DANDELION

Discourse Functions and Discourse Representation:
An Empirically and Linguistically Motivated,
Interdisciplinarily-Oriented Approach to Natural Language Texts

ESPRIT Basic Research Project 6665

Deliverable R2.3.2

Discourse Grammar for German

Wiebke Ramm, Annely Rothkegel,
Erich Steiner, Claudia Villiger

September 1995

— CLS Tilburg/Nijmegen —
— University of Edinburgh GMD-IPSI Darmstadt Universität Saarbrücken —
— Universidad Complutense Madrid —

Document ID	:	Deliverable R2.3.2: *Discourse Grammar for German*
Issue date	:	September 1995
Responsible	:	Universität Saarbrücken
Version	:	Final
Status	:	Public
WP number	:	WP2: Grammar integration
Distribution	:	DANDELION Consortium
Keywords	:	Discourse grammar for German

Discourse Grammar for German

Wiebke Ramm, Annely Rothkegel,
Erich Steiner, Claudia Villiger

Universität Saarbrücken

Abstract:

This document outlines central elements of a 'discourse grammar for German'. It is structured around two complementary perspectives on the topic: one proceeds from a global perspective and starts from a text as a whole with its text-type specific properties which find their characteristic reflections on local, sentence-level discourse organization. The second perspective takes the local, sentence-level options of discourse organization as a starting point in order to specify an interface to their global, textual motivations. The aim is to integrate the two views in a common architecture.

Contents

1 Introduction

The following contains deliverable R2.3.2 Discourse Grammar German. It is structured around the two perspectives foregrounded in the description of the task in the TA (p. 53): it systematically links the current coverage of the German grammars (cf. R2.1.3. [Steiner and Ramm, 1993] and P2.1.3. [Ramm and Teich, 1994]) to discourse through an extension of the semantic interface; a perspective we sometimes refer to as proceeding *inside-out; i.e., from grammar to text.* This linkage is formulated for a specified subset of the phenomena investigated in the empirical workpackages – especially those concerned with workpackages 1.1 and 1.2 – which are here taken as a starting point for our *outside-in, i.e., text-to-grammar perspective.*

Following the description of the task in TA p.55 in particular, we are giving a specification of semantic-functional constraints on grammatical selection particularly in the area of *Theme* (cf. Section 6). We also specify, in the earlier sections, what types of constraints are required from the perspective of the empirical work in that area, and to what extent the semantic functional constraints which we formulate are necessary and/or sufficient to respond to exactly those types of information.

In terms of how this report is structured, we start with a discussion of the relationship of work here in R. 2.3.2 with previous and ongoing work, first within DANDELION, yet then also outside the project. The following section (Section 3) addresses questions of the overall architecture of the discourse grammar, and how it is reflected in the development tool. We also attempt to specify clearly where in that overall architecture our work in the present cycle is situated. Section 4 *Discourse Patterns* takes up and summarizes the global, i.e., discourse perspective on types of information relevant for our purpose. It specifies, in particular, how global aspects of texts and text grammars influence the connectedness of thematic units. The next section on *(Con-)textual Parameters and Grammatical Selections* (Section 5) continues to pursue the *outside-in, i.e., text-to-grammar perspective*, specifying how text type/discourse purpose, subject matter, and chaining strategy work together to yield characteristic combinations of features in texts, and what constraints are thereby obtained for types of thematic development in particular. These contextual constraints and their influence on lexicogrammtical features were investigated and proposed by Julia Lavid [Lavid, 1994a] and emprically tested on a variety of English texts in R1.2.2b [Lavid, 1994b]. Of particular interest will be the discussion of the relationship between domain structure in texts and textual organization. Section 6 *From Grammar to Text* takes up the *inside-out, i.e., grammar-to-text* perspective, giving a specification of a semantic interface for constraints on lexicogrammatical selection arising from the global and discourse oriented perspective. We finish the current report with a discussion of future work (Section 7).

2 Relationship of this Report with Previous and Ongoing Work

The relationship of DANDELION work in this area with other work has previously received considerable attention (cf. R. 2.2.1 [Bateman *et al.*, 1994] and relevant work discussed there).

We propose therefore that, at this point, we limit ourselves to a brief consideration of the most immediately relevant work so as to focus attention to a maximum degree. We shall, however, briefly address work both within and outside DANDELION.

2.1 Work in the project

The aim of this section is to take stock of the empirical as well as the computational work relevant for the specification of a discourse grammar for German.

2.1.1 Empirical work: thematic development and chaining

A number of previous deliverables on thematic organization, chaining strategies, and information structuring (R1.2.2a [Rothkegel and Villiger, 1993], R1.2.1 [Lavid, 1994a], R1.2.2b [Lavid, 1994b]) have collected empirical data on textual organization and have specified how that global organization finds expression across the text and in constraints on lexicogrammatical selection. For texts of various types (to appear as R1.2.2c [Villiger, 1995]), attention has also been paid to questions of comparison between the languages represented in the project. The present deliverable would now seem to be the place where that information is brought together in an architecture which goes all the way from the more global levels of text organization to the realization of that organization in lexicogrammatical systems and structures.

2.1.2 The German grammar

Apart from work on global patterns of organization, another major input coming into the German Discourse Grammar at this point, is the coverage of the current German grammar fragment as documented in R2.1.3 [Steiner and Ramm, 1993] and P2.1.3 [Ramm and Teich, 1994]. That coverage, which developed largely as a response to needs arising from the earlier periods of work, is now, in turn, a source of possibilities and constraints for the discourse grammar. One important task to be performed for the following implementation deliverable (P2.3.2) will be to critically check the scope and limits of the Theme *specification* for German in the light of its empirical applicability (especially to different text types). Correspondingly, the coverage of the *implementation* has to be checked with respect to the demands of natural texts, in particular regarding texts of different types. This task will identify the point of departure for the inside-to-outside perspective on discourse grammar specification, i.e., starting from sentence grammar options in order to meet discourse requirements (cf. Section 6). It will be seen in that part of our work that, with the formulation of essentials of a semantic interface between grammar and text, or between *form* and *function*, the requirements arising out of empirical and global considerations on the one hand, and the potential of the lexicogrammar, on the other, are brought much closer together than before.

2.1.3 The specification of a discourse grammar development tool

The specification of a discourse grammar development tool in R2.2.1 [Bateman *et al.*, 1994] payed particular attention to the user requirements that would be brought to the workbench. The specification listed the empirical areas to be covered, the constructs necessary in such a coverage, inference mechanisms, and notational devices. Our current specification of a German Discourse Grammar stays within the limits posited there. The fragment posited here is one exemplification of the meta-stratally organized workbench advocated in chapters 3 and 4 of R2.2.1. Its specification has as yet no implementation to go with it, yet this aspect should receive more attention once the resources named in chapter 8 of R2.2.1 have been brought together.

2.2 Relationship with work outside Dandelion

As before, we shall largely focus here on work which contains accounts of text structure in German, and which bears some kind of methodological affinity with our own work (cf. also our Section 6 below).

To start with, in descriptive grammars of German there are statements available about types of textual motivation and their relationship to clause structure. One example for German would be [Engel, 1988, pp.33ff]. Such collections of correspondencies between text phenomena and lexicogrammatical constructions represent an important resource for work such as ours and have been taken into account in several places. Irrespective of one's particular theoretical orientation, it seems to us that any project in the area of representation of information in text and discourse must be able to offer some degree of compatibility with the richest descriptive resources of text phenomena of the language concerned.

Another type of source for German is [Weinrich, 1993], who not only correlates text phenomena and clause grammar, but proposes a systematic view of lexicogrammar as a resource for creating texts. This systematic view shares a substantial part of its underlying philosophy with an SFL approach, but is rooted in a different tradition and, therefore, different terminology; it has to be incorporated here cautiously so as not to add terminological confusion to an already multi-theoretical project. We have discussed several of Weinrich's features in previous reports concerning work towards the discourse grammar German.

Closest to what we are attempting within our *grammar-to-text* perspective (Section 6) is probably [Martin, 1992], who proposes a fairly comprehensive treatment of many areas of English text systems. Our own work in that direction is embedded within the same methodology and framework, yet *grammar-to-text*, rather than *text-to-grammar* as in Martin, and applied to a different language.

As another relevant publication from a related area we would like to mention the recently published paper on *Discourse structure, topicality and questioning* by J. van Kuppevelt [Kuppevelt, 1995] which focuses on the organisation of dialogical discourses in terms of topics and subtopics.

There is, of course, a world of other relevant work outside DANDELION, but this work has been considered and discussed in previous reports. At this stage, when in the project everything should be brought together and when the proper focus is more important than ever, we have chosen not to add yet another survey of the relevant literature in general.

The computational work relevant at this stage has been taken into account in R2.2.1 and what has been said there would not appear to be in need of a substantial update at this point.

3 General Architecture and Working Perspectives

The term 'discourse grammar' is one of the central terms within DANDELION and, at the same time, one of the most elusive, when it comes down to specifying in detail what it should cover. This holds in particular since it is one central (and ambitious) characteristic of DANDELION not to proceed from one uniform theoretical framework but to be *multi-theoretical* in aiming at the exploitation and integration of contributions from different theoretical approaches. An important step towards the specification of a 'discourse grammar of German' is thus to identify a model of the linguistic system which is general and flexible enough to house the various methodological and theoretical paradigms under which the single contributions have been developed and which at the same time provides the basis for a meaning- and powerful architecture for interfacing them with each other.

The model we have taken as our common point of departure is a *stratificational model* of the language system – see Figure 1 – which is in relevant aspects, but not exclusively, oriented at systemic-functional linguistics. We have adapted the figure from a similar one by John Bateman who has previously used this model to localize and interrelate the various working areas in DANDELION (cf. [Bateman, 1993]). The strata correspond to different levels of abstraction presupposed by systemic-linguistic theory, where higher strata are abstractions of lower ones and lower strata are realizations of the higher ones.

We can now relate our working directions to this architecture:

- Our *outside-in* perspective to be pursued in Sections 4 and 5 proceeds 'top-down' from higher levels of semantic abstraction to lexicogrammatical realization.
- The *inside-out* perspective to be taken in Section 6 works 'bottom-up' from lexicogrammatical options at the lower levels of abstraction, through the *local semantics* interface, to their textual motivations on the higher strata.

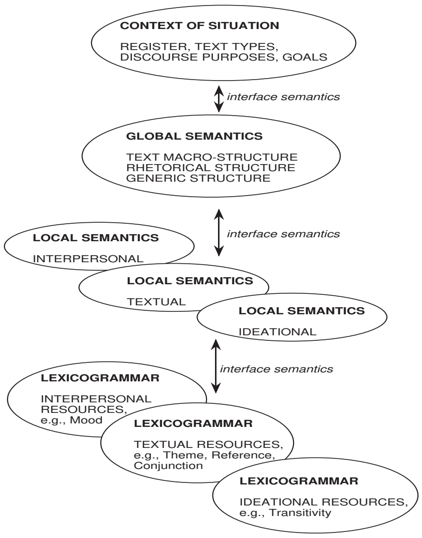

Figure 1: Language stratification

4 Discourse Patterns

In this and the following section we will proceed from a global, discourse perspective in order to identify potential components and resources for a discourse grammar responsible for representing and controlling discourse information. Since work in both sites in Saarbrücken has focused on aspects of thematic structure, the interaction of global/discourse and local/grammatical thematic structure is one question we are particularly interested in. It should be added at this point that the terms *local vs. global* are here used in two ways: one of these ways coincides with the *text vs. grammar* dimension, the categories of which are elaborated in Sections 4 and 6, respectively. The other way uses the terms *local vs. global* for subtypes of relationships *within the text level*, as in the present section.

4.1 Global and local aspects of texts and text grammars

What is the main difference between global and local aspects concerning texts? In general, 'global' means a complex framework of relationships between units which determine them together as one unit in its own right. The term 'local' concerns some effects which are restricted to relationships within a specific portion of a text, e.g. 'neighbourhood-relations'. Nevertheless it seems not to be very clear what consequences follow when distinctions such as 'global' and 'local' are applied to texts and their representation in text grammars. In addition, it will be interesting whether global aspects determine the application of specified local aspects or whether local aspects determine what a local structure finally will be! The question arises whether there are linguistic indicators which may provide some help for this distinction.

Furthermore, the term 'global' or 'global structure' is rather vaguely defined in so far as the particular kind of framework under consideration is not determined. There are at least two possibilities which again refer to two perspective in which ‘structure’ is considered:

- Global structure means some structure representing coherence/cohesion which covers the whole text (perspective of *data*).
- Global structure means a structure (hierarchy/sequence) which itself is determined by consistent categories of description and specified relationships between them (perspective of *text grammar*).

In optimal solutions both perspectives fit together: the representation itself contains a structure and at the same time, it covers all portions of the text. But when we look at real texts, then we have to cope with some mixtures of types of structures of texts. This is, e.g., the case when descriptive, narrative, and/or argumentative relationships occur in the same text. Then the normal procedure is to define one type of relationship to be the dominant one or we mix them together. Let us first consider descriptive or reporting styles of texts. Relevant relations of this kind link together two units which are connected in terms of relationships such as part-of, whole-of, property-of, instance-of, or some temporal sequences of events (pre, post, at the same time, overlapping, etc.). As neighbourhood-relations they represent local relationships. But the repetition of those relations finally constitute a global structure if it covers the whole text. This principle is also true for coherence relations such as RST-relations (e.g. a is motivation/reason/elaboration/... for b). By relating two text portions in terms of a nucleus and a satellite through some coherence relation a new portion is constructed which in turn can be a nucleus or satellite of a combination on a higher level of the hierarchy. This procedure is a bottom-up procedure and will be repeated until the whole text is described. In terms of cognition the procedure is data-driven.

On the other hand, there are top-down procedures which start from a perspective outside the text including some frameworks of relationships which derive from structurings of communicative tasks or of complex actions which are usually performed in particular situations (cf. generic definition of text types according to types of communication). Some of them seem to be represented in terms of schemas which contribute to a global structure of the whole text by repeating parts of it as well as the schema itself. The repetition of structural parts and the linguistic representation of it seems to provide us with some means for linking global and local aspects.

In the following we will analyse a German text in terms of an argumentative schema. We refer to the argumentation scheme of Toulmin [Toulmin, 1957] which has been developed for everyday argumentation. The example is a fragment of a newspaper text concerning the Golf war debate of 1991:

(1) Alle Welt hofft, daß es im Krieg um Kuweit nicht noch zu einem Einsatz von Atomwaffen kommt.
(2) Ein Beinahe-Atomkrieg ist der Golfkrieg aber schon jetzt:
(3) Die alliierte Luftflotte gibt an, das ‘irakische Atomwaffenpotential’ planmäßig zerstört zu haben;
(4) Sadam Hussein hat erklärt, die irakischen Mittelstreckenraketen könnten mit C-, B- und A-Waffen ausgerüstet werden;
(5) die israelische Regierung droht, bei einem Chemiewaffenangriff ”vernichtend” zurückzuschlagen,
(6) die amerikanische Regierung hat sich ähnlich geäußert.
(7) Noch nie seit Hiroshima war die Welt dem Einsatz nuklearer Massenvernichtungswaffen so nahe.

The Toulmin scheme relates arguments and a conclusion via a rule which allows this conclusion (generalization) and a supporting basis of norms (Figure 2).

