporal quantifiers With some markers, the temporal relation denoted by the marker can be quantified by a temporal adverb as in *kurz bevor* (shortly before) or *einige Stunden nachdem* (several hours after); others cannot be quantified: **einige Stunden sobald* (*several hours as soon as).

4 Generation perspective: Representing temporal markers

A representation of temporal markers suitable for generation purposes has to accommodate the following demands: First, it has to describe the semantic and pragmatic features of markers in a manner that supports a motivated choice between markers which can realize the same temporal relation. Second, it has to account for the constraints temporal markers impose on their syntactic and lexical contexts, thereby enabling interactions between marker choice and other sentence planning decisions where the order of decision-making is not fixed. In Grote and Stede (1998) we argue that such a flexible control is best realized by introducing independent modules for the different sentence planning tasks, such as proposed by Wanner and Hovy (1996), and that these modules should rely on declarative representations as much as possible. Therefore, we propose a discourse marker lexicon, i.e. an independent lexical resource that assembles specifically the information associated with discourse markers.

Traditional lexicology and grammars describe lexical entries along three features: semantic, pragmatic and syntactic dimensions (see section 3). From the production perspective, these features are to be classified with respect to when and where they come into play in the generation process; this amounts to a procedural view on the information coded in the lexicon. Following Grote and Stede (1998) we assume three categories in the marker lexion:

- **Applicability conditions**: The necessary conditions that need to be present in the input representation for the marker to be a candidate. Chiefly, this is the semantic/discourse relation to be expressed, and also (if applicable) features pertaining to presuppositions and intentions.
- **Combinability conditions**: The constraints that the marker imposes on its neighbouring linguistic constituents (the 'syntagmatic' dimension). These are syntactic constraints on subcategorization and semantic type constraints, which interact with other realization decisions in sentence planning.
- **Distinguishing features**: If preferential choice dimensions, such as style, brevity, etc., are attended to in the system, then these features serve to distinguish markers that are otherwise equivalent (the 'paradigmatic' dimension).

In the remainder of this section we describe lexicon entries for temporal markers along these lines.

4.1 Applicability conditions

Semantic conditions The semantic classes introduced in section 3.1 (simultaneity, anteriority and posteriority) turned out to be too coarse for generation purposes. Instead, one needs a more fine-grained representation of the semantics of temporal markers to support an informed choice among markers within the broad classes.

Allen's temporal interval relationships provide an adequate framework (Allen 1984), as already sug-

[3] The comma stands for 'is cotemporal', the underscore for 'precedes'.

gested by Dorr and Gaasterland (1995). Allen introduces seven basic temporal interval relationships, namely *equals(=), after(>), during(d), overlaps(o), meets(m), starts(s), finishes(f)*—and their inverses *<,di,oi,mi,si,fi*—that may exist between two events e_m and e_s. For instance, *overlaps*(e_m, e_s) as in (4b) implies that there is an intersection between the time at which e_m occurs and the time at which e_s occurs, but that neither event is a subset of the other.

Each temporal relation corresponds to one or several German temporal markers, for instance, *overlaps* may be expressed by the entire range of simultaneity markers given in table 1, except for *solange* and *kaum daß*. Conversely, the majority of the temporal markers can realize several temporal interval relations. Take the connective *nachdem* as in example (1), which can have the following meanings,

after(e_m, e_s) ∧ *meets-i*(e_m, e_s)

or *während* as in example (4b),

equals(e_m, e_s) ∧ *during*(e_m, e_s) ∧ *starts*(e_m, e_s) ∧ *finishes*(e_m, e_s) ∧ *overlaps-i*(e_m, e_s)

whereas *solange* has only one reading:

equals(e_m, e_s).

This adequately captures the semantic difference between *während* and *solange*. In the lexicon, the applicability conditions of a particular temporal marker are now described by listing the temporal interval relations it can realize.

Pragmatic conditions In section 3.1 we briefly discussed pragmatic features of temporal markers. For the time being, the lexicon supports the features style, with the values `neutral, brief, formal, archaic`, and intention. Its value `evaluative` indicates the speaker's (negative) attitude towards the kind of temporal relation holding between two events (Steube 1980; Buscha 1989).

