DANDELION

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

ESPRIT Basic Research Project 6665

Chaining

Annely Rothkegel, Claudia Villiger

September 1993

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

Contents

Document ID	:	Deliverable R1.2.2.a: *Chaining*
Issue date	:	September 1993
Responsible	:	Universität Saarbrücken
Version	:	Final
Status	:	Public
WP number	:	WP 1: Empirical studies
Distribution	:	DANDELION consortium, CEC
Keywords	:	Chaining, global text structure

1 Introduction

CHAINING refers to discourse phenomena in the area of text connectivity. The basic idea is that there are two perspectives which are reflected in chains of connectors between the succeeding parts of a text. One perspective concerns specific types of organizing information according to text-external principles (e.g. domain knowledge, text topic, communicative function), the other perspective considers text organization according to text-internal principles of linearity. Text-external principles lead to hierarchies which have to be transformed into sequences on the basis of text-internal principles (cf. distinction in [van Dijk, 1980]; cf. also the 'two-track'-theory in [Oversteegen, 1989] concerning the temporal structure in texts).

As we consider problems of CHAINING with respect to this relationship between text-external and text-internal principles we start with the discussion of the text-theoretical background. It determines the status of the specifications which contribute to the envisaged text grammar. We distinguish between two kinds of text grammars. One kind supplements sentence grammars by introducing contextual factors into the semantic/pragmatic interpretation of syntactic units of a sentence. The other type of text grammar concerns the relationships between text parts which are determined by local and/or global coherence. CHAINING represents a specific type of connecting text parts on the basis of global coherence. The corresponding specifications refer to the semantic level as well as to the pragmatic level of texts. In our model both of them are related to schemata which represent a coherent organization of information pieces. The role of CHAINING is to transform them into sequences according to particular sequential orders. The task of the text grammar is to specify the constraints for connecting text parts. Furthermore, it determines the consequences of those constraints for the description of sentences.

Our report tries to develop this idea in two ways. Section 2, Theoretical foundation, establishes a general framework in which the envisaged specifications can be embedded. In addition, this framework should provide some possibilities of integrating other aspects such as those elaborated in related DANDELION tasks (e.g. anaphora, the treatment of temporal connectives, thematic development, and finally an interface to sentence related grammars).

Section 3 reports a text-linguistic analysis of an authentic newspaper text. This study shows the linguistic effects of the principles introduced in section 2. It will be demonstrated how unmarked and marked chainings can occur at different hierarchical levels of text.

2 Theoretical foundation

2.1 Questions

It is assumed that coherence in texts goes beyond the relationship between two units (sentences or groups of sentences). The constitution of coherence is considered to be the constitution of a new view of information by means of a specified text structure. The text structure covers the whole text as a linguistic unit. This view of coherence is different from approaches such as [Hobbs, 1979], [Dahlgren, 1989], [Noel, 1990], [Norgard-Sorensen, 1992]. These models refer to single "coherence relations" between pairs of text parts. They may be explicit or implicit and concern relationships of a deeper text structure (e.g. knowledge about standards of sequences of events, logical connections between meanings, coreference) which may be composed to hierarchical structures also covering the whole text (e.g. rhetorical relations in [Thompson and Mann, 1987]). Our approach is complementary to these in the sense that it focuses on global aspects.

From this point of view, coherence means that connections between the parts of a text exist in such a way that they form together a text specific system of interrelationships by which information about objects, events or states of affairs is constructed and presented with regard to communicative purposes. This paper focuses on the following groups of questions:

(i) Semantic view of text structure: This field of questions concerns the basic units of discourse representation. There are two requirements. It should be possible to relate the semantic representation of the units to world knowledge and to the sentence grammar.

World knowledge comes in when we are dealing with the discourse topic. It is assumed that there are standards of topic organization which also determine the selections of pieces of information as well as the selections of the corresponding linguistic units. These are strongly related to units at sentence level. Therefore we argue for a close relationship between representations at discourse level and at sentence level. From this point of view, propositions in terms of predicate-argument-configurations seem to be a very proper representation format. Predicates as well as arguments may be interpreted in different ways according to the chosen interpretation model whether it refers to a text topic model or to a functional sentence grammar. Questions arise with respect to an adequate text-topic model. We will discuss them on the basis of schemata as information organizing instruments.

(ii) Pragmatic view of text structure: Whereas the semantic view provides some criteria for the determination of the basic units of discourse representation the pragmatic view aims at contextualization and its consequences for verbalization. This means that a certain propositional content may be linguistically realized in different ways according to distinct communicative purposes. At sentence level a consequence is lexical selection (lexicalization) and word order. With respect to discourse structure we have to deal with different orderings and connections of text parts. Similar to the semantic approach the pragmatic aspects will be formulated in terms of schemata which organize communicative purposes.

(iii) Global perspective: How can text structure be determined? We assume that there are relationships between text parts that can cover the whole text. The global perspective distinguishes content aspects and formal aspects. Content considerations are needed where text parts are determined by semantic or pragmatic schemata. Formal aspects arise from the hierarchical order of the text parts, which codetermines their interrelationship. In discourse, this order is realized in terms of sequencing. The global perspective therefore addresses questions concerning the transformation of hierarchies into sequences.