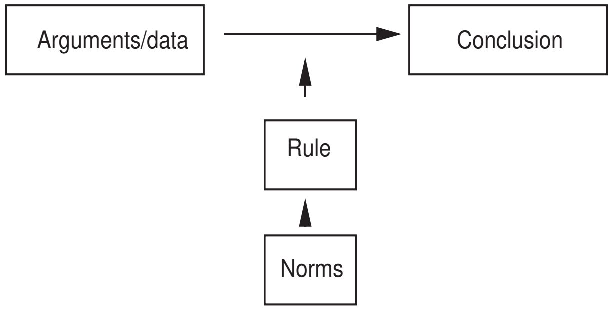

Figure 2: Argumentative Schema

Even if the 'rule' and the 'norms' are not explicitly expressed in the text (this is a very usual procedure) they are constitutive for the global structure of the text. They determine what assumptions can be interpreted as arguments and as conclusions. The argumentation is as indicated in Figure 3.

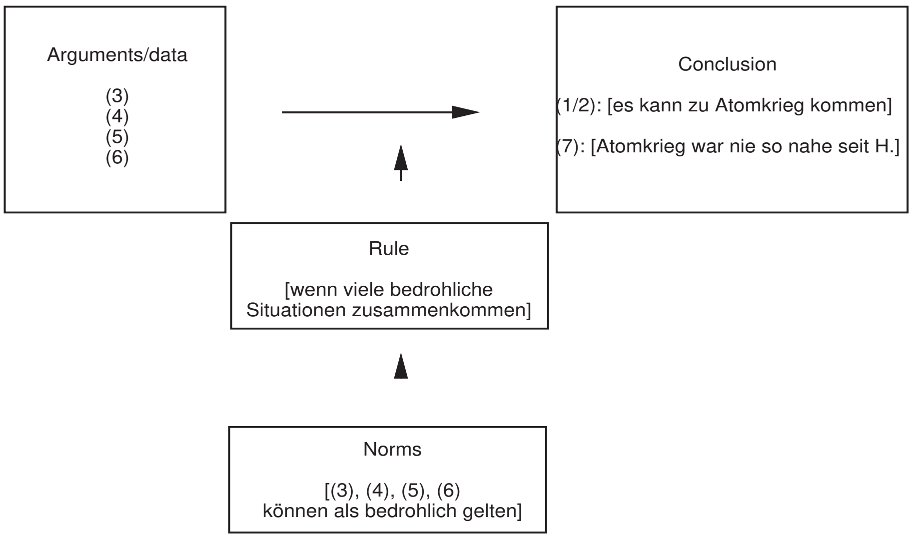

Figure 3: Argumentative Schema of the Golf war text

The following analysis tries to stress the linguistic indicators of this global structure:

	Indicator	**Type of category**
(1)	nicht noch	Conclusion
(2)	aber schon jetzt	
(3)	X1 gibt an Y1	
(4)	X2 hat erklärt Y2	Arguments
(5)	X3 droht Y3	
(6)	X4 hat sich ähnlich geäußert (droht) Y4	
(7)	noch nie	Conclusion

Indicators of the global structure

Let us now look at the local aspects of this structure. There are also repetitions in both parts of conclusion and arguments. The text form contains a frame (beginning and end of the text) which is represented by the conclusion which is repeated. On the linguistic level this is provided by the repetition of indicators of negation in combination with temporal aspects: nicht noch /aber schon jetzt ((1)/(2)) / noch nie (7). The arguments ((3)-(6)) are characterized by a parallelism of syntactic structure. By these indicators the distinction between portions of the conclusion and of the arguments is very clear also on the linguistic level.

4.2 Connectedness of thematic units

The question is what kind of impact follows from global and local aspects of the patterning of thematic units. It seems that the global structure helps to determine what the relevant units are. In an argumentative structure a high degree of relevance may be found in the conclusion. Concerning our example the main stress lies on expressions thematizing 'Atomkrieg' and its paraphrases: 'Einsatz von Atomwaffen' in (1), 'Beinahe-Atomkrieg' in (2), 'Einsatz nuklearer Massenvernichtungswaffen' in (7). They are connected in terms of global aspects through being part of the conclusion, in linguistic terms through the parallel application of negation and temporal relations. With respect to the arguments we think that the joining marker consists of some semantic property of the predicates even if the 'sentence topics' (X1 - X4) represent the participants of this war. On a syntactic level it is the parallelism of sentence structure which also marks the connectedness between the states of affairs. In this case the global aspect provides the kind of structure whereas the repetition and its linguistic reflection is provided by local aspects.

In terms of the semantic interface we shall be proposing in Section 6, the contextualization of (2) with a non-nuclear participant (more specifically, the CLASSIFIER), or the repeated contextualization of (3) to (6) with a nuclear participant (more specifically, the SENDER), are controlled by the global aspects postulated for our example here.

5 (Con-)textual Parameters and Grammatical Selections

5.1 Overview of this section

In this section we want to follow our 'text-to-grammar' perspective with a closer look on the relation between global and local text organization. Our particular interest lies in the investigation of correlations between text and text-type specific discourse parameters such as discourse purpose, subject matter, and chaining strategy and their sentence-level reflections in terms of characteristic theme and focus structures.as proposed and investigated by Julia Lavid (cf. [Lavid, 1994a]). The following section (5.2) will summarize the results of the analyses of German texts of different text types presented in [Villiger, 1995]. Four sample analyses will be repeated in order to illustrate the types of correlations found in different text types. The same 4 texts will be analysed again in Section 5.3 under a further perspective: here the point of departure is the inherent/ontological structuring of the domain underlying a text and its influence on sentence-level textual organization.

Our aim is to use these text-type specific correlations for the specification of interstratal constraints expressing the interdependency between higher-level discourse organization and lexicogrammatical realization at sentence level. The semantic interface providing the link in our current architecture is worked out in Section 6 below.

5.2 Theme, discourse topic and information structuring in German texts

In this section we want to present the results of the analysis of German texts. We concentrate on the question in how far textual parameters, especially *discourse purpose*, *text type*, *subject matter*, and *chaining strategies*, have an influence on grammatical selections of *local/sentence theme* (i.e. *contextualization* in terms of our section 6) and *sentence focus*. We first sketch out the correlations between these textual parameters and grammatical selections we found in the analysis of German text. Then we illustrate these correlations with the help of four German sample texts from various text types.

In our analysis we proceed from the assumption, that there is a dominant global goal, the *discourse purpose* underlying each text. This discourse purpose covers the major intention a writer has when producing a text (cf. [Neuendorff, 1988], [Moehn, 1991]). *Text types* are prototypical patterns of global text organization, with which a particular discourse purpose can be linguistically realized. The text typology we apply is based on text-internal criteria (cf. [Linke *et al.*, 1991]) which refer to the form or content of texts (eg. subject(s) of the text, development of the global theme, patterns of text structure, lexical peculiarities). The typology consists of five text types [1]: narrative, descriptive, argumentative, expository, and instructive text type.

Apart from the discourse purpose and the corresponding text type the *subject matter* of a text has an influence on the text structure (cf. [Lavid, 1994b]). For this reason we selected texts which belong to the same text type but cover various subject matters (cf. Table 1

[1]This division in five types is also employed by [Hansen, 1989], [Virtanen, 1992], [Werlich, 1979]

below).

The *chaining strategy* employed in a text provides a specific line which steers the reader through the text. Following the proposal presented in [Lavid, 1994a] we assume that particular chaining strategies are applied, when particular discourse purposes are pursued and particular subject matters have to be unfolded. Table 1 (the information presented in the table is adopted from [Lavid, 1994a]) shows the correlations we found in our German sample texts concerning text type, subject matter and chaining strategy.

Discourse purpose/ text type	**Subject matter**	**Chaining strategy**	**Category of chaining strategy**	**Texts**
narrative	events	temporal	time	history of cities
narrative	participants	participant	topical	short biographies
descriptive	place relations	spatial	location	passages from travel guides
argumentative	complex facts and ideas	through-argument or counter-argument	conjunctive	newspaper articles
expository	classes of objects, generic concepts	characterization	topical	encyclopedia entries
instructive	steps in a procedure	sequential	time and/ or process	recipes chapters from instruction manuals

Table 1: Correlations of discourse parameters

The *global theme* of a text is defined in accordance with the specific chaining strategy of the text. These elements of a text, which relate the incoming information to what is already known, provide orientation to the reader. On the other hand *local themes* or *sentence themes* are defined according to the criteria presented by [Steiner and Ramm, 1993]. The determining factor is the position of the text elements in a sentence or a clause. This definition of local/sentence theme closely corresponds to the notion of *Vorfeld* (eg. [Hoberg, 1981]. It was our interest to analyse the correlation between global and local themes in texts of various text types. Further we were interested in the correlations which exist between the metafunctional type of local themes (*experiential*, *interpersonal*, or *textual*) and text types in our sample texts. In terms of our semantic interface of Section 6 below, it is the **Category of chaining strategy** in Table 1 which links up directly with the most delicate classes in our semantic interface for theme choice.

A further issue was to investigate the effect of global textual parameters on the realization of *sentence focus*, which is understood as those concepts in a sentence/clause on which the speaker wants to draw the attention of the hearer (cf. [Lavid, 1994a]). Our Section 6, while not concentrating on sentence focus itself very extensively, does discuss *focussing* as a semantic correlate. We adopted the focus categories introduced by [McCoy and Cheng, 1990]. They distinguish between object, attribute, setting, event and action nodes. We added the category of fact nodes as it was proposed by [Lavid, 1994a]. In the following, we illustrate the results of the analysis of narrative, descriptive, expository, and argumentative texts through presenting a sample analysis for each text type.

5.2.1 Narrative texts: 'Ferdinand VII'

Characteristically, texts of the narrative text type are always (implicitly or explicitly) temporally organized (cf. [Faigley and Meyer, 1983]). For the analysis of texts belonging to this text type we selected two short biographies and two texts which were concerned with the history of cities (see also [Lavid, 1994b]). In three of the four texts the global organization of the text was attained by the means of a temporal global chaining strategy. The temporal elements which build up the global chain are in boldface in the sample text below.

Ferdinand VII.

(01) Ferdinand VII., König (1808 und seit 1814), geb. San Ildefonso **14.10.1784**, gest. Madrid **29.9.1833**;
(Ferdinand, King (1808 and since 1814), b. San Ildefonso 14.10.1784, ob. Madrid 29.9.1833)
(02) Sohn KARLS IV. und MARIA LUISAS von Parma.
(Son of KARLS IV. and MARIA LUISAS of Parma.)
(03) **1807** ließ sich Ferdinand von Gegnern des Min. GODOY in Beziehungen zu Frankreich verwickeln.
(In 1807 Ferdinand became involved in relations with France by adversaries of minister Godoy.)
(04) Diese 'Kronprinzverschwörung von El Escorial' und das Zerwürfnis des span. Hofes erleichterte
(This 'Crown Prince Conspiracy' of El Escorial and quarrels at the Span. Court facilitated the occupation)
NAPOLEON I. **1808** die Besetzung Spaniens, die über die 'Revolution von Aranjuez' (Ausrufung
(of Spain by NAPOLEON I. in 1808, which lead from the 'Revolution of Aranjuez' (Ferdinand)
Ferdinands zum König, Abdankung des Vaters im März) zur Entthronung der Bourbonen im Mai 1808
(proclaimed king, abdication of the father in March) to the dethronement of the Bourbons in Bayonne)
in Bayonne (Thronverzicht Ferdinands und seines Vaters) und zum spanischen Unabhängigkeitskrieg führte.
(in May 1808 (abdication of the throne by Ferdinand and his father) and to the Spanish War of Independence.)
(05) Ihn verlebte Ferdinand, vom Volk als Symbol des Freiheitswillens angesehen und herbeigesehnt (deshalb
(During the war, Ferdinand lived at the French Valencay Palace considered and longed for by people as a)
Beiname 'el Deseado', 'der Herbeigewünschte'), auf dem französischen Schloß Valencay.
(symbol of the claim for freedom (therefore the appellation 'El Deseado', the desired).
(06) Erst im **März 1814** kehrte Ferdinand zurück, hob am 4.5. die Verfassung von 1812
auf und regierte als absoluter Herrscher.
(Only in March 1814 Ferdinand returned, abolished the Constitution of 1812 and ruled as absolute monarch.)
(07) Die Liberalen ließ er mit größter Härte verfolgen, die Mönchsorden, Inquisition und die
(The liberals were persecuted with toughness, the monastic orders, the inquisition and torture)
Folter wiederherstellen;
(were restored;)
(08) die liberale Revolution von **1820** wurde mit französischer Hilfe niedergeschlagen **(1823)**.
(the liberal revolution of 1820 was suppressed with help of the French (1823).)
(09) **Nach dem Tod seiner dritten Frau**, der sächsischen Prinzessin JOSEPHA **(gest. 1829)**,
(After the death of his third wife, the Saxon Princess JOSEPHA (ob. 1829) Ferdinand married)
heiratete Ferdinand MARIA CHRISTINA, Tochter FRANZ' I., des Königs
(MARIA CHRISTINA daughter of FRANZ I, King of both Sicilies, who gave birth to the)
beider Sizilien, die ihm **1830** die Thronerbin ISABELLA II. gebar.
(heir of the throne ISBELLA II in 1820.)
(10) Um dieser die Thronfolge zu sichern, hob er durch die Pragmatische Sanktion **(1830)** das
(In order to ensure her succession to the throne, he abolished the Sal. law of succession through)
sal. Erbfolgerecht auf.
(the 'Pragmatic Sanction' (1830))
(11) Die Thronansprüche seines Bruders Don CARLOS führten zu den für CARLOS
(The claim to the throne by Ferdinand's brother DON CARLOS lead to the 'Karlistenkriegen' which)
erfolglosen Karlistenkriegen.
(were unsuccessful for CARLOS.)
(12) Wegen Ferdinands schwerer Krankheit leitete seit **1832** seine Frau die Regierungsgeschäfte.
(Because of Ferdinand's serious illness his wife led the government since 1832.)

The discourse purpose of the text is the presentation of the life of *Ferdinand VII* and the historical events in which he was involved. The text is temporally structured by presenting the dates of historical events (cf. boldface in text above and column global chaining strategy in Table 2). All text segments with the exception of 02, 05, 07 and 11 contain temporal elements. These temporal elements ensure the progression of the text. The global pattern of the text (cf. Figure 4) which we developed following [Heinemann and Viehweger, 1991] illustrates the parts into which the text can be divided.

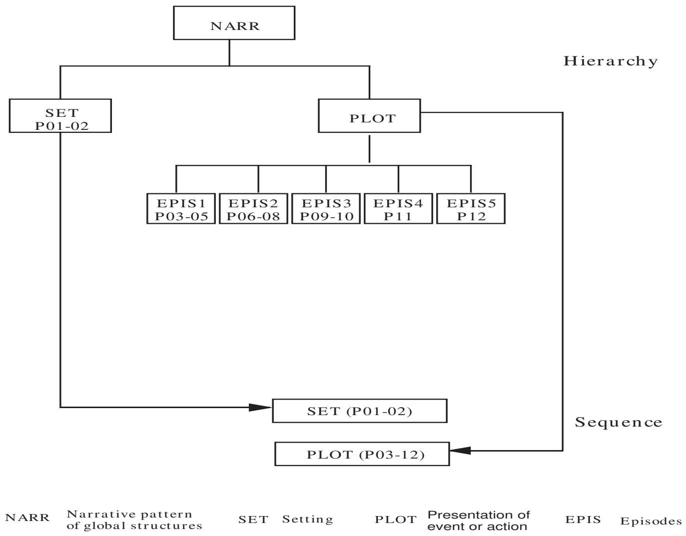

Figure 4: Global pattern of the text 'Ferdinand VII'

In 01 the period of time which is covered by the text is determined (see also time axis in Figure 12 in Section 5.3). The dates of birth and death of Ferdinand demarcate beginning and end of the temporal chain, which is developed in the text. Text segments 01 and 02 form the unit setting (SET) of the text. The date *1807* in local theme position signals the beginning of the first episode (EPIS1), which covers text segments 03-05. The next temporal text element in local theme position *Erst im März 1814* designates the beginning of the second episode (EPIS2, text segments 06-08). The third episode (EPIS3, text segments 09 and 10) is introduced by the temporal element *Nach dem Tod seiner dritten Frau*, which is a component of the global chain, but also an element of the local chain (Table 2 column local chaining strategy). This local chain is created by referring to Ferdinand and members of his milieu. It basically warrants the coherence of the text, while the global chain ensures the progression of the text. In one of the sample texts (biography of 'Heinrich Böll') the topical chain was considered as the global chain, which chiefly organizes the text. This chaining strategy was probably preferred, because in case of the inherent structure of the subject matter (Life and especially work of Heinrich Böll) the use of a topical chaining strategy was likely. However, in the text *Ferdinand VII* and in the other texts of the narrative text type the global organization of the texts is attained by means of the temporal chaining strategy.