4.2 Combinability conditions

Combinability conditions appear as constraints in the lexicon entries of individual markers. In the present lexicon, constraints are described using the following features:

Aktionsart The Aktionsart plays a central role during the lexicalization of events: Candidate verbs are, among others, selected due to their Aktionsart. Aktionsart features are usually stored in the lexicon entries of verbs, and are thus available to sentence planning. To represent these constraints, we turn to Bussmann (1990) for the major Aktionsarten in German (see also section 3.2).[4] At present, the lexicon supports a subset of Bussmann's Aktionsarten, namely `stative, durative, iterative, semelfactive, causative` and `resultative`.

[4]There is no generally accepted and well-defined set of Aktionsart features; we opted for Bussmann (1990) because these features are supported by the lexicalization component we intend to use (Stede 1996).

Aspect Grammatical aspect is encoded using the feature values `perfective` and `imperfective`.

Tense We argued above that marker choice relates to the underlying temporal structure—as expressed in terms of the Reichenbachian threefold description of time–and not to a particular grammatical tense (see also Ehrich (1987)). Temporal constraints in the marker lexicon will thus be described using the BTS notation, and defining the legal linear orderings of E, R and S of the related events. For instance, *als* in its simultaneous meaning imposes the constraint E(e_m),E(e_s)_S, which can be realized by all grammatical tenses that meet this constraint.

Mapping this representation into grammatical tense requires knowledge on how to map pairs of Basic Tense Structures to the tense structure of complex German sentences, as described in Hornstein (1990) for English (Complex Tense Structures, CTS) and extended by Dorr and Gaasterland (1995) to cover intervals, too. Since we envision independent modules for the different sentence planning tasks that posit their choices as constraints, the tense selection process need not concern us.

Syntactic structure Possible values are `pp` (prepositional phrase) and `subord` (subordinate clause); both refer to the realization of the event that acts as temporal reference point.

Quantification The lexicon contains the two values `quantifiable` and `not-quantifiable`.

4.3 The shape of the lexicon

The possible values for the applicability and combinability features can now be used in the lexicon to describe individual temporal markers. Table 2 gives the lexical representations for most of the German anteriority markers and the posteriority marker *before*. Similar representations have been developed for the other marker classes given in table 1: simultaneity and posteriority. Feature values for individual markers have been identified by analysing marker occurrences in the LIMAS corpus (Glas 1975); as such, they mainly reflect marker *usage*. We then compared our marker descriptions to results from research literature (see section 3). Note that combinability conditions can apply to main and subordinate clause/prepositional phrase separately, hence some feature values are prefixed with `mc:`, `sc:` and `pp:` to mark their scope. If a marker involves no constraint for a particular feature, the slot in the table remains empty.

Table 2 contains an informal description of the lexicon entries; the formal representation depends on the actual sentence planner used in text production, see Grote and Stede (1998) for a preliminary proposal.

Feature	*nachdem* (after)	*nach* (after)	*sobald* (as soon as)	*kaum daß* (no sooner)	*bevor* (before)
applicability					
– denotation	after(e_m, e_s)∧ meets-i(e_m, e_s)	after(e_m, e_s)∧ meets-i(e_m, e_s)	meets-i(e_m, e_s)	meets-i(e_m, e_s)	before(e_m, e_s)∧ meets(e_m, e_s)
combinability					
– Aktionsart	sc:resultative∧ sc:iterative∧ sc:semelfactive	pp:resultative∧ pp:iterative∧ pp:semelfactive			sc:non-durative
– aspect	sc:perfective mc:imperfective				
– tense	$E(e_s)\_E(e_m)\_S$∧ $E(e_s)\_E(e_m), S$∧ $E(e_s)\_S\_E(e_s)$		$\{E,S\}(e_m)$ $=\{E,S\}(e_s)$∧ $E(e_m)\_E(e_s)\_S$	$E(e_s)\_E(e_m)\_S$∧ $E(e_s), E(e_m)\_S$	$\{E,S\}(e_m)$ $=\{E,S\}(e_s)$
– syntax	subord	pp	subord	subord	subord
– quantifier	quantifiable	quantifiable	not-quantifiable	not-quantifiable	quantifiable
preferences					
– style	neutral	brief	neutral	formal	neutral
– intention				evaluative	

Table 2: *Lexicon entries for some German temporal markers*

4.4 Selecting temporal markers

This section briefly addresses the issue of selecting an appropriate temporal marker during text production using the discourse marker lexicon. We will focus on the anteriority markers.