(iv) Dynamic perspective: The forming of the text structure is characterized by directionality which is a feature of the sequential order of texts. Sequences are constructed by a continuous attaching of text parts. It is assumed that this movement is reflected in terms of connections between the text parts. CHAINING is a kind of attaching parts in order to complete a sequence which is based on topic schemata and/or communication schemata.

In this section these four aspects are dealt with in more detail and are related to other approaches. In addition, they are discussed in relation to the envisaged text grammar (representation) and in relation to a sentence grammar.

2.2 Semantic view of text structure (STS)

The semantic level is considered insofar as it is important for the embedding of pragmatic aspects. Therefore we introduce an abstract carrier-structure which allows the attachment of all text-relevant specifications. Such a structure is the representation of a text list in terms of a list of propositions (cf. [van Dijk, 1980], macrostructures). This text list is based on the following ideas.

There are two properties of text structure which are fundamental: recurrency and connectivity (cf. [de Beaugrande and Dressler, 1981]). Recurrency provides some stability of the content structure by means of repetition of the same units (lexical or semantic units). By contrast, connectivity provides some possibilities of expansion and progression of content structure. This is achieved by introducing different and/or new lexical and/or semantic units. Both principles are involved in the process of building a text structure. They also have an impact on text-thematic development. This development differs from the theme-rheme development (cf. "thematic progression" in [Daneš, 1974], or the "referential movement" in [Klein and von Stutterheim, 1991]). It is assumed that there is a text-theme which provides the possibility of combining parts in a particular way and which refers to the overall motivation for producing the text. A text is not produced in order to express reference to a specific object (or event, etc.) but merely to construct new connections by combining the referents with further information. Therefore it seems not to be helpful to consider a text theme only in terms of reference and/or coreference. Instead it should include a more complex object which is represented in terms of a thematic question (according to the quaestio-approach by [Klein and von Stutterheim, 1991], [Von Stutterheim, 1992]). The idea is that a text consists of the answers which relate to an abstract thematic question. This delimits the domain of the world knowledge which is relevant for this particular text. A coherent sequence of sentences is necessary in order to answer it. Normally the basic form is a wh-question which meets these requirements:

- What happened at a specified time with regard to x?
- Why did x react in such and such manner?
- What are the reasons for the positive/negative value of a specified object?

The question is where do thematic questions come from? First of all, it is assumed that they are not totally arbitrary. They play a central role in communication and with regard to this they correspond to standards of information transfer. They are text-external and correspond to knowledge frames as structures of experiences and expectations in communication (cf. [van Dijk, 1980], [Metzing, 1980], [Minsky, 1975], [Minsky, 1990], [Tonfoni, 1990]). With respect to the text structure these thematic questions introduce classes of slots for predicates and argument-roles which are filled in the text. What are the consequences of such an approach for the representation of the semantic text structure?

We think that it will be necessary to establish some kind of propositional basis which allows the structuring of the content in terms of a predicate-argument-configuration which corresponds to the structure of the thematic question. On the one hand this propositional basis should be very general and on the other hand it should offer the possibility to express very specific information. In order to fulfill these requirements we introduce the concept of 'text role' (TROL). It is assumed that the wh-question provides different slots concerning a specific domain which are to be filled through the various constituents of the sentences. From this point of view, text roles have the function of an interface between the text-external thematic structure and the text-internal sentence structure of the single text parts. A TROL corresponds to a slot which is determined by the configuration of the thematic question. In formal terms a TROL represents an argument within the propositional configuration as an attribute-value pair. This allows specifications at different levels corresponding to the particular theoretical approaches, e.g. categories of the domain, of semantic roles (deep cases), of conceptual categories or of lexical units (in our example in section 3 we will refer to lexical units).

The form of the basic propositional configuration (p) is the following:

[1] p = (PRED
    (TROL1 = value (X = value (...)))
    (TROL2 = value (X = value (...)))
    ...
    (TROLn = value (X = value (...))))

P represents the basic unit of the semantic text structure (STS). This text structure is then composed of a list of Ps based on the assumption that the real text structure is a sequential one (see section 2.5).

The form corresponds to:

```
[2] STS = (list
        (p1 (PRED ..... ) )
        (p2 (PRED ..... ) )
        ...
        (pn (PRED ..... ) ) )
```

Concerning the text-thematic connections between the various configurations of a text we have to rely on the two principles of text constitution mentioned above (recurrency and connectivity). In order to construct a stable text structure text roles are successively repeated. If this is achieved by using the same referents (or partially the same referents) this can be described in terms of anaphorical relations (which can also be represented in this model).

Informational and text-thematic progression results from the construction of specific connections between the same or different units. These units concern single text roles as well as whole propositions. In order to grasp these connections we supplement our TROL-configurations with a CON-component which represents the different types of connectivity.

The semantic text structure can be represented as a list of propositions which are linked by a CON-component. STS then looks as follows:

```
[3] STS: (list
       [CON = value ( ... ) ]
       (P1 (PRED
                 (TROL1 = value ( ... ), [CON = value (...) ] )
                 ... ) )
       [CON = value ( ... ) ]
       (P2 (PRED
                 (TROL1 = value ( ... ), [CON = value (...) ] )
                 ... ) )
                 ...
       [CON = value ( ... ) ]
       (Pn (PRED
                 (TROL1 = value ( ... ), [CON = value (...) ] )
                 ... ) ) )
```

The specifications of CON are described in section 2.5.