Similar to all other texts of this text type the category of all local themes is simple and experiential (cf.Table 2). As the subject matter of the text (presentation of historically relevant aspects of the life of Ferdinand VII) offers sufficient possibilities to structure the text with experiential themes (especially the themes of the temporal chain) there is no need to use textual or interpersonal themes.

The density of information is partly very high in the foci of the text segments (especially 04, 06, and 09). As was to be expected in a narrative text, all foci are classified as event nodes with the exception of the focus in 02, which we classified as an object node (cf. Table 2). The focus tree of this text again illustrates the structure of the text emphasized above (cf. Figure 5).

Obviously, the text segments 03, 04 and 05, text segments 06, 07, 08 and text segments 09 and 10 form units. In all English and German texts we have analysed the focus trees of narrative texts consist of an event mother node which branches out in event daughter nodes.

Parts of Table 2 will reappear again in Section 6, but then with the categories of our textual semantic interface added to replace the reference to the *grammatical category of theme* still employed here.

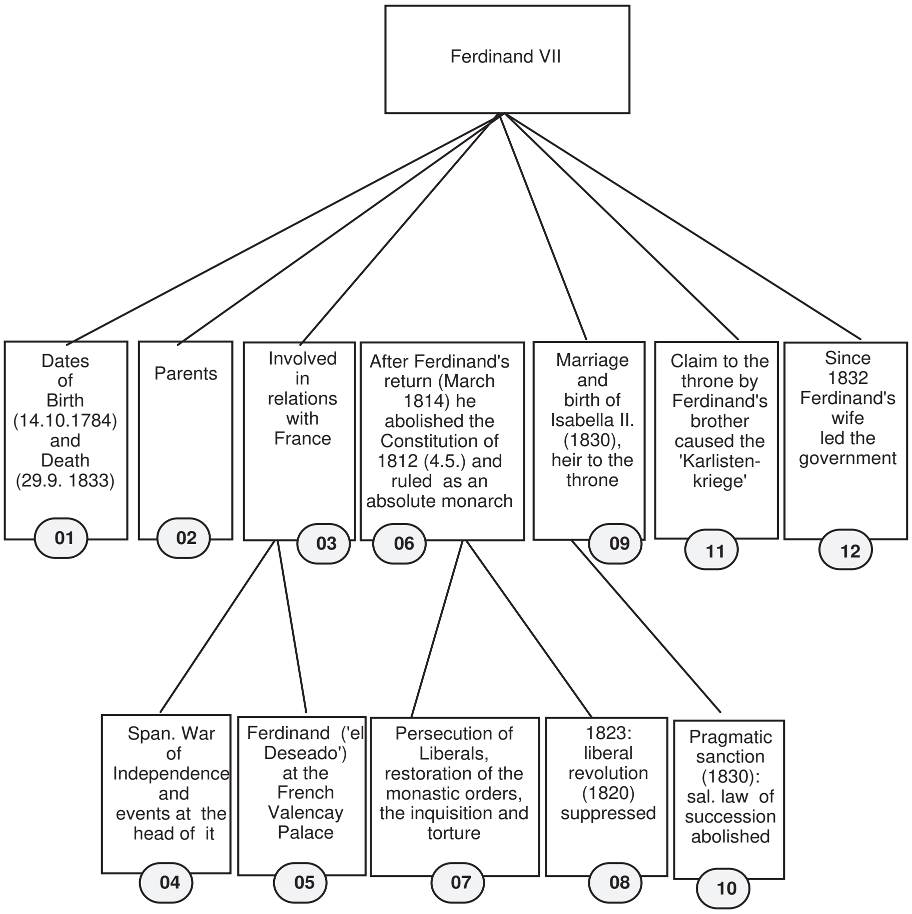

Figure 5: Focus tree of the text 'Ferdinand VII'

	Text level				Sentence level				Correlation
	Realization of global chaining strategy	**Category of global chaining strategy**	**Realization of local chaining strategy**	**Category of local chaining strategy**	**Local theme**	**Category of local theme**	**Focus**	**Focus category**	
(01)	14.10.1784 ... 29.9.1833	time	Ferdinand VII.	top	Ferdinand VII., König (1808 und seit 1814)	sim, exp	Date of birth (14.10.1784) and death (29.9.1833)	event	p: gc-f p: lc-lt
(02)			Sohn	top	Sohn Karls IV. und Maria Luisas von Parma	sim, exp	Parents	attr	non
(03)	1807	time	Ferdinand	top	1807	sim, exp	Involved in relations with France	event	gc-lt p: lc-f
(04)	1808	time	Ausrufung Ferdinands ... Thron- verzicht Ferdinands	top	Diese ‘Kronprinz- verschwörung von El Escorial’ und das Zerwürfnis des span. Hofes	sim, exp	Span. War of Independence and events at the head of it	event	p: gc-f p: lc-f
(05)			Ferdinand	top	Ihn	sim, exp	Ferdinand (‘el Deseado’) at the French Valencay Palace	event	p: lc-f
(06)	März 1814	time	Ferdinand	top	Erst im März 1814	sim, exp	After Ferdinand’s return (March 1814) he abolished the Constitution of 1812 (4.5.) and ruled as an absolute monarch	event	p: gc-lt p: lc-f p: gc-f
(07)			er	top	Die Liberalen	sim, exp	Persecution of Liberals, restoration of the monastic orders, the inquisition, and torture	event	p: lc-f
(08)	1820 (1823)	time time			die liberale Revolution von 1820	sim, exp	1823: Liberal revolution (1820) (1820) suppressed	event	p: gc-lt p: gc-f
(09)	Nach dem Tod seiner dritten Frau ... (gest. 1829) ... 1830	time time	seiner Ferdinand	top top	Nach dem Tod seiner dritten Frau, der sächsischen Prinzessin JOSEPHA (gest. 1829)	sim, exp	Marriage and birth of Isabella II. (1830), heir to the throne	event	p: gc-lt p: gc-f p: lc-lt p: lc-f
(10)	(1830)	time	er	top	Um dieser die Thronfolge zu sichern	sim, exp	Pragmatic sanction (1830): sal. law of succession abolished	event	p: gc-f p: lc-f
(11)			seines Bruders	top	Die Thronan- sprüche seines Bruders Don CARLOS	sim, exp	Claim to the throne by Ferdinand’s brother caused the ‘Karlistenkriege’	event	p: lc-lt
(12)	seit 1832	time	Ferdinands seine Frau	top top	Wegen Ferdinands schwerer Krankheit	sim, exp	Since 1832 Ferdinand ’s wife led the government	event	p: gc-f p: lc-lt p: lc-f

top = topical, sim = simple, exp = experiential, attr = attribute; non = no correlation, (p:) gc-lt = (Partial) correlation of global chaining strategy and local theme, (p:) gc-f = (Partial) correlation of global chaining strategy and focus, (p:) lc-lt = (Partial) correlation of local chaining strategy and local theme, (p:) lc-f = (Partial) correlation of local chaining strategy and focus

Table 2: Chaining strategies, local theme, and focus

5.2.2 Expository texts: 'Drongos'

The discourse purpose of expository texts is definition and/or explanation and/or illustration of objects or facts. Related text types are the descriptive and the argumentative text type. E.g., [Werlich, 1979]'s examples for expository texts are definition, explication, abstract, interpretation, and essay. As examples for expository texts we have chosen definitions (encyclopedia entries). The analysis of English and German expository texts revealed that the most common global chaining strategy used in definition is the *characterization strategy* (cf. R1.2.2b [Lavid, 1994b] and R1.2.2c [Villiger, 1995]). This strategy is also applied in the sample text below.

(01) **Drongos** (Dircruridae) sind star- bis hähergroße Singvögel mit kurzen Beinen, breitem,
(Drongos are song-birds from starling size to jay size with short legs, broad and strong beak and forked tail,
kräftigem Schnabel und Gabelschwanz vorwiegend altweltlich-tropisch (Afrika, Südasien,
Australien und Neuginea) verbreitet.
ranged mainly in tropical part of the Old World (Africa, South-Asia, Australia and New-Guinea).)

(02) **Die 20 Arten der Familie** unterscheiden sich durch Vorhandensein oder Fehlen von
Hauben, verlängerten Schwanzfedern und gekräuselten oder samtigen Schmuckfedern.
(The 20 species of the family differ in presence or absence of crests, lengthened tail feathers and curled or velvet plumes.)

(03) **Das Gefieder** ist meist schwarz mit mattem grünem oder stahlblauem Glanz.
(The plumage is mostly black with matt green or steel blue luster.)

(04) **Die Drongos** sind Wartenjäger,
(Drongos are perchers,)

(05) **sie** fangen Insekten im Fluge oder nehmen sie Würgern ähnlich vom Boden oder von
(they hawk for insects or pick them from the ground like shrikes or from twigs)
Zweigen ab.

(06) In der Brutzeit leben **sie** paarweise in abgegrenzten Revieren, die sie sehr heftig gegen
(In breeding time they live in pairs in demarcated breeding areas, which they vehemently defend
ihresgleichen und gegen andere, größere Vögel verteidigen.
against her own kind and against other bigger birds.)

(07) **Sie** bauen freistehende Schalennester und legen 2-5 gefleckte Eier.
(They build exposed bowl nests and lay 2-5 spotted eggs.)

(08) **Männchen und Weibchen** kümmern sich um die Pflege der Brut.
(Male and female birds take care of the brood.)

(09) Neben verschiedenartigen, meist rauhen Rufen verfügen **einige Arten** auch über einen,
(Apart from various, mostly hoarse calls, some species also have a loud twittering song at their disposal.)
lauten zwitschernden Gesang.

(Lexikon der Tiere, S. Fischer 1970)

In this text the discourse purpose is to explain the relevant features of Drongos. The characterization strategy is applied to organize the global pattern of the text. It starts with a general definition (GENER) of the object (01) (cf. Figure 6), which refers to genus, characteristic features of body constitution and geographical distribution of Drongos.

The object Drongos is introduced in this text segment in local theme position. In 02 the family of Drongos is addressed, and in 03 a whole-part relation is set up by describing the plumage of Drongos. Apart from recurrence in the definite form (04), pronominalization (05-07) and substitution (08) are employed to create the topical chain.

Features (FEAT) mentioned in the text are ontology (FEAT1 in 02), the plumage (FEAT2

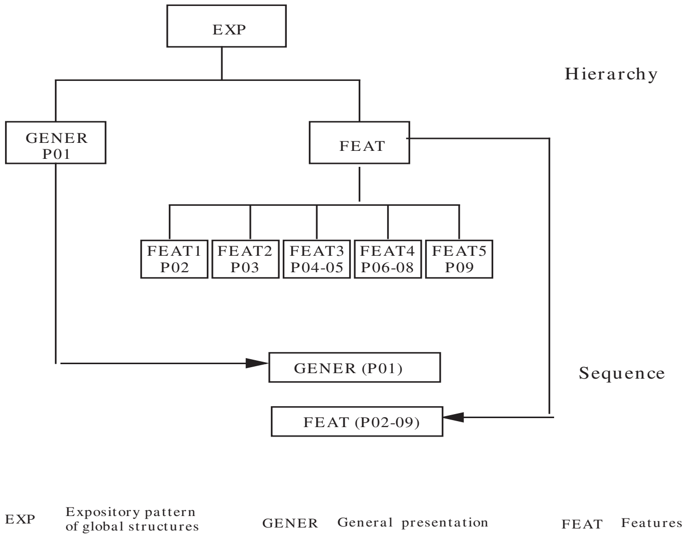

Figure 6: Global structure of the text 'Drongos'

in 03), food and manner of hunting (FEAT3 in 04-05), breeding (FEAT4 in 06-08) and the calls and songs (FEAT4 in 09) of Drongos (cf. (6)). These features of the part FEAT could also be arranged in a different way, because the subject matter does not demand for a special structuring of the text. Only the features mentioned at the beginning of the text (genus and biodiversity) seem to be some kind of conventionalized at least in German definitions of Drongos [2].

Concerning the relation of the global chaining strategy and the selection of local themes we notice an almost complete correlation of the element which expresses the global chaining strategy and the local theme of a sentence (cf. Table 3). Only in the last segment there is no correlation. The reader is lead along this line of local themes. Either the subject of the text is repeated or new features (02, 03, 09) are mentioned in local theme position. There is no need to employ textual or interpersonal themes for structuring the text. Therefore the category of all local themes is simple and experiential.

The focus mother node of this text is the object *Drongos* and the children nodes are all attribute nodes of this mother node (Figure 7). Correspondingly focus shift candidates for the sentence foci are attributes from the object *Drongos.*

The density of information is partly very high. In text segment 01 e.g., information about three features (size, genus characteristic features of body constitution, and geographical distribution) is presented, or in 02 biodiversity and further characteristic features of Drongos are explained. This probably is a result of the little space, which is provided to present the features of Drongos.