In our scenario, generation starts from a conceptual representation which contains the facts that must be reported in the text and their position in time. Let us assume the following very simple input structure:[5]

```
e1:  arrive(he,home,19:14)
e2:  watch(he,TV,19:15,22:30)
```

The first event precedes the second event, but they 'meet' at one point in time. Now, the first step is to determine the applicable temporal relations. Two interpretations are possible, depending on the discourse context and focus structure, which we have not dealt with so far: Focussing on the earlier event would yield the temporal relation *meets*(e_m, e_s), with $e_m = e_1$, focussing on the later event the relation *meets-i*(e_m, e_s), with $e_m = e_2$. Matching this against the lexicon entries in table 2 would produce *bevor* in the former case, and *nachdem, nach, sobald, kaum daß* as candidate realizations for the latter interpretation. Possible verbalizations are:

(8) (a) *Bevor er Fernsehen guckte, ist er nach Hause gekommen.*
(Before he watched TV, he has come home.)
(b) *Sobald er nach Hause gekommen war, guckte er Fernsehen.*
(As soon as he had come home, he watched TV.)
(c) *(Direkt) nachdem er nach Hause gekommen war, guckte er Fernsehen.*
((Right) after he had come home, he watched TV)
(d) *Nach dem Heimkommen guckte er Fernsehen.*
(After coming home he watched TV.)
(e) *Kaum daß er nach Hause gekommen war, guckte er Fernsehen.*
(As soon as he had come home, he watched TV.)

Assuming the anteriority interpretation (8b-e), how does a generation system choose among the four remaining alternatives? We argued above that we envision a modular architecture where independent sentence planning modules posit their constraints regarding tense selection, lexicalization, syntactic realization, etc.[6] In case no constraints are put forward by the sentence planning modules, *sobald* (8b) would be selected, as it is the most specific and at the same time neutral realization. If, however, a quantifier is to be included, then *nachdem* would be chosen (8c). If brevity is a stylistic concern, and the process in the subordinate clause can be deverbalized, a phrasal realization with the preposition *nach* is selected (8d). If, on the other hand, a more formal realization is the overall goal given to the generator, *kaum daß* (8e) would be chosen. In these cases, marker choice would posit constraints (as given in the combinability slot in table 2) on all other sentence planning decisions.

So far, we only considered a perfective aspect in the subordinate clause. Once we change aspect to imperfective, a realization including *nachdem* is no longer an option, compare **Nachdem er nach Hause kam, hat er Fernsehen geguckt* (After he came home, he has watched TV). *Sobald* would be an adequate realization. Likewise, changing the Aktionsart from resultative to durative, as in *Sobald er schläft, guckt*

[5] This is an abridged representation. We will eventually represent the facts as *SitSpecs* (Stede 1996), which will be annotated with temporal information. During lexicalization—as one task in the sentence planning phase—*SitSpecs* are mapped onto semantic representations (*SemSpecs*).

[6] This approach differs from Dorr and Gaasterland (1995) who impose a strict order on the selection of tense, aspect and connecting word.

sie Fernsehen (As soon as he sleeps, she watches TV) would rule out *nachdem*. With the resultative variant *einschlafen* (fall asleep) both markers are possible. Finally, if a constraint is posited that the tense has to be 'present', *kaum daß* would not be available.

5 Conclusion and Outlook

Temporal markers have neither received much attention in NLG, nor has a principled account of marker selection as such been introduced. In this paper we presented a general framework for representing German temporal markers for generation purposes. We identified some of the features required to describe applicability conditions, constraints and preferences, and proposed a declarative lexical resource that makes it possible to treat temporal markers and other linguistic means as mutual constraints at the sentence planning stage. Now, we need to examine individual temporal markers more closely and incorporate the temporal marker lexicon into a text generation system.

For the purpose of this paper, we have assumed that temporal relations are always explicitly signalled, and thus limited our study to marker selection. Marker occurrence, however, is an important issue. First, Hitzeman et al. (1995) argue that there exist temporal defaults of the kind "An event will occur just after a preceding event"; this renders the introduction of explicit markers superfluous. Second, we have only assumed pairs of time-stamped expressions, but have ignored that they usually occur in a larger discourse situation where other kinds of coherence relations might hold between events. For instance, all causal coherence relations have some temporal implicature; still, one does not want a temporal marker to signal a VOLITIONAL-CAUSE, even though cause and effect are temporally related.