The concept of the text question has consequences for descriptions of theme-rheme-relationship. Without discussing this question in detail here, it should be mentioned that a relatively simple integration is possible. Slots which are introduced by the position of the wh-part belong to rhematic units, whereas the central text roles are representatives of the thematic parts.

2.3 Pragmatic view of text structure (PTS)

The text-thematic structure is normally embedded in a communication-oriented environment. This has effects on the specifications of the values of the text roles. It also influences the form of the thematic text question and consequently the selection of text roles as well as the whole organization of the propositional structure. This aspect touches one of the central topics of the project as it concerns the phenomenon of the same propositional content in different linguistic realizations.

It is assumed that different communicative purposes lead to different effects. The question arises how these relationships can be grasped in terms of a linguistic representation (cf. [Rothkegel, 1993].

We assume that there are standards of communicative purposes which are reflected in schemata of verbal presentation. For instance, an object/event/state of affairs can be a subject of description, report, narration, argumentation, etc. Correspondingly, there are different styles of presentation which are - more or less - standardized in terms of communication schemata (our example in section 3 refers to an argumentative schema).

Communication schemata are standardized organizational structures of how to embed information into the actions and interactions of the language users. Information pieces are grouped according to particular functions in communication. E.g., argumentative functions are characterized by data arguments and a conclusion which is derived from a rule and optionally from a backing part (cf. [Toulmin, 1959]) which in texts are mostly not verbalized. This argumentation schema consists of three types of information pieces and their relationships. With respect to PRO- and CONTRA- argumentation it has to be expanded. Descriptive or narrative schemata can be described in a similar way (cf. [Heinemann and Viehweger, 1991]).

The idea of schemata brings in text-external organization patterns which have a strong impact on text organization.

2.4 Global perspective

Schemata have to do with complex organizational structures. Whereas the semantic and pragmatic view concerns text-external factors, the text form itself provides constraints of organization. On the other hand, there are strong relationships between those text-external factors and text-internal formal constraints which we describe in terms of text schemata.

Schemata of connected pieces of information are well known in cognitive AI-approaches (e.g. frame theory in [Minsky, 1990], [Minsky, 1975], [Metzing, 1980],[Tonfoni, 1990]) as well as in text linguistics [van Dijk, 1980], [de Beaugrande and Dressler, 1981], [de Beaugrande, 1984]). In this paper the text linguistic aspects are emphasized as they focus on the text specific organization of information which may be distinguished from organization types of world knowledge. This distinction is motivated by the assumption that texts are communicative entities which provide information according to contextual purposes.

Thematic questions as well as presentation schemata refer to larger units such as the text as a whole or text portions which can be delimited. The idea of delimitation is based on the assumption of a structured grouping of portions in a text. This means that specified positions in the linear order of the text (see list of propositions) can be functionally interpreted. Similar to word order at sentence level, functions at text level may be determined according to positions (or they have to be marked lexically).

In our view text schemata are interesting for two reasons:

(a) They cover a text (or a delimited text part) as a whole unit which is representative of a hierarchical structure.

(b) They provide a sequential order of the running text according to weighted positions which are relevant for the organization of the various portions of information.

Both properties have effects on the connectivity in texts. As there are different kinds of connectivity relationships to be distinguished in dependence of global text schemata we will look at them in more detail.

A very standard order of text positions is that of three parts: beginning, middle, and end. This pattern is relatively simple, but it has strong effects on phenomena of text organization (cf. sentence order according to Vorfeld, Mittelfeld, Nachfeld). The studies of [Brandt and Rosengren, 1991] show clearly that there are strict constraints on how portions of the hierarchical structure are distributed in the sequential order. For instance, neighbours of the same hierarchical level cannot be separated in the sequential order.

Text schemata may refer to the semantic level of texts as well as to the pragmatic level. Semantic text schemata have been developed in terms of macrostructures ([van Dijk, 1980]) or in dynamic terms such as "thematic development" ([Brinker, 1992]) or "thematic progression" ([Daneš, 1974]). They have in common that the particular relationship between single nominal parts of subsequent sentences (or propositions) are considered. This may concern anaphorical relations based on recurrency or general semantic relationships such as 'whole-part', 'super-/sub-ordination', 'global-detailed', etc. To this group also belong some of the ideational types of 'rhetorical relations' such as 'part-of', 'specification-of', 'sequence-of'. These types of relationship can also be used for organizing the text theme. This is the case if the text positions BEGINNING(1), MIDDLE(2), END(3) are specified by particular instances of these relations.

A very standard form for ordering text thematic pieces is that of 'WHOLE-PART' in a relatively fixed manner as shown in the following structure:

(1) WHOLE

(2) PART1, PART2, ..., PARTn

(3) WHOLE

Position (1) and (3) can be alternatively realized, but a mixture of WHOLE instances and PART instances is not allowed as presented in the following figure.