[2]In other German encyclopedias the definition of Drongos also starts with these features

	Text level		Sentence level				Corre-lation
	Realization of global chaining strategy	**Cate-gory of global chaining strategy**	**Local theme**	**Category of local theme**	**Focus**	**Focus cate-gory**	
(01)	Drongos	top	Drongos	sim, exp	General description of appearance	obj	gc-lt
(02)	Die 20 Arten der Familie	top	Die 20 Arten der Familie	sim, exp	Distinctive features of species	attr	gc-lt
(03)	Das Gefieder	top	Das Gefieder	sim, exp	Colour of plumage	attr	gc-lt
(04)	Die Drongos	top	Die Drongos	sim, exp	Percher	attr	gc-lt
(05)	sie	top	sie	sim, exp	Food and manner of hunting	attr	gc-lt
(06)	sie	top	In der Brutzeit	sim, exp	Demarcated breeding area, defence against other birds	attr	gc-f
(07)	Sie	top	Sie	sim, exp	Nests	attr	gc-lt
(08)	Männchen und Weibchen	topic	Männchen und Weibchen	sim, exp	Parental care in breeding time	attr	gc-lt
(09)	einige Arten	top	Neben verschieden-artigen, meist rauhen Rufen	sim, exp	Calls and song	attr	p: gc-f

top = topical, sim = simple, exp = experiential, non = no correlation, obj = object, attr = attribute, (p:) gc-lt = (Partial) correlation of global chaining strategy and local theme, (p:) gc-f = (Partial) correlation of global chaining strategy and focus

Table 3: Chaining strategies, local theme, and focus

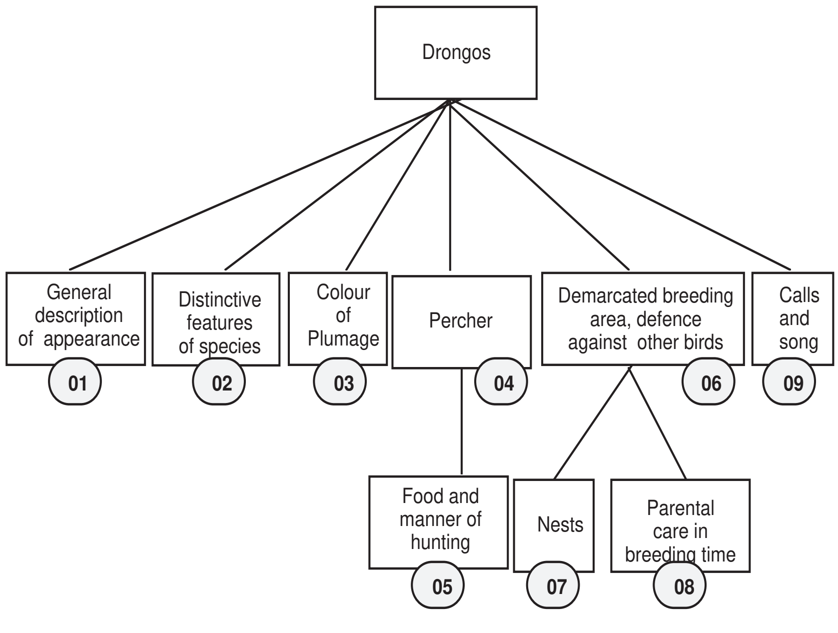

Figure 7: Focus tree for the text ‘Drongos’

5.2.3 Descriptive text: 'Seville'

For the descriptive text type we selected texts describing place relations (see also the analysis of English descriptive texts [Lavid, 1994b]). The global text structure of the texts presented a spatial chaining strategy.

(01) **Sevillas Zentrum** liegt **östlich eines Seitenkanals des Rio Guadalquivir**, der
(The Centre of Seville is situated east of a side canal of the Rio Guadalquivir)
die Stadt etwa **in Nord-Süd-Richtung** durchzieht.
(which runs through the city roughly from north to south.)
(02) Hauptstraße ist die Avenida de la Constitucion;
(The main street is the Avenida de la Constitucion.)
(03) **in ihrer unmittelbaren Umgebung** liegen mit Kathedrale und Giralda sowie der
(In its immediate surroundings, the most important sights of the city,)
Alcazaba die bedeutendsten Sehenswürdigkeiten der Stadt.
(the cathedral, the Giralda, and the Alcazaba, are situated.)
(04) **Östlich** schließt sich das Barrio de Santa Cruz an, Sevillas lauschiges Vorzeigeviertel.
(In the east, the Barrio de Santa Cruz, Seville's secluded showpiece quarter, borders on the city.)
(05) Die Avenida de la Constitucion beginnt **im Süden** am Verkehrsknotenpunkt Puerta de Jerez
(The Avenida de la Constitucion begins in the south at the Puerta de Jerez junction)
und mündet **im Norden** in den Doppelplatz Plaza San Francisco/Plaza Nueva;
(and leads into the double square Plaza San Francisco/Plaza Nueva in the north.)
(06) **hier** liegt auch das Geschäftsviertel um die Haupteinkaufsstraße Calle Sierpes.
(Here also the shopping centre around the main shopping street, Calle Sierpes, is situated.)
(07) **Südlich des engeren Zentrums** erstrecken sich der Park Parque de Maria Luisa
(South of the immediate centre the park Parque de Maria Luisa and)
und das Weltausstellungsgelände von 1929, die Plaza de Espana.
(the site of the world fair 1929, the Plaza de Espana, are located.)
(08) **Jenseits des Gualdalquivir** sind zwei ehemals selbständige Siedlungen zu abendlichen
(Beyond the Guadalquivir two formerly independent settlements)
und nächtlichen Anlaufadressen avanciert: das volkstümliche Barrio de Triana
(have developed into places to go to in the evenings and at night: the traditional Barrio de Triana,)
auf Höhe des Zentrums und, **südlich anschließend**, das neuzeitlichere
(which is on a level with the centre and, bordering on this area in the south,)
Barrio de los Remedios auf Höhe des Parque de Maria Luisa.
(the more modern Barrio de los Remedios, which is on a level with the Parque de Maria Luisa.)

(T. Schröder: Andalusien, S. 332/333)

In the sample *Seville* above the subject matter of the text is the centre of Seville. It is the discourse purpose of this text to describe this object. The centre of Seville is introduced in text segment 01. In the same segment two further spatial clues are introduced: *östlich eines Seitenkanals des Rio Guadalquivir* and *in Nord-Süd-Richtung*. Through these spatial elements the global chain of the text is anchored. Consequently 01 forms the unit setting (SET) in the graphical presentation of the global pattern of the text (Figure 8).

The units in which the single properties of Seville are presented are designated as PROP in Figure 8. The spatial chain is pursued through the whole text and ends with the spatial elements *Jenseits des Gualdalquivir*, *auf Höhe des Zentrums* and *südlich anschließend* which ly beyond the area, which was demarcated in the first text segment (see also Figure 12 in Section 5.3). This signals that the description of the centre of Seville is terminated and that the discourse purpose is attained.

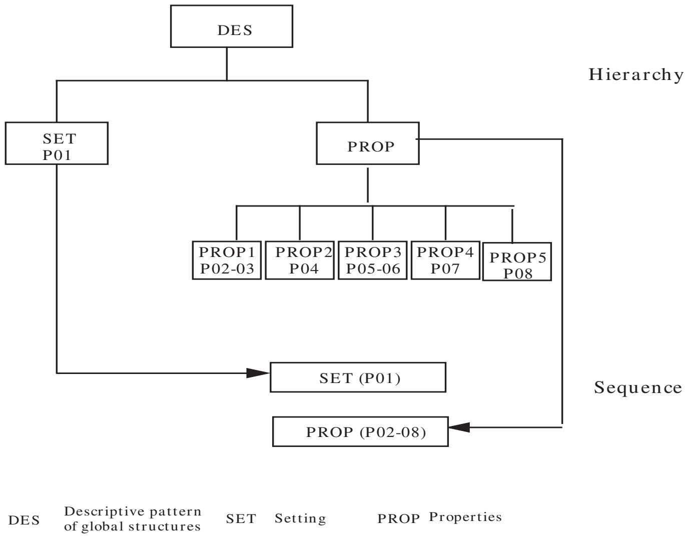

Figure 8: Global structure of the text 'Sevilla'

Most spatial elements of the text are presented in local theme position. Therefore, the local theme category of all text segments is simple and experiential. Text segments which do not contain such a spatial local theme are linked with text segments which deliver such a reference point (text segment 02 and 03 and text segment 05 and 06).

As was to be expected in a descriptive text the mother focus node of the text is an object node and all other foci are attribute nodes referring to this object node (Figure 9).

While local themes consist of simple elements referring to a landmark, the foci of the text segments are often more complex (cf. focus tree above and Table 4). In 03, e.g., the various sights in the surroundings of the main street (cathedral, Giralda and Alcazaba) are presented. Since foci also contain further landmarks, there is not only a correlation of the elements of the global chaining strategy and local themes, but also a partial correlation of elements of this chain with foci (cf. Table 4, 01, 05, and 08).

The analysis of this descriptive text illustrates that the discourse purpose (describe an object) and the subject matter (presentation of the centre of Seville) not only have an influence on the global structure of the text, but that they also have consequences for the organization of the text at sentence level. While sentence themes figure as pointers and indicate landmarks, foci present information referring to these landmarks.

5.2.4 Argumentative texts: 'Nostalgiekarte Jahrgang 1992'

In argumentative texts it is the discourse purpose to substantiate an assertion which is presented in the text (cf. e.g. [Heinemann and Viehweger, 1991],[Herbig, 1992]). Therefore, a conclusion is presented and arguments which give the reasons for this conclusion are specified. The sample text below is an article published in the commentary part of a German newspaper. Similar to the descriptive sample text, spatial information (political geography

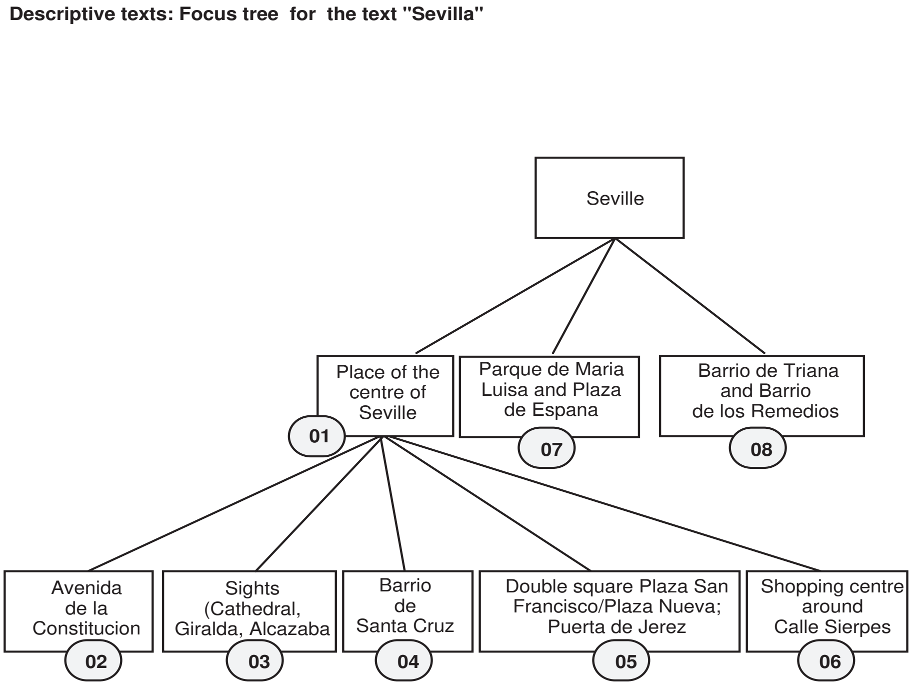

Figure 9: Focus tree of the text 'Sevilla'

	Text level		Sentence level				Corre-lation
	Realization of global chaining strategy	**Cate-gory of global chaining strategy**	**Local theme**	**Category of local theme**	**Focus**	**Focus cate-gory**	
(01)	Sevillas Zentrum ... östlich eines Seitenkanals ... Nord-Süd-Richtung	spat spat spat spat	Sevillas Zentrum	sim, exp	Place of the centre of Seville	obje	gc-lt p: gc-f
(02)			Hauptstraße	sim, exp	Avenida de la Constitucion	attr	non
(03)	in ihrer unmittel-baren Umgebung	spat	in ihrer unmittel-baren Umgebung	sim, exp	Sights (Cathedral, Giralda, Alcazaba)	attr	gc-lt
(04)	Östlich	spat	Östlich	sim, exp	Barrio de Santa Cruz	attr	gc-lt
(05)	im Süden ... im Norden	spat	Die Avenida de la Constitucion	sim, exp	Double square Plaza San Francisco/Plaza Nueva; Puerta de Jerez	attr	gc-f
(06)	hier	spat	hier	sim, exp	Shopping centre around Calle Sierpes	attr	gc-lt
(07)	Südlich des engeren Zentrums	spat	Südlich des engeren Zentrums	sim, exp	Parque de Maria Luisa and Plaza de Espana	attr	gc-lt
(08)	Jenseits des Gualdalquivir ... auf Höhe des Zentrums ... südlich anschließend	spat spat spat	Jenseits des Gualdalquivir	sim, exp	Barrio de Triana and Barrio de los Remedios	attr	gc-lt p: gc-f

spat = spatial, sim = simple, exp = experiential, obj = object, attr = attribute, non = no correlation, gc-lt = Correlation of global chaining strategy and local theme, (p:) gc-f = (Partial) correlation of global chaining strategy and focus

Table 4: Global chaining strategy, local theme, and focus

of Germany) has to be presented, but as the discourse purpose is argumentative, this is done in a different way.

(01) Nostalgie-Karte Jahrgang 1992
(Nostalgia-Map of the Year 1992)
(02) So war die politische Geographie einmal zu fernen Zeiten.
(This is how the political geography used to be a long time ago.)
(03) Deutschland noch nicht vereint,
(Germany not yet united,)
(04) der Saar-Lor-Lux-Raum ein weißer Fleck auf der Landkarte.
(the Saar-Lor-Lux-area a white spot on the map.)
(05) Zu fernen Zeiten?
(A long time ago?)
(06) Mitnichten!!!
(Far from it!!!)
(07) Die oben abgebildete Deutschlandkarte findet sich im neuen Taschen-Terminkalender
(The map of Germany shown above is published in the new pocket-agenda 1992)
1992 der Sparkasse Saarbrücken.
(of the savings-bank Saarbrücken.)
(08) Dort hat man noch nicht mitbekommen,
(Where no-one has **so far** *noticed)*
(09) daß Deutschland um die Kleinigkeit von fünf neuen Bundesländern größer geworden ist.
(hat Germany has grown by the trifling bit of five new Bundesländer.)
(10) **Zudem** gibt es jenseits der alten DDR-Grenze noch andere Städte als Leipzig und Berlin,
(Moreover, there are still other cities apart from Leipzig and Berlin beyond the former GDR-frontier,)
so zum Beispiel Rostock, Dresden, Magdeburg oder Saarbückens Partnerstadt Cottbus.
(so for example, Rostock, Dresden, Magdeburg, or Cottbus, the twin city of Saarbrücken.)
(11) **Außerdem** scheint den Herren von der Sparkasse entgangen zu sein,
(Besides , it seems that the gentlemen of the savings-bank didn't realize)
(12) daß am Ende des Terminello-Jahres 1992 der europäische Binnenmarkt steht.
(that at the end of the Terminello-Year 1992 the Single European Market will come into force.)
(13) **Nicht zuletzt** vermittelt das Kärtchen den Eindruck,
(Last but not least the little map suggests,)
(14) daß Saarbrücken der Nabel (Alt-)Deutschlands zu sein scheint.
(that Saarbrücken was the navel of (the former) Germany.)
(15) Je nach anatomischer Sichtweise kann es aber auch ein anderes Körperteil sein.
(But depending on the anatomical point of view it might also refer to another body part.)

(Saarbrücker Zeitung 14/14.12.1991, p. 5)

According to the Toulmin-scheme presented in 4.1 the global pattern of the text (cf. Figure (10)) consists of the parts conclusion (CONC) and arguments (ARGU). The latter is further divided into the pro-argumentative (PRO) and the counter-argumentative part (CONTRA). The conclusion of the text is realized implicitly and has to be reconstructed with the help of indications in the text. In this text a clear indication is given by the juxtaposition of *Nostalgie* and *Jahrgang 1992* in the headline (01) of the text. If a map is nostalgic, it must be in some respects not current with the reality it should represent. The counter-argumentative part of the text (02-07) constists of two arguments which present the perspective of the savings bank. Additionally, this part of the text has an embedding function (cf. [Brinker, 1992]). It presents the specific knowledge that is needed to understand the following pro-argumentative part. Therefore, in the counter-argumentative part simple and experiential themes are dominant. They create a topical local chain (cf. Table 5 column *local chaining strategy*). The textual theme in 02 refers to the following text segments and signals the

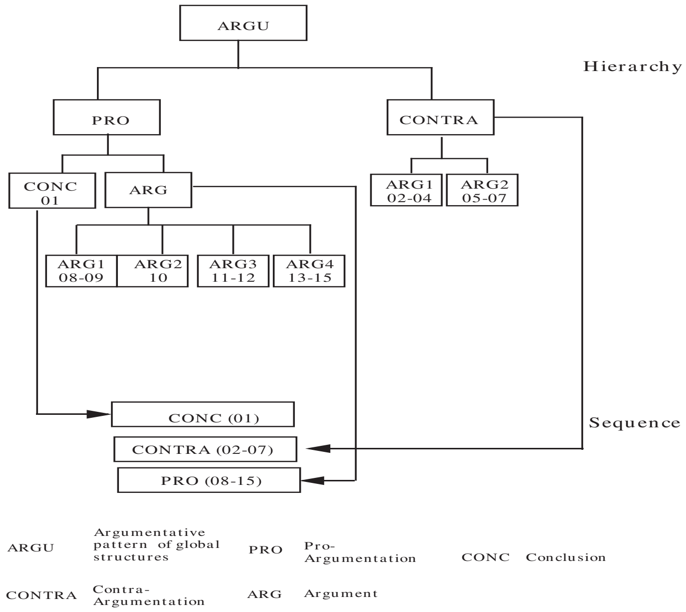

Figure 10: Global structure of the text 'Nostalgie-Karte Jahrgang 1992'

beginning of the counter-argumentative part 02.