Finally, future work needs to address the interaction of marker choice and temporal adverbs, as these are the means to realize the simple/progressive distinction in German.

Acknowledgement Thanks to Manfred Stede and two anonymous reviewers for helpful comments on earlier versions of this paper.

References

J. Allen. Towards a general theory of action and time. *Artificial Intelligence*, 23(2), 1984.

R. Bäuerle. Temporalsätze und Bezugspunktsetzung im Deutschen. In B. Handwerker (ed.) *Fremde Sprache Deutsch*. Tübingen: Gunter Narr, 1995.

J. Buscha. *Lexikon deutscher Konjunktionen*. Leipzig: Verlag Enzyklopädie, 1989.

H. Bussmann. *Lexikon der Sprachwissenschaft*. Stuttgart: Körner, 1990.

J. Delin, D. Scott, A. Hartley. Pragmatic congruence through language-specific mappings from semantics to syntax. In *Proc. of the 16th Conference on Computational Linguistics*, Copenhagen, 1996.

B. Dorr, T. Gaasterland. Selecting tense, aspect and connecting words in language generation. In *Proc. of the 14th International Joint Conference on Artificial Intelligence*, Montreal, 1995.

V. Ehrich. The generation of tense. In G. Kempen (ed.) *Natural Language Generation: New Results in Artificial Intelligence, Psychology and Linguistics*. Dordrecht: Martinus Nijhoff Publishers, 1987.

M. Elhadad, K.R. McKeown. Generating connectives. In *Proc. of the 13th Conference on Computational Linguistics*, Helsinki, 1990.

M. Gagnon, G. Lapalme. Prétexte: A generator for the expression of temporal information. In *Proc. of the 4th European Workshop on Natural Language Generation*, Pisa, 1993.

R. Glas. Ein Textkorpus für die deutsche Gegenwartsprache. In: *Linguistische Berichte* 40, 1975, pp 63-66.

B. Grote, M. Stede. Discourse marker choice in sentence planning. In *Proc. of the 9th International Natural Language Generation Workshop*. Niagara-on-the-Lake, Canada (to appear).

G. Helbig, J. Buscha. *Deutsche Grammatik: Ein Handbuch für den Ausländerunterricht*. Berlin, Leipzig: Langenscheidt, Verlag Enzyklopädie, 1990.

J. Hitzeman, M. Moens, C. Grover. Algorithms for analysing the temporal structure of discourse. In *Proc. of the Proceedings of the 6th International Conference of the European Chapter of the Association for Computational Linguistics*, Dublin, 1995.

N. Hornstein. *As Time Goes By*. Cambridge, Mass.: MIT Press, 1990.

A. Lascarides, J. Oberlander. Temporal connectives in a discourse context. In *Proc. of the 6th Conference of the European Chapter of the Association for Computational Linguistics*, Utrecht, 1993.

M. Moens, M. Steedman. Temporal ontology and temporal reference. *Computational Linguistics*, 14(2), 1988.

O. Rambow, T. Korelsky. Applied text generation. In *Proc. of the Conference on Applied Natural Language Processing*, Trento, 1992.

H. Reichenbach. *Elements of Symbolic Logic*. London: Macmillan, 1947.

D. Rösner, M. Stede. Customizing RST for the automatic production of technical manuals. In R. Dale et al. (eds.) *Aspects of Automated Natural Language Generation*. Berlin: Springer, 1992.

M. Stede. *Lexical semantics and knowledge representation in multilingual generation*. Doctoral dissertation. Published as Technical report CSRI-347, Dept. of Computer Science, University of Toronto, 1996.

A. Steube. *Temporale Bedeutung im Deutschen*. studia grammatica XX. Berlin: Akademie-Verlag, 1980.

K. Vander Linden, J. Martin. Expressing rhetorical relations in instructional texts: a case study of the purpose relation. *Computational Linguistics*, 21(2), 1995.

L. Wanner, E. Hovy. The HealthDoc sentence planner. In *Proc. of the 8th International Workshop on Natural Language Generation*, Herstmonceux Castle, 1996.