(1) -

(2) PART1, PART2, ..., PARTn

(3) WHOLE

or

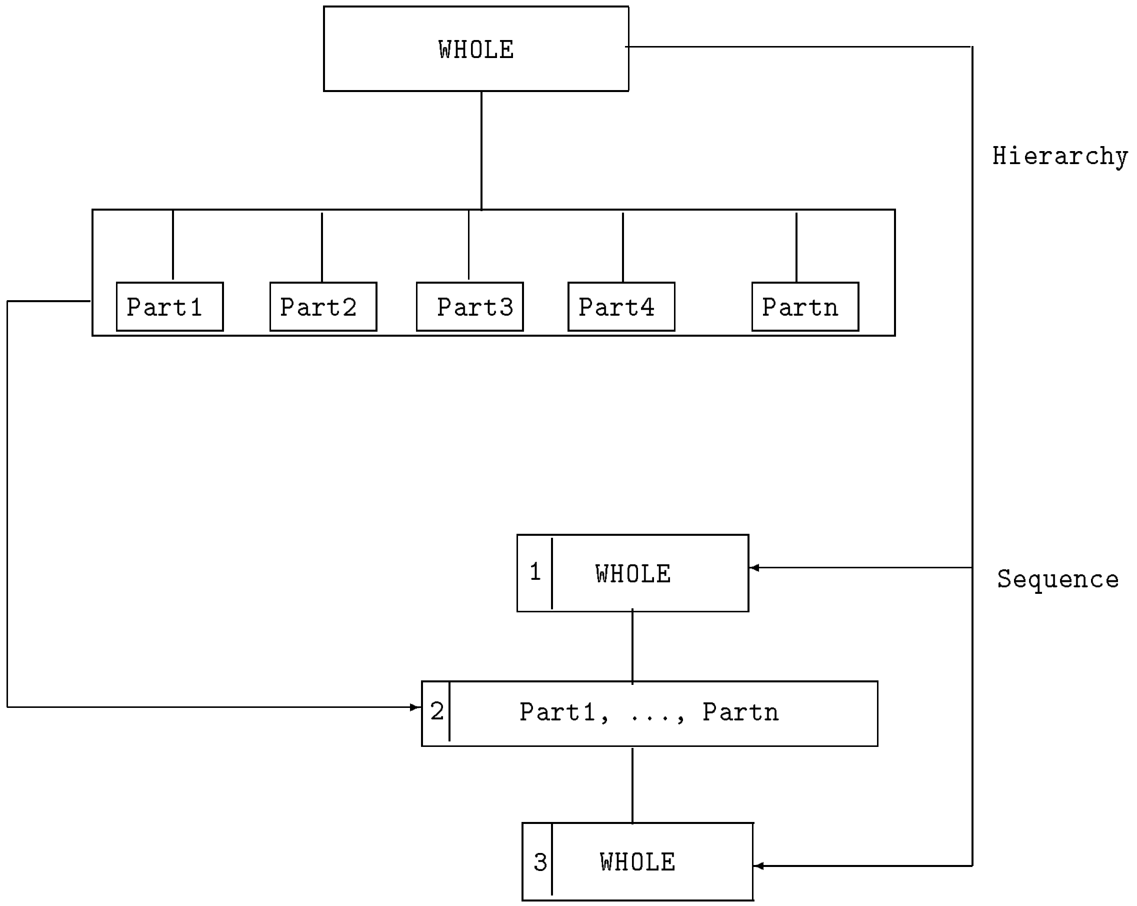
(1) WHOLE

(2) PART1, PART2, ..., PARTn

(3) -

The relation of hierarchical and sequential structure is shown in figure 1.