In the pro-argumentative part of the text (08-15) arguments, which reflect the opinion of the author are presented. In this part the structuring of the text is characteristic for argumentative texts. In 10, 11, and 13 textual themes in local theme position create the global chain. These conjunctive expressions *Zudem*, *Außerdem*, and *Nicht zuletzt* enumerate the arguments which are set against the perspective of the savings bank presented in text segments 02-07.

Concerning focus and focus shift candidates all foci are classified as fact nodes (cf. Table 5). The focus tree (see Figure 11) again illustrates the global structure of the text. While the focus mother node of 02 with the children nodes 03 and 04 reflects the perspective of the savings bank, the focus mother node of 06 with it's children nodes 10, 12, 13, 14 represents the foci in which the author argues against the perspective of the saving banks.

In this argumentative text the global structure of the text is created with the help of textual themes in local theme position. They are necessary to attain the argumentative discourse purpose of the text. On the other hand the first part of the text, the counter-argumentative part, is structured with the help of a local chain of simple and experiential themes which arise from the subject matter of the text.

5.2.5 Correlations between discourse parameters and the selection of sentence theme and sentence focus

The analysis of four sample texts of various text types presented the correlations which exist between discourse parameters such as discourse purpose, subject matter, and chaining strategy on the one hand and the selection of sentence theme and sentence focus on the

	Text level				Sentence level				Corre-lation
	Realization of global chaining strategy	**Cate-gory of global chaining strategy**	**Realization of local chaining strategy**	**Cate-gory of local chaining strategy**	**Local theme**	**Category of local theme**	**Focus**	**Focus cate-gory**	
(01)			Nostalgie-Karte Jahrgang 1992	top	Nostalgie-Karte Jahrgang 1992	sim, exp	Nostalgia Map	fact	non
(02)					So	sim, tex	Past political geography	fact	non
(03)			Deutschland	top	Deutschland	sim, exp	Unity of Germany	fact	lc-lt
(04)			der Saar-Lor-Lux-Raum	top	der Saar-Lor-Lux-Raum	sim, exp	Importance of Saar-Lor-Lux area	fact	lc-lt
(05)							Relation of past and actual		non
(06)							political geography		non
(07)			Die oben abgebildete Deutschland-karte	top	Die oben abgebildete Deutschland-karte	sim, exp	Source of map	fact	lc-lt
(08)			Dort	top	Dort	sim, exp	Facts not considered in the map	fact	lc-lt
(09)					daß Deutschland	sim, exp	Five new Bundesländer	fact	non
(10)	Zudem	coun			Zudem	sim, tex	Cities beyond the former border to the GDR	fact	gc-lt
(11)	Außerdem	coun			Außerdem	sim, tex	Responsibility for the map	fact	gc-lt
(12)					am Ende des Terminello-Jahres	sim, exp	Introduction of the Single European Market	fact	non
(13)	Nicht zuletzt	coun			Nicht zuletzt	sim, tex	Impression which the map provides	fact	gc-lt
(14)			Saarbrücken	top	Saarbrücken	sim, exp	Overestimation of Saarbrücken	fact	non
(15)					Je nach ana-tomischer Sichtweise	sim, exp	Correct estimation of Saarbrücken	fact	non

coun = Counter-argument: Thesis-Antithesis, topi = topical, sim = simple, exp = experiential, tex = textual, non = no correlation, gc-lt = Correlation of global chaining strategy and local theme, (p:) gc-f = (Partial) correlation of global chaining strategy and focus,

Table 5: Chaining strategies, local theme, and focus

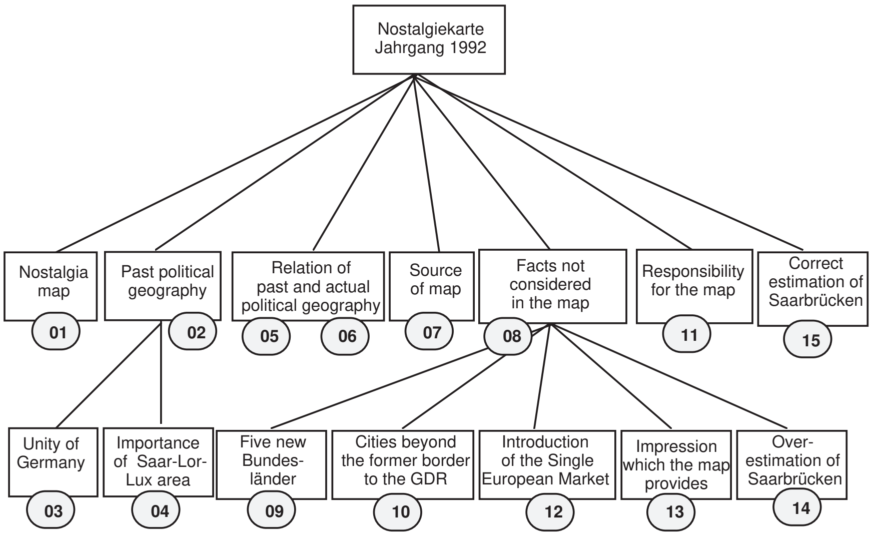

Figure 11: Focus tree of the text ‘Nostalgie-Karte Jahrgang 1992’

other hand. While in the narrative, the descriptive, and the expository sample texts simple experiential themes were selected, the argumentative text was at least partly structured with the help of simple textual themes, which were realized in local theme position [3]. In all texts the organization of the global structure was reflected by way of elements of the particular global chain in local theme position. Corresponding to the specific subject matter of the texts, particular sentence foci were selected (see also section 5.3).

5.3 Domain structure and text organization

5.3.1 The conceptual structure of the sample texts

In this section, our focus will be on the role of *conceptual information* and its lexicogrammatical realization in different types of texts. We will have a closer look at the inherent, ontological structuring of the domain of each of our four sample texts in order to see how the particular 'subject matter' the text is about influences its organization and how conceptual structure in combination with discourse purpose correlates with the specific textual parameters investigated above. In the following we want to analyse in more detail what the specific conceptual organization of the texts is and how it surfaces linguistically in terms of textual parameters such as theme.

5.3.2 Narrative text: 'Ferdinand VII'

As the analyses of English as well as German narrative texts have shown (cf. R1.2.2b [Lavid, 1994b] and R1.2.2c [Villiger, 1995]) *temporal structuring* is a central characteristic of this type of text which can be reflected linguistically, for instance, by a temporal chaining strategy, the selection of temporal elements as theme and the selection of events as focus nodes. According to these features, also our 'Ferdinand' text is a typical narrative text (see again Table 2 in Section 5.2).

Figure 12[4] gives a graphical representation of the organization of the knowledge domain underlying the text[5]. Its conceptual information basically consists of event concepts associated with certain participants. The concepts have an inherent temporal ordering along a time axis (roughly covering the time span between 1784 and 1833). In addition to temporal ordering there are other relations expressing some kind of causality between the events (such as 'cause', 'enablement', 'purpose', etc.), linking events to causal chains spanning 2 to 3 sentences. These types of conceptual relations are indicated in the figure by arrows (↓) (without further differentiating between them). The chains of events structure the domain into 5 chunks of information which are perceived in the global text structure as the episodes the plot of the text consists of (cf Figure 4). Within the sentences, the causal relations are frequently realized as theme-rheme (or theme-focus) patterns as for instance in (04), (09),

[3]texts of the instructive text type often show a similar realization of sentence themes

[4]We would like to thank Eva Treppmann for preparing the English versions of the domain structure representations of the four sample texts

[5]The numbers in grey circles relate to the clause numbering of the text.
Concepts printed in *italics* refer to lexical material taken from the source texts.

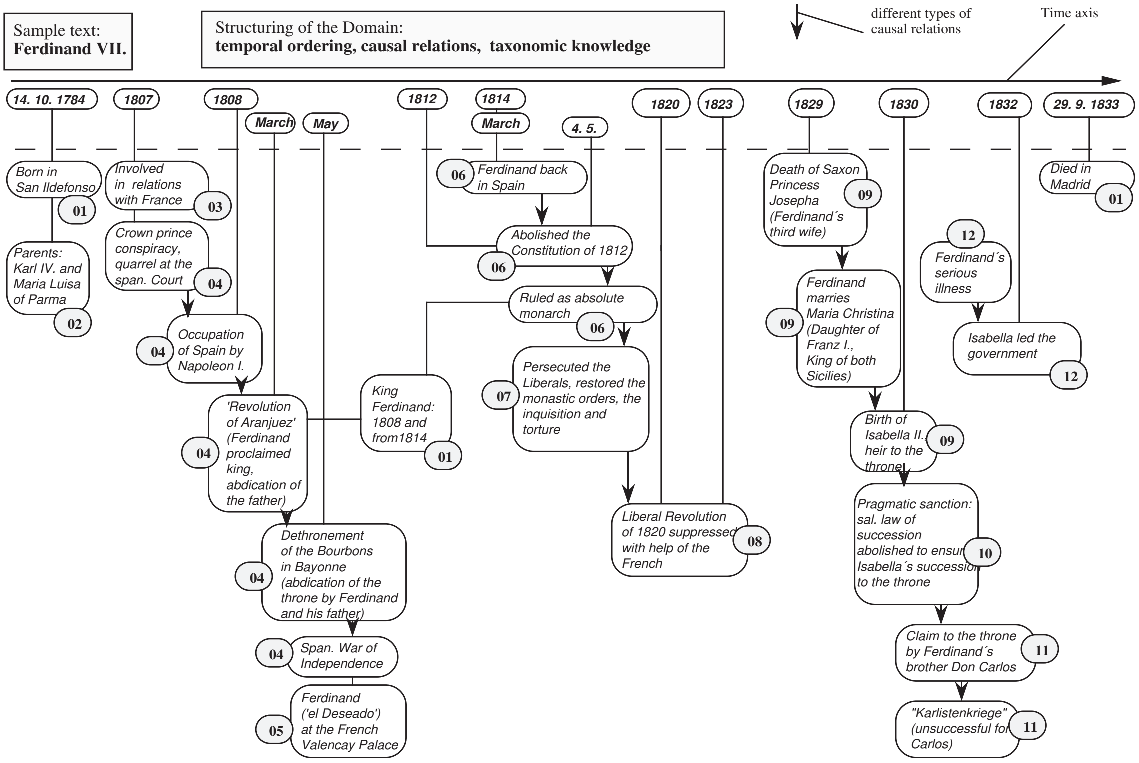

Figure 12: Domain structure of 'Ferdinand VII' text

(10), (11), or (12) and are explicitly indicated by appropriate lexical and grammatical means (e.g., the infinitival clause with 'um zu' in (10), the prepositional phrase in (12) and the lexical meaning of the process as in (04), 'erleichtern', or (11), 'führen zu'). Causality between the sentences has rather to be inferred by applying general knowledge of the world and history as in (06)/(07), (07)/(08), (09)/(10), and (10)/(11). As a further type of conceptual information necessary for the coherent interpretation of the text we should mention taxonomic knowledge about family, monarchical, and governmental relations.

5.3.3 Expository text: 'Drongos'

Our expository sample text, an entry in a German lexicon of animals describing 'Drongos', a species of song birds, deals with *taxonomic information* about the subject matter. This concept taxonomy of the underlying domain, formally a specialization hierarchy of attributes characterizing the concept 'Drongos', is clearly reflected in the structure of the text. The information presented in the text relates to biological features and characteristics of the species. Each of these features (such as size, genus, body constitution, geographical distribution, food, reproduction, and voices) is dealt with in 1 – 3, mostly adjacent, sentences. The thematic structure of the majority of the sentences is organized such that the theme refers to 'Drongos', either recurrently as in (01) and (04), via pronominalization as in (05) and (07), or by means of substitution with related concepts as in (02) and (08). The elaboration of the various properties of the concept 'Drongos' is performed in the rhematic part of the sentences. Only

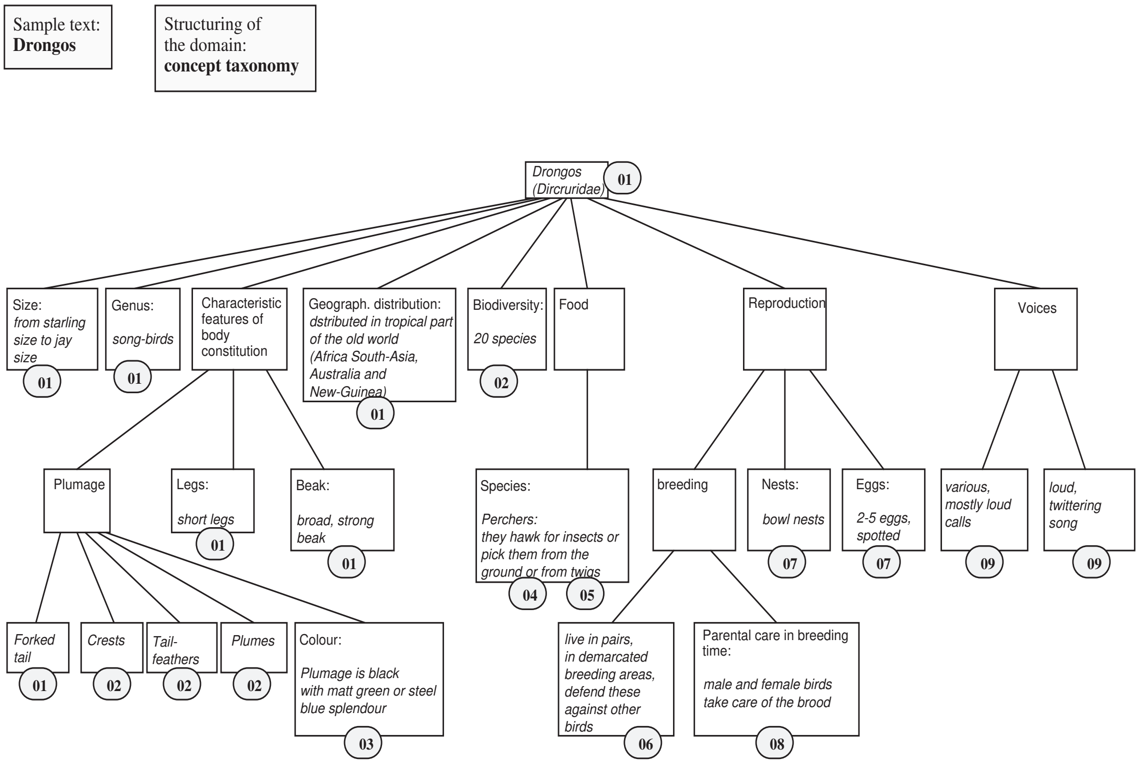

Figure 13: Domain structure of 'Drongo' text

in (06) and (09) the global theme is not sentence-thematic but some attribute (concerning breeding and calls, respectively) associated with it. To sum up, in this text we can observe a clear correlation between conceptual and textual structure: the text starts with introducing the top-level domain concept and develops by traversing conceptual links connecting general with more specific attributes.