Figure 1: Global schema



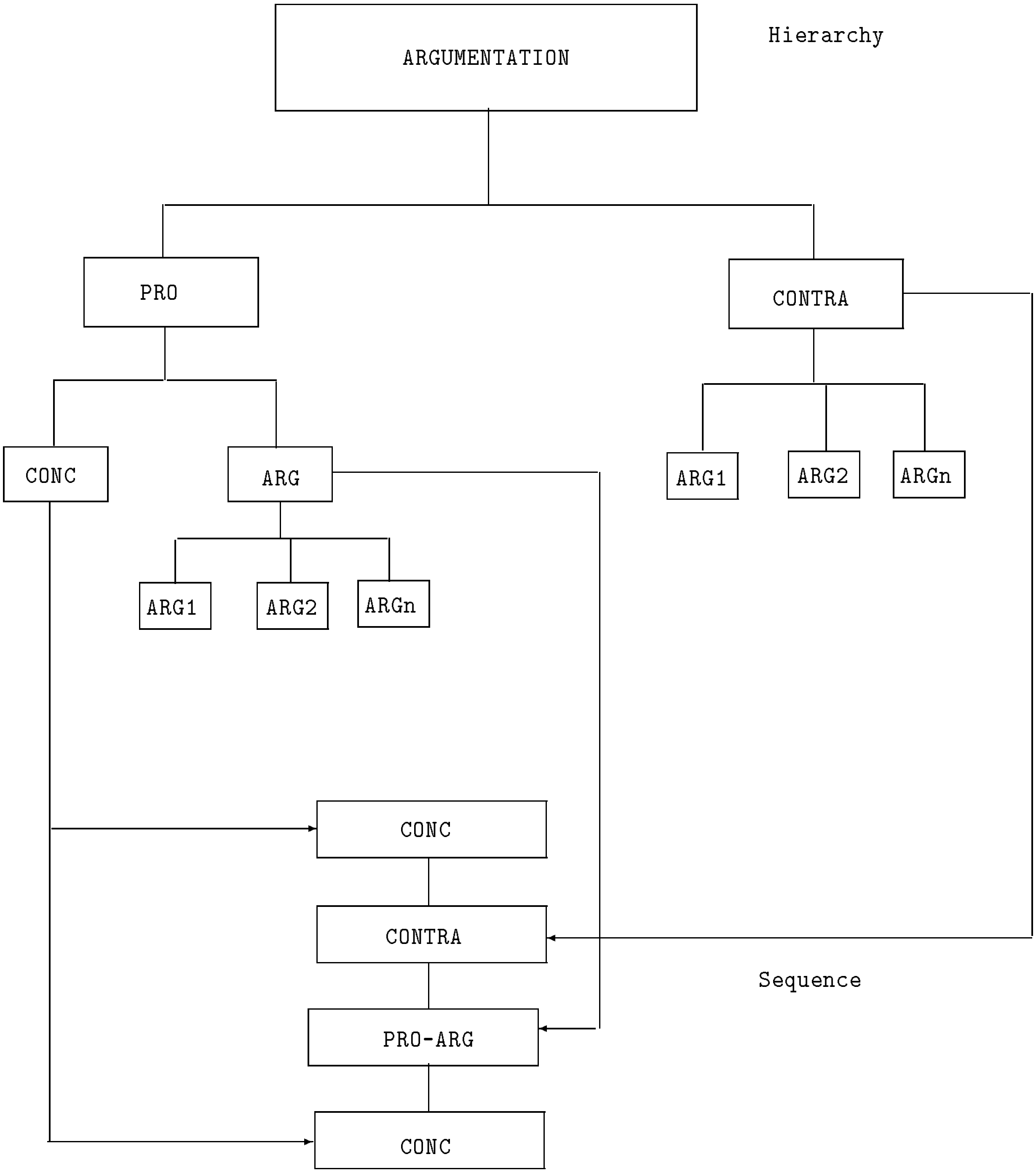
The determination of text positions has an impact on the text list. It is assumed that the placement of the propositions is controlled by global schemata and this in turn has effects on the realization of the principle of connectivity and its linguistic expressions. But it is not sufficient to rely on semantic schemata in natural texts. There is another higher level of control by schemata at the pragmatic level. One type of those schemata is the argumentative schema which also reflects this three-part-structure and which will be discussed in more detail.

Van Dijk [van Dijk, 1980] describes global pragmatic text structures in terms of 'super-structures'. His example of an argumentative text structure shows a hierarchy of arguments (based on the Toulmin schema) with the conclusion missing. This is typical of advertising texts: the increased cognitive effort and attention when the language user has to infer the conclusion makes the advertisement more effective. This strategy is not appropriate for newspaper texts, where the writer is concerned mainly with giving information and possible conveying a particular attitude. The readers here should be addressed in such a manner that their possibly contrary opinions are taken into account. In this case argumentation is a presentation modus for favouring a specific perspective of the text writer. This leads to a dialogically oriented argumentation which contrasts a PRO-aspect and a CONTRA-aspect. This view of argumentation is an interactional one which is complementary to a semantic (monological) view of following the inferential line of arguments.

PRO- and CONTRA- parts form a hierarchy which is reflected in the sequential order as represented in figure 2.

The global schemata form a basic structure which has effects on the organization of thematic and functional development. The verbal realization according to the canonical form can be considered to be unmarked. Marked realizations are bound to specifications which are (additionally) reflected at the lexico-grammatical level of the text (see also [Schaeffner *et al.*, 1987], [Schaeffner and Wiesemann, 1988]).

Figure 2: Argumentative schema



2.5 The dynamic perspective

Connectivity can be seen in conjunction with directionality. This concerns the successive forming of the text structure whereby the process of succession is reflected in markers of connectivity. We will now discuss some principles of directionality (see also [de Beaugrande, 1984], [Rothkegel, 1993]).

Directionality in linear ordered structures such as text structures refers to two directions: left-right and right-left. 'Left-right' directionality is forward-oriented and aims at the completion of the global structure. The corresponding markers are of such a kind that they fulfill the chosen global schema (e.g. argumentation schema). 'Right-left' directionality is backward-oriented. It aims at the attachment of a unit to the preceding unit. The relationship between both units is a local one. The corresponding markers refer to single relations of different sorts. They may become an integral part of the global structure, but this is not necessarily so. It is an open question whether there are standards of relationships between global and local connections.

Another question concerns the linguistic units which can be connected. Candidates for global and local connectivity (CON) are whole propositions as well as text roles (cf. semantic text structure in section 2.2). The specifications of the CON-component are represented as follows:

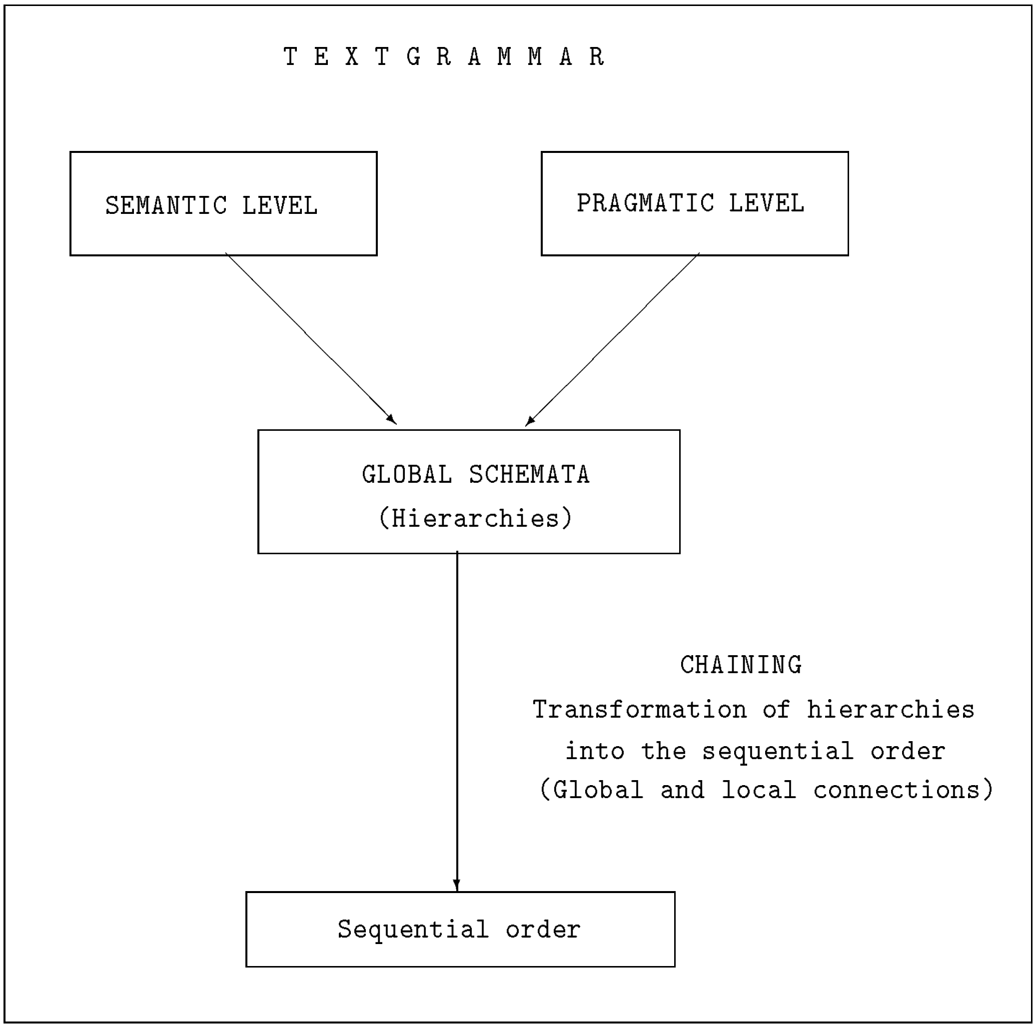
```
[4] (CON =
        (GLOBAL =
                (SCHEMA = [...])
                (RELATION = [...])
                (CONNEC = { list (connec)}))
        (LOCAL =
                (RELATION = [...] )
                (CONNEC = { list (connec)})))
```

Global connectivity is specified by the type of schema and candidates of connectors (CON) of this schema. Local connectivity is specified by the type of relation and the candidates of connectors of this relation.

Correspondingly, CHAINING is expressed in terms of the CON-component which reflects the type of relationship as well as a potential chain of connectors ({list (connec)}). It is assumed that there are distinct patterns of connecting propositions according to thematic development, interactional schemata and global text structures. The filling of those patterns in terms of natural language expressions is a matter of future empirical work. The study in section 3 gives an example of how to proceed in order to elaborate a relatively completed list of patterns for German.

In figure 3 the interplay of the different levels of text and the function of CHAINING is recapitulated:

Figure 3: Text grammar



2.6 Relationship between text grammar and sentence grammar

In the following paragraph some reflections are made about how to link the text grammar and the sentence grammar. These ideas are very provisional and refer to cases such as the example analysed below. The general concept is that an interface between text- and sentence grammar should provide the connection of those specifications of both levels of description which are related. From the perspective of text grammar this concerns all the specifications which have lexico-grammatical effects. We can distinguish two kinds of dependencies:

(a) Text grammar internal effects include the ordering of whole propositions and selection and order of the corresponding connectors introduced by the global schema. It may also concern connectives of local relationships when they contribute to the global structure.

(b) Text grammar external effects result from syntactic properties of the connectives themselves.

In our example (section 3) the parallel structure of the arguments leads to the repeated introduction of the connectives as topics in the syntactic structure of the sentences. This is a fact which is controlled by text grammar internal effects. But these effects have text grammar external consequences. The topic position (Vorfeld) - in its turn - has effects on word order of the sentence concerned.