5.3.4 Descriptive text: 'Sevilla'

The most characteristic feature of this text is its reflection of the *spatial organization* of the underlying domain. As already pointed out in [Lavid, 1994b], this is a typical property of many descriptive texts and is in this case leading to the incremental creation of a *cognitive map* of the domain, 'centre of Seville'. The centrality of the domain structure for the construction of the meaning of the text is mirrored in its linguistic appearance, also with respect to thematic choice: the sentence themes in this text constantly are *ideational/experiential* elements realized as subject theme ((01) and (05)), subject complement (02), or circumstantial ((03), (04), (06), (07), and (08)). In each of the clauses the rhematic part (at the same time representing the focus) elaborates on the specific spatial concept introduced as theme, i.e., adds certain *attributes* in form of other spatial concepts in order to build up a spatial representation of the domain. The sequence of concepts verbalized as theme allows the reader of the text to navigate through a cognitive map of the domain by strictly keeping to a *spatial chaining strategy*. As with the expository sample text, we thus get a relatively

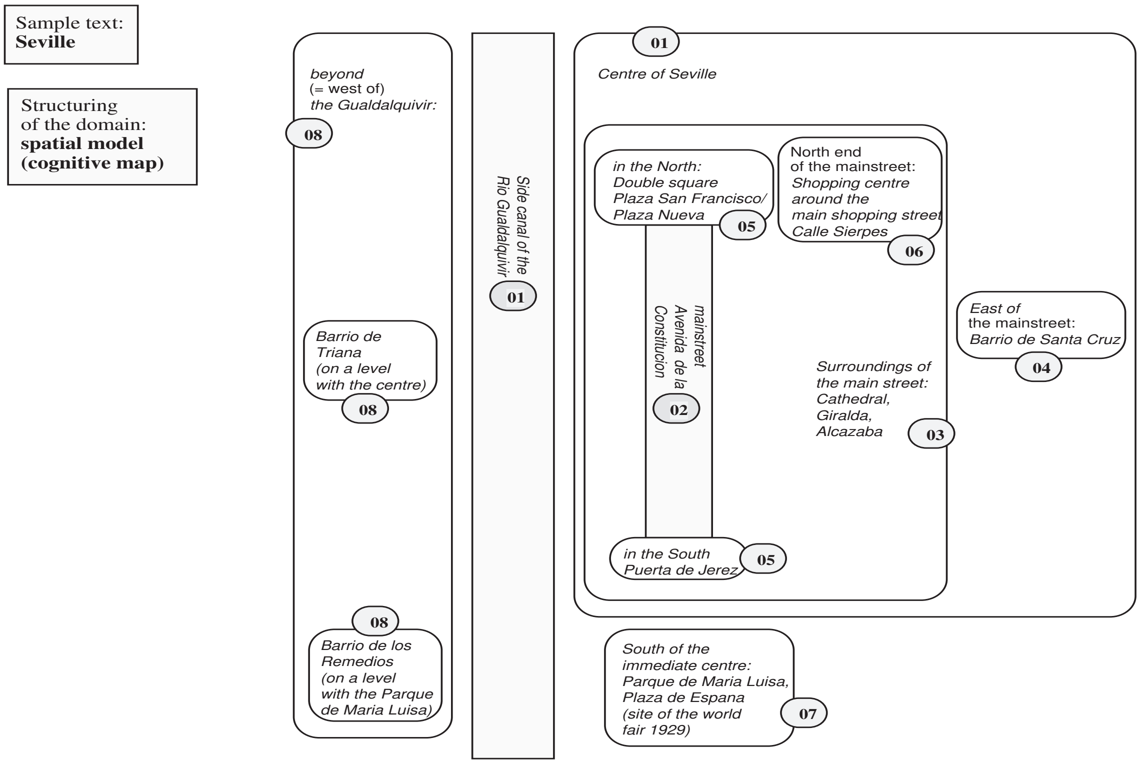

Figure 14: Domain structure of 'Sevilla' text

direct correlation between text and domain structure.

5.3.5 Argumentative text: 'Nostalgiekarte Jahrgang 1992'

The last sample text is argumentative, a satirical article published in the commentary part of a German newspaper.

This text deals with spatial information represented in form of some map as well but here the domain is not the main responsible for the structuring of the text. In this case the underlying discourse purpose is not to inform the reader about some state of affairs as in the previous examples but rather to argue in favour of some opinion taken by the author, i.e., it is argumentative. This is clearly reflected in the linguistic structure of the text: propositions (02) – (07) represent the *contra-argumentation*, the presentation of facts/arguments against which the author is going to argue[6]. (08) – (15) develop the author's opinion, the *pro-argumentation*. The *conclusion*, a further constitutive building block of argumentative texts, in this text is presented in its headline. The linguistic means exploited differ significantly from those in the descriptive text example: regarding theme choice, a clear distinction can be made between the counter-argumentative ((02) – (07)) and the pro-argumentative part ((08) – (15)) of the text: in the former *ideational/experiential* themes are predominant,

[6]For a detailed description of patterns of global text organization in different types of texts see [Rothkegel and Villiger, 1993; Rothkegel *et al.*, 1994]

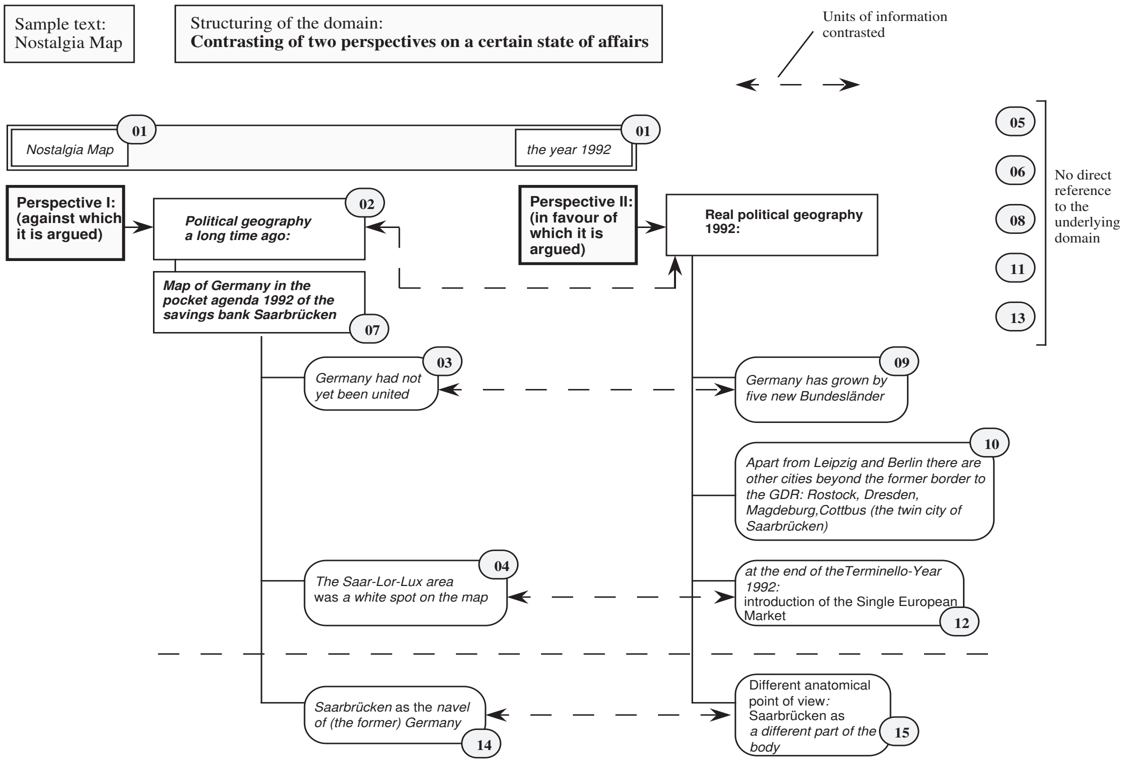

Figure 15: Domain structure of 'Nostalgiekarte' text

e.g., in (03), (04), and (07). Only in (02) we find a *textual* theme, a conjunctive element, cataphorically referring to the following propositions. In (05) and (06), in our interpretation, theme is elided, presupposing something like 'War es so .....' (06) and 'So war es ...' (07), respectively. In the pro-argumentation, in contrast, *textual* themes (realized by conjunctive elements), representing an argumentative chain predominate as in (10), (11), and (13). An interpretation would be that the task of the contra-argumentative discourse segment is to present the background information (which is inherently ideational) on which the following pro-argumentation builds. Thus, in the first segment the local contextualization - what sentence theme is responsible for - is mainly ideational whereas in the following it is basically textual: in this segment the author is going to articulate his/her position. In order to reach this communicative effect each of the arguments are explicitly integrated into an *argumentative chain* signalled by appropriate textual elements ('zudem' – 'außerdem' – 'nicht zuletzt'). To sum up, in this text it is the *line of argumentation* which structures the discourse which is guided by *interpersonal* rather than ideational information. In order to convey this type of textual meaning, the author has to make extensive use of textual devices such as connectors some of which moreover appear in theme position.

5.3.6 Conceptual relations and textual coherence

We will now try to compare the results of the domain structure analyses of the four sample texts. What is quite obvious is that the underlying domain does not play the same role in the four sample texts. In general, we can say that the argumentative sample text behaves differently regarding the underlying conceptual information communicated than the other three texts. In the former it is the line of argumentation, i.e. the specific *perspective on the subject matter chosen by the author* that structures the text as a whole. In the latter group of texts it is the *naturally given organization of the domain* itself (or at least a certain aspect of it) that is reflected in the structuring of the text. In terms of metafunctional types of information, we can say that the narrative, the expository, and the descriptive text is basically structured by *experiential* information, whereas in the argumentative sample text the role and perspective of the author as an *interpersonal* parameter becomes relevant for the organization of the text. This influences, for instance, the sentence-thematic organization: experiential themes are highly predominant in the first three text types whereas textual themes are used to ensure a coherent interpretation of the argumentative text.

6 From Grammar to Text

In this section, we give an outline of a semantic interface between the grammatical systems that we developed for the German grammar (cf. [Steiner and Ramm, 1993; Ramm and Teich, 1994]) and the textual (and in that sense functional) motivations for *theme selection* that arise out of empirical work (cf. particularly DANDELION empirical deliverables under R 1.1. and R 1.2.). We begin with a brief review of generalized *grounding devices* on sentence level, one of which is *Theme* as discussed here (Section 6.1). After thus briefly considering the context of our present work, we shall proceed in Section 6.2 to a review of how, in our current architecture, *empirical* or *textual* motivations for Theme Choice are linked to grammatical realization through a *semantic interface*, constituting the *local semantics* in terms of Figure 1 in Section 3. In the following subsection (6.3), we shall first repeat our central *grammatical* Theme system, that of *Theme-Type*, and then develop a first semantic system for Theme in German, constituting the interface between textual motivation and clause grammar.

6.1 Grounding devices on sentence level

The semantics of *Theme* may be seen to be systematically related to a few other areas in the following way (cf. TA pp.32ff, [Steiner, 1991]): the way how ideational and interpersonal information and interaction are presented in texts is dependent on a small number of dimensions along which chunks are related to each other. Two of these dimensions (*focussing* and *thematizing*) are part of the *text base*, whereas two others (*ranking* and *taxis*) are situated by Halliday and Matthiessen within the ideation base (cf. [Halliday and Matthiessen, in press]). All of them, though, are concerned with creating background-foreground structures in discourse and text:

- **Focussing:** Some element of information associated with a situation to be encoded is focused relative to some other elements. Focus here is a semantic category, in parallel, but not in a one-to-one relationship with, the grammatical term *new* as distinguished from *given*. Focus may be understood in contextual terms as one or more of *new, surprising, narrowing of attention, contrast to expectation, saliency in a number of ways*

... or some more elements of a potentially open list. This area is situated in the text base proper and is most directly realized in the grammar of German in systems of *information distribution* and, in interaction with those, *Theme*. *Focus* as used here is a semantic category, parallel to the category *contextualization* used in Section 6.3.2 below. The more delicate distinctions within that category will be the particular *focus categories* referred to in Section 5.2, such as *object, attribute,..* etc. Their place in a *focussing system*, if we were to write it, would be, approximately, where the most delicate categories are in our current Theme semantics (Figure 17) (cf. features *participant* or *circumstance* in that system).

- **Ranking:** Each element in a situation (events, entities) can be encoded in different places of a part-whole hierarchy, i.e. as process, complement, circumstance etc. in terms of the clause grammar (cf. [Bateman and Paris, 1989]) , cf. (1) to (3):

  (1) The fire destroyed the buildings.

  (2) The cause of the destruction of the buildings was the fire.

  (3) Due to the fire, the buildings collapsed.

  In SFL the existence of such correlated ways of grammaticalization is often referred to as one kind of *grammatical metaphor*, and it is of particular relevance in an explanation of translation and other processes in multilingual text production. What this kind of textual organization does is to provide a *background-foreground* structure in terms of part-whole relationships between events. One notion of *information focus*, particularly important in translation (cf. [Doherty, 1991]) may be related to this distinction, rather than to any thematic or given vs. new distinctions. It is, for example, in this sense, and in this sense only, that in many constructions German tends towards *end shift of information*, whereas English tends towards *mid shift of information.* Ranking is situated in the ideation base in [Halliday and Matthiessen, in press] and is most directly realized grammatically in the areas of *transitivity, diathesis, aktionsart.*

- **Taxis:** Taxis provides a third type of grounding distinction, this time in terms of a *dependency structure* and in terms of *head vs. dependent* in an SFL sense. To continue our examples (1) to (3) above, we get (4) and (5):

  (4) Because there was a fire, the buildings collapsed.

  (5) There was a fire. So the buildings collapsed.

  Parallel to *Taxis*, we get the systems of clause combining relations under *expansion* and *projection*, which we shall not go into here (cf. Dandelion's coherence relations as in R1.1.2a [Bateman and Rondhuis, 1994]). The ideationally-oriented ones of these are situated in the ideation base, forming a sort of generalized logico-semantic organizing principle for the language as whole. Another subclass of these relations, those associated with the structure of discourse and argumentation, are best situated in the text base.

- **Thematizing:** This is the fourth type of grounding, and the one we are mainly concerned with here. It consists in *contextualizing* the message currently in progress, i.e. in making something in a clause *thematic*, rather than *rhematic*. Obviously, a *grammatically thematic* and *semantically contextualizing* variant of (4) above would be (6):

(6) The building collapsed because there was a fire.

For each of the four types of grounding mentioned above, there is one directly related type of realization (*intonation, constituency, (systemic) dependency, linear precedence*, respectively), but other lexicogrammatical resources are available as well, most notably *lexis*, and all of these types of realization interact in obvious ways. A particularly interesting interaction in the case of *Theme* is, of course, that between intonation and linear precedence.

According to the view expressed in this section, we would expect two major systems in the text base *focusing, thematizing/ contextualization*, with our particular attention going out to the system of *contextualization* at the moment. These would interact heavily with ranking and taxis/logico-semantic relations. Another major part of the text base will be taken up by the semantics of cohesion, but this is largely outside the focus of the present attempts.