In a text grammar only text grammar internal effects should be represented in the CON-component. This is achieved by adding a specification type of TEGIE (text grammar internal effects) which may be sub-specified by sentence related constraints such as TOPIC, FOREGROUND, BACKGROUND, etc. :

```
[5] ( CON =
        (GLOBAL =
                (SCHEMA = [...])
                (RELATION = [...])
                (CONNEC = {list (connec)})
                (TEGIE = [...]))
        (LOCAL =
                (RELATION = [...])
                (CONNEC = { list (connec)})
                (TEGIE = [...])))
```

3 Empirical work (text analysis)

3.1 Method

In this section we present an analysis of a newspaper text according to the principles sketched above. After the presentation of the text itself and a rough translation, the first step of our analysis is to decompose the text into a propositional representation of the sentences (semantic text structure). The aim of this representation is to provide a rough determination of semantic units which may correspond to predicates, text roles, and possibly other units (such as argument roles which do not belong to the set of slots established by the text theme). It is a representation level which can be considered as an intermediate level between a text representation and a syntactic sentence representation.

Due to syntactic complexity there may be some problems of decomposition such as the treatment of attributes or of embedded propositions. Both can take the function of an own proposition which has to be attached to the superordinated proposition. This means that a text role can be filled by a whole proposition.

In a second step we will analyse the argumentative schema of the text. We will finally give a short description of the effects of the global text structure on the realization of the semantic text structure.

3.2 Example: Nostalgiekarte Jahrgang 1992

(01) Nostalgie-Karte Jahrgang 1992

(02) So war die politische Geographie einmal zu fernen Zeiten.
(03) Deutschland noch nicht vereint,
(04) der Saar-Lor-Lux-Raum ein weißer Fleck auf der Landkarte.
(05) Zu fernen Zeiten?
(06) Mitnichten!!!
(07) Die oben abgebildete Deutschlandkarte findet sich im neuen Taschen-Terminkalender 1992 der Sparkasse Saarbrücken.
(08) Dort hat man **noch nicht** mitbekommen,
(09) daß Deutschland um die Kleinigkeit von fünf neuen Bundesländern größer geworden ist.
(10) **Zudem** gibt es jenseits der alten DDR-Grenze noch andere Städte als Leipzig und Berlin, **so zum Beispiel** Rostock,Dresden, Magdeburg oder Saarbrückens Partnerstadt Cottbus.
(11) **Außerdem** scheint den Herren von der Sparkasse entgangen zu sein,
(12) daß am Ende des Terminello-Jahres 1992 der europäische Binnenmarkt steht.
(13) **Nicht zuletzt** vermittelt das Kärtchen den Eindruck,
(14) daß Saarbrücken der Nabel (Alt-)Deutschlands zu sein scheint.
(15) Je nach anatomischer Sichtweise kann es **aber** auch ein anderes Körperteil sein.

Rough translation of Nostalgie-Karte Jahrgang 1992:

(01) Nostalgia-Map of the Year 1992

(02) This is how the political geography used to be a long time ago.
(03) Germany not yet united,
(04) the Saar-Lor-Lux-area a white spot on the map.
(05) A long time ago?
(06) Far from it!!!
(07) The map of Germany shown above is published in the new pocket-agenda 1992 of the savings-bank Saarbrücken.
(08) Where no-one has **so far** noticed
(09) that Germany has grown by the trifling bit of five new Bundesländer.
(10) **Moreover**, there are still other cities apart from Leipzig and Berlin beyond the former GDR-frontier, **so for example** Rostock, Dresden, Magdeburg, or Cottbus, the twin city of Saarbrücken.
(11) **Besides**, it seems that the gentlemen of the savings-bank didn't realize
(12) that at the end of the Terminello-Year 1992 the Single European Market will come into force.
(13) **Last but not least** the little map suggests,
(14) that Saarbrücken was the navel of (the former) Germany.
(15) **But** depending on the anatomical point of view it might also refer to another body part.

3.3 Analysis of Propositions

```
STS: (list
     (P01 (PRED = _
                    (TROL1 = Nostalgie-Karte Jahrgang 1992)))
     (P02 (PRED = war
                    (TROL1 = so) [REFERENCE P03 and P04]
                    (TROL2 = die politische Geographie)
                    (TROL3 = einmal zu fernen Zeiten)
     (P03 (PRED = noch nicht vereint
                    (TROL1 = Deutschland))
     (P04 (PRED = ein weißer Fleck (sein)
                    (TROL1 = der Saar-Lor-Lux-Raum)
                    (TROL2 = auf der Landkarte)))
     (P05 (PRED = __
                    (TROL1 = Zu fernen Zeiten)?)
     (P06 (PRED = _
                    (TROL1 = Mitnichten)!!!)
     (P07 (PRED = findet sich
                    (TROL1 = die oben abgebildete Deutschlandkarte)
                    (TROL2 = im neuen Taschen-Terminkalender 1992
                             der Sparkasse Saarbrücken)))
     (P08 (PRED = hat noch nicht mitbekommen
```