6.2 Textual motivations for Theme choice and the need for a semantic interface

The German grammar work in the project so far has focused on thematic options. The grammatical systems implemented in the grammars of that area together with the inquiries addressing the text plan are formulated in the specifications and the implementation report of the German grammars (R2.1.3. [Steiner and Ramm, 1993] and P.2.1.3. [Ramm and Teich, 1994]). We shall here repeat one particularly comprehensive inquiry for the sake of illustration:

Choice: options under EXPERIENTIAL:
Inquiry: Do you want to focus the hearer's attention strongly on referent (I),
and the REGISTER/MODE/MEDIUM of discourse is WRITTEN
(i.e. simply using stress with unmarked word order is not available);
or do you want to downgrade the current proposition;
or do you want to present your thematic information as focused at the same time;
or do you want to focus on new information;
or focus on a contrasting semantic role of referent (I);
or do you want to introduce a sub-set of Referent (I);
or do you encode R(I) as *unexpected*;
or do you want to contrast or strongly identify R(I);
or do you want to correct some aspect of R(I) or ascribe a salient property to R(I);
or do you want to introduce, resume, or end R(I) as a paragraph topic?
Then: classify ENGLISH-THEME as PREDICATED and
ONE CONSTITUENT and
GERMAN-THEME as PREDICATED or NON-PREDICATED, OTHER..

Observe that, whether or not the classificatory outcome here will be *predicated* or otherwise, at this point in the classification, we already know that we are dealing with an entity which is *thematic* (in one of the senses of that term), and which is *experiential* as well, because the grammar at that point specifies experiential themes only, i.e. previous features are inherited.

Because of previous work in the empirical parts of the project, and because of the additional sources where we took the information about textual motivation for our particular grammatical choice from, we can be confident about a few properties of the conditions used in our inquiry:

As these types of motivation do not stem from ad-hoc introspection, nor from some isolated examples, but from extensive textual and psycholinguistic experiential studies (for detailed references cf. R.2.1.3.), we may assume that they systematically correlate as textual motivations with the choice of the particular grammatical construction of *predicated theme* in English, or *predicated Theme* or *non-predicated, other* in German (all of which would be *marked themes*, cf. German grammar). In other words, there is a good chance that they will not represent isolated and idiosyncratic cases. On the other hand, depending on the text type investigated, and depending on the delicacy of our types of textual motivation, our list of textual motivations seems relatively open and also fairly *indirectly related* to any individual grammatical construction, such as *predicated theme* here. In order to close the methodological gap indicated by these facts, our architecture contains a level of *semantic systems* which are abstracted out from various interrelated areas of grammar, yet still more generalized than any list of textual motivations for a lexicogrammatical construction.

We also assume that there is a many to many relationship between grammatical and semantic categories, i.e., that any given grammatical construction is related as co-classification with a set of semantic categories, each of which, in turn, will have lexicogrammatical correspondences (or realizations) other than the one just under investigation.

Our inquiry addresses the following textual notions: being thematic, focus of hearer's attention, written vs. spoken medium (and here we might add other register factors, such as high vs. low degree of formality, or a particular macro- or micro-level goal of the passage which we are just in), downgrading of a proposition, current vs. non-current proposition, i.e. one of the dynamic aspects of discourse, focus of information by speaker, new information, (contrast) and semantic role, referent, sub-set of a (class of) referents, being (un-)expected, strong identification of a referent, correcting, ascription of a salient property, introduction, resumption or ending of a paragraph topic ... In other words, coming from inside out, i.e. from grammar through semantics to text, we can identify a set of categories which our semantics of Theme should realize, and we can state that, while the grammatical construction of predicated theme is thus shown to have a wide range of semantic correlates, and en even wider range of textual motivations, it may safely be assumed that hardly any of these motivations will affect the choice of Theme only. In particular, we already know that all of these interact in intricate ways with word order in general, and with intonation, to name only two other grammatical areas.

6.3 Representation

We shall initially repeat our maybe central grammatical classification of clauses in terms of Themes in our system network of *Theme Type* (Figure 16). After this, we shall give an outline of the semantics of Theme in German, based on our earlier DANDELION work in grammar. The semantic categories developed there should be the ones figuring in the text plan. The semantics, as an *interlevel* (cf. [Halliday *et al.*, 1964]) faces two ways: inwards as an abstraction from lexicogrammatical categories, and outwards as a generalization of individual functional categories for text building of the type addressed in inquiries.

A methodological principle following from this is that the categories we are developing here are *semantic* categories, with the realization statements coming in terms of grammatical classes.

6.3.1 The grammar of Theme for German

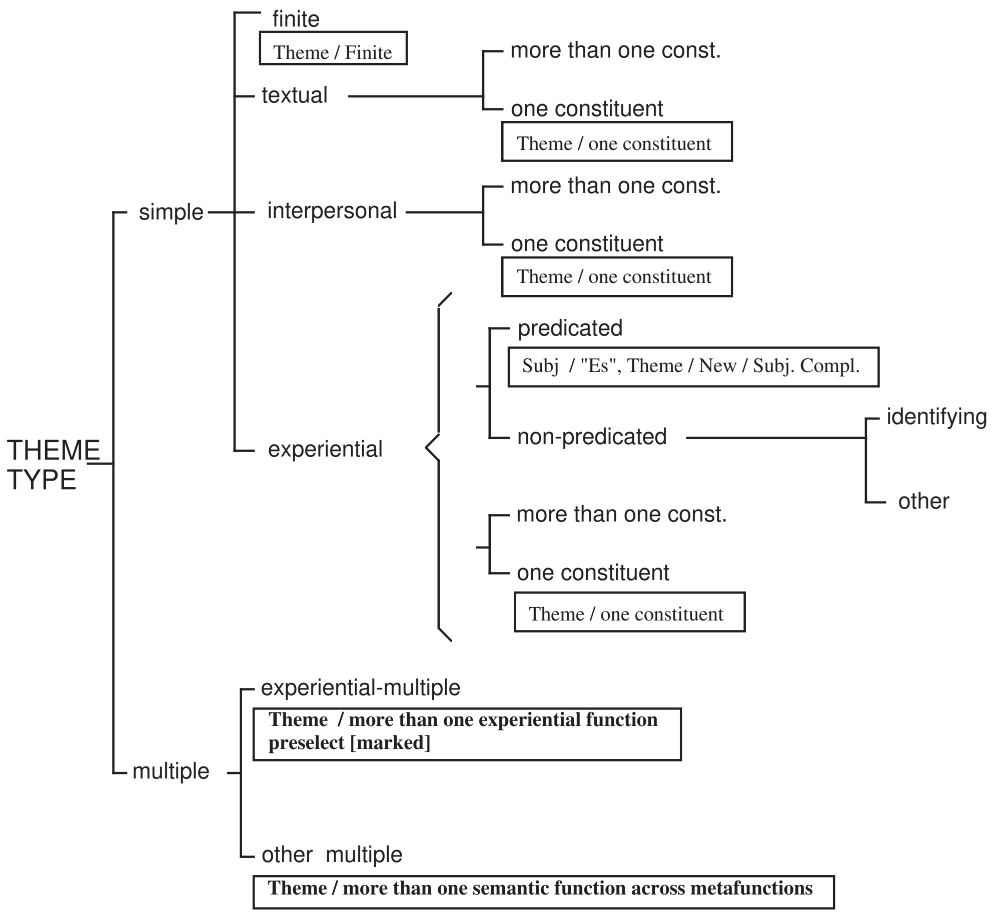

Figure 16: Theme types in German

As we stated above, additional grammatical systems for Theme are given in DANDELION deliverable R2.1.3. and P2.1.3..

6.3.2 The semantics of Theme for German

For each *figure*, to use Halliday and Matthiessen's semantic term for what in the congruent case corresponds to a clause lexicogrammatically (cf. [Halliday and Matthiessen, in press],

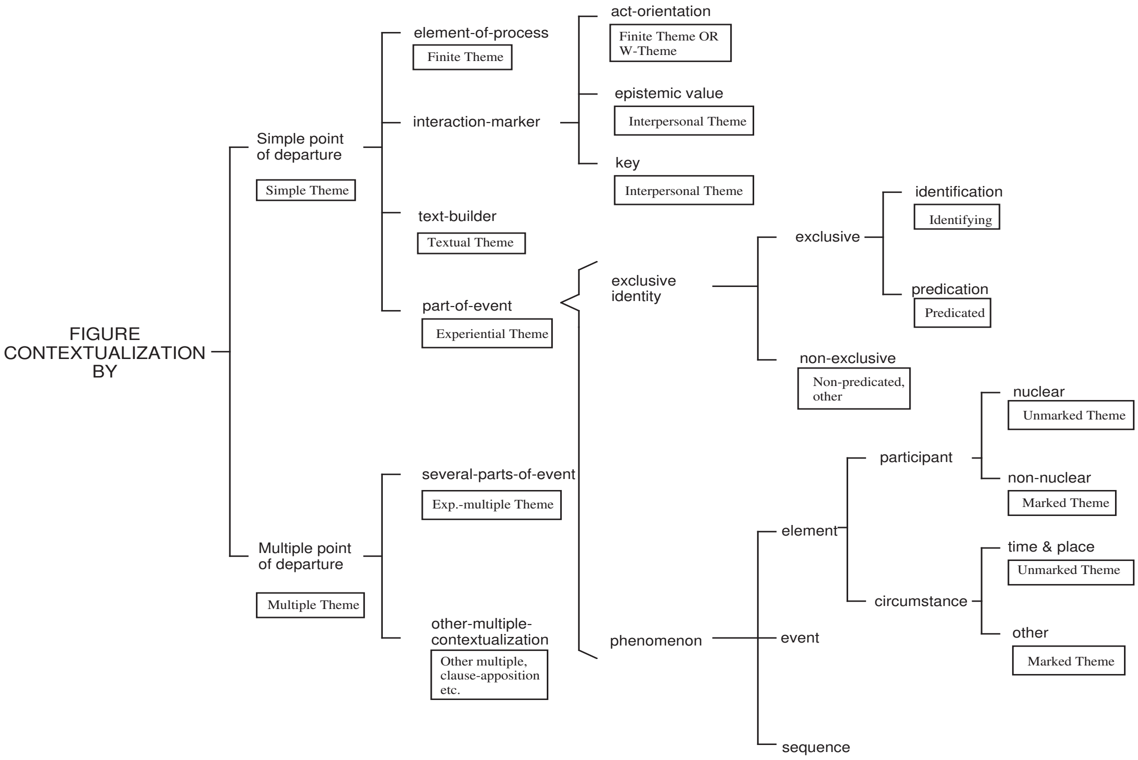

Figure 17: Theme semantics

there is an element of the message which serves as its point of departure, which contextualises it. This point may be *simple*, i.e. embracing only one of *element-of-process, text-builder, interaction-marker, part-of-event*, or it may be *multiple*, i.e. embracing one of *several-parts-of-event, other-multiple-contextualization.* The lexicogrammatical realization in the first case is a *simple Theme*, in the second case a *multiple Theme.* At this first point of classification of semantic types, observe, as in all of the more delicate options in our Theme semantics, that the types referred to by the features are *semantic types*, and that the classes referred to in realizational statements (in boxes under the types) are *grammatical* types. This relationship realizes methodologically the status of the semantics as an *interface* between discourse and lexicogrammar.

Contextualisation by (simple point of departure OR multiple point of departure)

We shall now proceed one step in delicacy further:

simple (element-of-process OR interaction-marker OR text-builder OR part-of-event)

multiple (several-parts-of-event OR other-multiple-contextualization)

Simple points of departure for messages in German seem to offer at least the following four options:

- *element of process*: The lexicogrammatical realization of this is a *finite theme* in terms of the German grammar fragment, e.g.in clauses of the type *Es hat noch nie jemand hier gewohnt, Es wohnt hier schon lange niemand mehr, Hat hier auch schon lange niemand mehr gewohnt, so....* The semantic motivation for this class may be twofold, at least: In so far as the finite element realizes the features *mood, person*, it can be seen as one of the choices under *interaction marker, act orientation* discussed below, i.e. it is exactly the *mood* or *person* meanings that are being used for contextualization at this point. *Person*, by the way, appears to be 3rd obligatorily, but that also has to do with the grammatical resistance of referents with high *identifiability* against positions of relative *newness* in clauses (cf. [Steiner and Ramm, 1995]). Where the finite element is conflated with the lexical verb, it carries semantically an element of process, which is why we are discussing it here. The additional function of the *finite* in German to encode *tense, diathesis* is not further discussed at this point, yet clearly part of the overall picture.

- *interaction-marker.* One lexicogrammatical realization of this is *interpersonal Theme*, e.g. *Leider hat er das getan, Ehrlich gesagt freue ich mich, Vielleicht regnet es morgen.*

- *text builder* The lexicogrammatical realization of this choice would be a *textual theme*, e.g. *Dennoch trat die Vorhersage nicht ein, Und dies tut auch nichts zur Sache, Was diesen Einwand betrifft, (so) ist er nicht stichhaltig.* The fact that a thematic clause has, like any clause, an internal experiential meaning, as a *betreffen* process in our last example, should not distract us from the fact that within its clause complex, it has a *textual* function. The distinction, of course, is blurred in many cases, classically with pronominal adverbs such as *daher, daran, etc.*. Another important observation here is the following: *Text builder* is, like the others introduced here, a *semantic* category which is not in a one-to-one relationship with a grammatical category such as *textual Theme.* A *text builder* is, for example, any clause related through one of the *textual* coherence relations (rather than the more experiential ones) to another, and the question of whether or not it is grammatically encoded as a *textual Theme* is a separate question. There are, for example, means of contextualizing a figure which do so predominantly through a *tactic relationship*, whereas others may do so through constituency.

- *part of event.* Realization: *experiential theme*, e.g. *Die Vorhersage trat nicht ein.*

Multiple points of departure offer the following possibilities:

- *several parts of event*; realization: *experiential-multiple* Theme, e.g. *Die Angelegenheit ueberbetonen wollen wir nicht*, or *source-path-goal* Themes and *telescope* Themes, yet these phenomena follow from the general status of such constituents.

- *other multiple contextualisation*; realization: either *other-multiple Theme* or *apposition, Voranstellung*, e.g. *Dennoch, wir haben keine Wahl*, *Was Eure Chancen angeht, ich sehe schwarz* (Apposition) vs. *Frank, ich sehe schwarz!* (multiple Theme). But both grammatical devices are means of *other multiple contextualization*, i.e. this is an instance of the many-to-many relationships between semantics and lexicogrammar.

Let us move another step further in delicacy, looking at more delicate options under *interaction marker* and under *part of event* in particular. In both areas we shall see that the relationship between semantics and grammar becomes somewhat more indirect than before, and also more interesting from the point of view of methodology and architecture.

$\boxed{\text{interaction marker (act orientation OR epistemic value OR key)}}$

If a message is contextualized by a marker of its interactional value along one of several dimensions, and if the dimension is *act orientation*, then we get a thematization of some mood-indicating element grammatically, e.g. a straight imperative (rather than an *indirect one*), or a straight interrogative (rather than an indirect one), or a W-element in case of a content question. The lexicogrammatical realization of *act orientation* is thus *Finite Theme* or a subtype of *unmarked Theme AND interrogative.* As a correlate, this type of approach claims that if we have selected, in the *interaction base*, say, *demand goods and services* AND we select in our text base *point of departure, simple, part of event, participant*, then we shall select some indirect command which does not have imperative mood grammatically.

If a message is contextualized by the *epistemic value* assigned by the speaker to the message, then the realization is *interpersonal Theme.* The semantic constraints thus enforce the presence of an adjunct expressing an epistemic value, and the conflation of this with Theme position, rather than, say, a modal auxiliary in position of the Finite. However, the latter is entirely a grammatical and thus more language specific question. The text plan at this point will only state that there is an expression of epistemic value and that it should serve as contextualization of the message (cf. [Bateman, 1989], [Steiner, 1992]).

If a message is contextualized by the *key* of that message, i.e. *overtone or undertone etc.*, then the lexicogrammatical realization would be again an *interpersonal Theme*, i.e. in the grammar, and depending on the language, a particle or other lexical element expressing key in Theme position (e.g. Definitiv ist dies der Fall/ *Positively, this is the case*).

We can thus see how the realization of indirect vs. direct commands and questions, as well as certain choices in epistemic value and in emotional attitude (key) are controlled by this semantic system - an important point of entry into questions of congruence and metaphor etc..

Moving on to the possibilities under *contextualization by, simple point of departure, part of event*, we can see at first that the lexicogrammatical realization of the semantic class *part of event* is that the grammatical clause will have an *experiential Theme.* Then, as for messages which have as point of departure some part of event, there seem two conjunctively related sets of distinctions:

$\boxed{\text{(exclusive identity (exclusive (identification OR predication) OR non-exclusive)}}$

AND

$\boxed{\text{(Phenomenon (element OR event OR sequence))}}$

In cases in which an *exclusive identity* is assigned to the point of departure, and as subtype *identification*, the lexicogrammatical realization is an *identifying Theme*, when the subtype is *predication*, the realization is *predicated theme.* Alternative realizations in German would seem to be *marked Themes* of the type exemplified in (1):

(1) Auf den Schrank hast Du die Tabletten gelegt

At least they also assign a type of exclusive identity through their narrow focus and marked intonation. They do not assign that exclusive identity by using the grammatical metaphor of a relational process with the verb *sein*, so that the marked Themes with narrow focus may have to be excluded from this category, yet this is subject to some further inquiry.

If the choice here is *non-exclusive**, then the realization is *non-predicated, other* in terms of the grammatical Theme systems for German.

The other major semantic choice in the area of *part of event*, and one conjunctively related to *exclusive identity*, seems to be as to the Phenomenon ([Halliday and Matthiessen, in press]) to be chosen as point of departure. The major options here are *element, event, or sequence.*

Under *element*, the more specific options are *participant, circumstance.* Observe, by the way, that this reflects the fact that we are talking about part-of-event Themes here only, i.e. not interaction markers or text builders. This is important because it is mirrored in the grammar of German (and of English) by the fact that interpersonal and textual themes usually cannot come into the scope of sentence negation and are not predicated or identified.

Another possible question is why we do not include a semantic subtype here *process*, because (parts of) the process can be thematized to different extends in English and German. The answer is that we cover this subtype in our less delicate type *element of process* above, the reason being that, if some element of the process becomes Theme, it does not fall into any of the classes to which *exclusive identity* can be assigned in German, and is therefore dealt with above, rather than here, where it might rather belong in terms of experiential categorizations.

Under the semantic type *participant*, the options are *nuclear vs. non-nuclear.* The lexicogrammatical realization of making the nuclear participant a Theme is an *unmarked Theme*, the realization of making a non-nuclear participant the Theme is *marked Theme.* Observe that in the grammar, an understanding of this area involves going into the fairly complex interplay between process types and hierarchies of being inherent (Thematic Hierarchies, Participant Role Hierarchies) – yet what matters in the semantics is only a notion of *core vs. non-core.*

If under the category *element*, we have a semantic Theme which is *circumstance* rather than *participant*, there is a further subdistinction into *time and place* vs. *other*, the first of which again goes together a grammatical realization as *unmarked Theme*, the second with *marked Theme.* Now, this would be evidence to suggest that we could model the semantics here in a slightly more general way, suggesting two simultaneous systems of *centrality (core, non-core)* AND *element-type (participant, circumstance)* which would postulate the same semantic types, but in a more general way. This, however, should at the moment not be decided, because the two formulations are equivalent with respect to their extension into the grammar.

If the semantic type of the clause in terms of Theme is *simple point of departure, part of event, phenomenon, event*, then the grammatical realization is *Theme rank, constituent on different rank, higher.* If the semantic type of Theme is a *sequence*, then the grammatical realization is a clause complex.

6.4 A semantic reformulation of category of local theme in Table 2

We shall now repeat the relevant columns of Table 2 , but with the *Category of local theme* reformulated in terms of our new semantic interface. For the semantic types, we shall only specify the features which are necessary for identification, rather than giving the full feature descriptions. Occasionally, we shall add some feature in brackets which is one step in delicacy beyond the terminal features in our current Theme semantics. These are completely predictable, though, from the small sets of participant roles and circumstantial roles.

	Local theme	category of local theme	semantic type of local contextualization
(01)	Ferdinand VII., König (1808 unnd seit 1814)	sim, exp	participant, nuclear
(02)	Sohn Karls IV. und Maria Luisa von Parma	sim, exp	participant, nuclear
(03)	1807	sim, exp	circumstance, time and place
(04)	Diese Kronprinzenverschwörung...	sim, exp	participant, nuclear, (cause)
(05)	Ihn	sim, exp	participant, non-nuclear, (range)
(06)	Erst im März 1814	sim, exp	circumstance, time and place
(07)	Die Liberalen	sim, exp	participant, non-nuclear, (affected)
(08)	Die liberale Revolution	sim, exp	participant, nuclear, (affected)
(09)	Nach dem Tode...	sim, exp	circumstance, time and place
(10)	Um dieser ...	sim, exp	circumstance, other
(11)	Die Thronansprüche ...	sim, exp	participant, nuclear, (cause)
(12)	Wegen ...	sim, exp	circumstance, other, (reason)

This is our current account of the semantics of Theme in German from an inside-out perspective. In the semantic classes that we have proposed here, we hope to provide the methodological link to work on *discourse patterns* and *discourse parameters* as discussed above. The implications of *text types*, for example, for text organization in terms of *Themes* should now be formulated in terms of the *semantic classes* proposed here, rather than grammatical classes directly. The column **Category of chaining strategy** in Table 1 contains the semantic classes which are subtypes of *participants, circumstances, text-builders* in terms of Figure 17 above. The discourse researcher should thus be liberated from having to deal with unnecessary lexicogrammatical detail in their work, being able to concentrate on semantic correlates of their textual i.e. functional categories. In terms of our general architecture (cf. Figure 1), we have been proposing a module of the textual local semantics, working *inside - out* from the lexicogrammar.

7 Future Work

In this closing section we will make a few remarks concerning the way in which we intend to integrate and combine the complementary approaches to aspects of discourse organization presented in the last sections.

A relatively straightforward step is to employ the text-type specific correlations between global and local discourse characteristics of a text such as those outlined in Section 5 for text planning and text generation: the recognition of information types at the higher levels of abstraction such as text type, discourse purpose and subject matter can be modelled as constraints on the selection lower-level features responsible for the lexicogrammatical organization on sentence level. In this way they can be seen as a counterpart of the fragment of the local discourse semantics specified in Section 6.3.2 in that the global constraints on discourse organization provide a basis for controlling choices between the various semantic options for theme choice.

As one key method for pushing us closer to an integration of the components developed so far into one workbench we intend the specification and, to the extent that this is possible, implementation of one complete cycle from a text representation of a sufficiently small fragment of text through semantic constraints to lexicogrammatical realization. Subject to final agreement on this between DANDELION partners, that specification/implementation would include:

- a text plan/discourse segment for the passage
- the global, discourse semantics of Theme
- the local, sentence semantics of Theme
- the lexicogrammar of Theme
- a component covering the dynamic aspects of the fragment
- inquiries, or their re-implemented equivalents
- output

Some space should also be given to a detailed consideration of the question where and how in our present architecture we deal with questions of *multilinguality*: The challenge of multilinguality is not one of the central ones of the project, yet it is one that forms part of the background to a substantial amount of DANDELION'S work and it is certainly hoped that the project will be able to contribute to a further understanding of the issue in meaningful ways.

References

[Bateman and Paris, 1989] John A. Bateman and Cecile L. Paris. Phrasing a Text in Terms a User can Understand. In *Proceedings of the Eleventh International Joint Conference on Artificial Intelligence*, Detroit, Michigan, USA, 1989.

[Bateman and Rondhuis, 1994] John Bateman and Klaas Jan Rondhuis. Coherence Relations: Analysis. Deliverable R1.1.2a (Draft) of WP 1.1.2 'Coherence Relations', ESPRIT Basic Research Project 6665 DANDELION, GMD-IPSI Darmstadt, Darmstadt, July 1994.

[Bateman *et al.*, 1994] John Bateman, Jo Calder, Renate Henschel, and Erich Steiner. Specification of a Discourse Grammar Development Tool. Deliverable R2.2.1 (Preliminary Version) of WP 2 'Grammar Integration', ESPRIT Basic Research Project 6665 DANDELION, GMD-IPSI Darmstadt, Darmstadt, October 1994.

[Bateman, 1989] John A. Bateman. Upper Modelling for Machine Translation: A Level of Abstraction for Preserving Meaning. Technical report, IAI-Eurotra-D, Saarbrücken, FRG, 1989.

[Bateman, 1993] John A. Bateman. The Dandelion Project – an interim report. Technical report, GMD/IPSI, 1993.

[Brinker, 1992] Klaus Brinker. *Linguistische Textanalyse: Eine Einführung in Grundbegriffe und Methoden.* Schmidt Verlag, Berlin, 1992.

[Doherty, 1991]
Monika Doherty. Informationelle Holzwege. Ein Problem der Übersetzungswissenschaft. *Zeitschrift für Literaturwissenschaft und Linguistik*, 21(84), 1991.

[Engel, 1988] Ulrich Engel. *Deutsche Grammatik.* Groos, Heidelberg, 2nd edition, 1988.

[Faigley and Meyer, 1983] Lester Faigley and Paul Meyer. Rhetorical theory and readers' classification of text types. *Text*, 3:305–325, 1983.

[Halliday and Matthiessen, in press] Michael .A.K. Halliday and Christian M.I.M. Matthiessen. *Construing experience as meaning.* Foundations of Communication and Cognition1. Mouton De Gruyter, Berlin etc., in press.

[Halliday *et al.*, 1964] Michael A.K. Halliday, A. McIntosh, and Peter Strevens. *The linguistic sciences and language teaching.* Longman, London, 1964.

[Hansen, 1989] G. Hansen. *Textlinguistische Analyse von Gebrauchstexten.* Nyt Nordisk Forlag Arnold Busck, Kopenhagen, 1989.

[Heinemann and Viehweger, 1991] Wolfgang Heinemann and Dieter Viehweger. *Textlinguistik. Eine Einführung.* Niemeyer, Tübingen, 1991.

[Herbig, 1992] Albert F. Herbig. *Sie argumentieren doch scheinheilig. Sprach- und sprechwissenschaftliche Aspekte einer Stilistik des Argumentierens.* Peter Lang, Frankfurt am Main, 1992.

[Hoberg, 1981] Ursula Hoberg. *Die Wortstellung in der geschriebenen deutschen Gegenwartssprache.* Hueber, München, 1981.

[Kuppevelt, 1995] Jan van Kuppevelt. Discourse structure, topicality and questioning. *Linguistics*, 31:109–147, 1995.

[Lavid, 1994a] Julia Lavid. Thematic Development in Texts. Deliverable R1.2.1 of WP 1.2.1 'Thematic Structure', ESPRIT Basic Research Project 6665 DANDELION, Universidad Complutense de Madrid, Madrid, October 1994.

[Lavid, 1994b] Julia Lavid. Theme, Discourse Topic, and Information Structuring. Deliverable R1.2.2b of WP 1.2.2 'Theme, discourse topic, and information structuring', ESPRIT Basic Research Project 6665 DANDELION, Universidad Complutense de Madrid, Madrid, October 1994.

[Linke *et al.*, 1991] Angelika Linke, Markus Nussbaumer, and Paul R. Portman. *Studienbuch Linguistik.* Niemeyer, Tübingen, 1991.

[Martin, 1992] J.R. Martin. *English Text. System and Structure.* John Benjamin's, Amsterdam, 1st edition, 1992.

[McCoy and Cheng, 1990] K. McCoy and J. Cheng. Focus of attention: constraining what can be said next. In C. Paris et al., editor, *Natural language generation in artificial intelligence and computational linguistics*, pages 103–124. Kluwer, Dordrecht, 1990.

[Moehn, 1991] Dieter Moehn. Instruktionstexte. ein Problemfall bei der Textidentifikation. *Germanistische Linguistik*, 106-107:183–212, 1991.

[Neuendorff, 1988] Dagmar Neuendorff. Textsorte als Handlung - zu einigen Aspekten einer prozeduralen Textsortenbeschreibung. *Neuphilologische Mitteilungen*, LXXXIX:539–546, 1988.

[Ramm and Teich, 1994] Wiebke Ramm and Elke Teich. German Grammar Fragment. Implementation of Theme for German in the KOMET-PENMAN Text Generation System. Deliverable P2.1.3 of WP 2.1.3 'Grammatical Coverage German', ESPRIT Basic Research Project 6665 DANDELION, University of the Saarland, Saarbrücken, July 1994.

[Rothkegel and Villiger, 1993] Annely Rothkegel and Claudia Villiger. Chaining. Deliverable R1.2.2a of WP 1.2.2 'Information Structuring', ESPRIT Basic Research Project 6665 DANDELION, University of the Saarland, Saarbrücken, October 1993.

[Rothkegel *et al.*, 1994] A. Rothkegel, M. Schmitt, and C. Villiger. CHAINING - eine Strategie zum Aufbau der sequentiellen Textstruktur. *Sprache und Datenverarbeitung*, 1994.

[Steiner and Ramm, 1993] Erich Steiner and Wiebke Ramm. Grammatical Coverage German. Deliverable R2.1.3 of WP 2.1.3 'Grammatical Coverage German', ESPRIT Basic Research Project 6665 DANDELION, University of the Saarland, Saarbrücken, October 1993.

[Steiner and Ramm, 1995] Erich Steiner and Wiebke Ramm. On Theme as a grammatical notion for German. *Functions of Language*, 2.1, 1995.

[Steiner, 1991] Erich Steiner. *A Functional Perspective on Language, Action, and Interpretation.* Mouton de Gruyter, Berlin and New York, 1991.

[Steiner, 1992] Erich Steiner. Some Remarks on a Functional Level for Machine Translation. *Language Sciences*, 14(4), 1992.

[Toulmin, 1957] Stephen Toulmin. *The uses of argument.* Cambridge University Press, Cambridge, 1957.

[Villiger, 1995] Claudia Villiger. Theme, discourse topic, and information structuring in German texts. Deliverable R1.2.2c (Draft) of WP 1.2.2 'Theme, discourse topic, and information structuring', ESPRIT Basic Research Project 6665 DANDELION, University of the Saarland, Saarbrücken, in preparation 1995.

[Virtanen, 1992] Tuija Virtanen. *Discourse Functions of Adverbial Placement in English. Clause-Initial Adverbials of Time and Place in Narratives and Procedural Place Descriptions.* Abo Akademis forlag, Abo, 1992.

[Weinrich, 1993] Harald Weinrich. *Textgrammatik der deutschen Sprache.* Dudenverlag, Mannheim, 1993.

[Werlich, 1979] E. Werlich. *Typologie der Texte: Entwurf eines textlinguistischen Modells zur Grundlegung einer Textgrammatik.* Quelle und Meyer, Heidelberg, 1979.