(TROL1 = dort) [REFERENCE (P07), TROL2]
(TROL2 = man)
(TROL3 = daß (P09))))
(P09 (PRED = größer geworden ist
(TROL1 = Deutschland)
(TROL2 = um die Kleinigkeit von fünf neuen Bundesländern)))
(P10 [CONg = (UNIT: (P09)
REL: ADDITION
CONNEC: **zudem**
ORDER: CONNEC = TOPIC)]
(PRED = gibt
(TROL1 = es)
(TROL2 = jenseits der alten DDR-Grenze))
(TROL3 = noch andere Städte als Leipzig und Berlin
(TROL3' = Rostock, Dresden, Magdeburg oder
Saarbrückens Partnerstadt Cottbus
[CONl (UNIT: (TROL3, TROL3')
REL: EXEMPLIFICATION
CONNEC: **zum Beispiel**
ORDER: CONNEC = TOPIC)])
(P11 [CONg = (UNIT: (P10)
REL: ADDITION
CONNEC: **außerdem**
ORDER: CONNEC = TOPIC)]
(PRED = scheint entgangen zu sein
(TROL1 = den Herren (von der Sparkasse)))
(TROL2 = daß (P12)))
(P12 (PRED = steht
(TROL1 = am Ende des Terminello-Jahres 1992)
(TROL2 = der europäische Binnenmarkt)))
(P13 [CONg = (UNIT: (P11)
REL: ADDITION
CONNEC: **nicht zuletzt**
ORDER: CONNEC = TOPIC)]
(PRED = vermittelt
(TROL1 = das Kärtchen)
(TROL2 = den Eindruck)
(TROL3 = daß (P14)))
(P14 (PRED = zu sein scheint
(TROL1 = Saarbrücken)
(TROL2 = der Nabel (Alt)-Deutschlands))
(P15 (PRED = kann sein
(TROL1 = je nach anatomischer Sichtweise)
(TROL3 = es)))
[CONl: (UNIT = (P14)
REL: ADVERSATIVE

CONNEC: aber
ORDER: CONNEC not = TOPIC)])
(TROL2 = auch ein anderes Körperteil)

List of abbreviations:
CONNEC: connector (lexical entry)
CONg: global connector
CONl: local connector
ORDER: word order
P: proposition
PRED: predicate
REL: relation (between units)
TOPIC: topicalization
TROL: text role
UNIT: unit

3.4 The global structure of the text

The global structure of the text is developed on the basis of the semantic text structure presented in section 3.3.

In the case of an argumentative schema of a text one would expect that a conclusion is drawn by the author. As mentioned in section 2.2 this conclusion is often not realized explicitly but 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 of the text. If a map is nostalgic, it must be in some respects not current with the reality it should represent. This is in conflict with the expectation that an object labelled with a date should somehow pertain to that date. The reader thus can infer from the headline already that the presented map will be evaluated negatively. We can assume, then, that the conclusion of the text is evident in the headline and thus represents the first element of the global structure of the text. This element can be paraphrased as follows:

CONCLUSION: The presented map should be evaluated negatively. (P01)

From proposition P02 to proposition P04 a description of Germany is given which is no longer valid. Corresponding to this the predicate of P02 is realized in past tense and the situation is reflected lexically in TROL3 of P02 (*einmal zu fernen Zeiten*). For this part of the text we can formulate a paraphrase for the next section of the global structure:

the map reflected the geographical situation of Germany (P02 - P04)

For the institution that has published the map (Sparkasse Saarbrücken) this situation still seems to be relevant. Corresponding to this part of the text we can develop the next element of the global structure:

for the savings bank Saarbrücken it is true that: the map in question reflects the geographical situation of Germany (P05 - P07)

These two arguments do not aim at the implicit conclusion of the text (cf. section 2.4) and thus form the CONTRA-POSITION of this argumentative text (cf. argumentative schema figure 2). 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 argumentation. Consequently it is placed at the beginning of the text and the text-roles of this part of the text refer to objects/events/states which play a central role in the text. In our example the expressions *Karte* and *Deutschland* form the core of such text-roles (P03 TROL1, P04 TROL2, P 07 TROL1). These expressions are repeated in the further text and therefore contribute to the stability of the text (cf. section 2.2). Thus the global structure of the text is extended as follows:

CONCLUSION: The presented map should be valued negatively.
CONTRA-POSITION: the map reflected the geographical situation of Germany (P02 - P04)
for the savings bank Saarbrücken it is true that:
the map in question reflects the geographical situation of Germany (P05 - P07)

In this part of the text, connectors are not realized explicitly, but connections are construed implicitly by reference (cf. reference relation in P02).

In the next section of the text, the arguments against the assumption of the savings bank Saarbrücken are presented. Together they form the PRO-POSITION of the text. The text-roles in this part should evolve in line with the thematic progression of the text (cf. section 2.2). A new subject of consideration is presented in nearly each of the following propositions (cf. section 3.3 *um die Kleinigkeit von fünf neuen Bundesländern* (P09 TROL2), *jenseits der alten DDR-Grenze* (P10 TROL2), *noch andere Städte als Leipzig und Berlin* (P10 TROL3, *der europäische Binnenmarkt* (P12 TROL2), *der Nabel (Alt-)Deutschlands* (P14 TROL2). In order to explain the function of these propositions to formulate the thematic question of the text (cf. section 2.2) is helpful. Corresponding to the above presented first part of the global structure (CONCLUSION) the following question seems to be central:

THEMATIC QUESTION: 'What are the reasons for the negative evaluation of the map? '

The respective propositions give answers to this question:

- because Germany has grown by the trifling bit ...(P09)
- because there are still other cities beyond the former GDR-frontier...(P10)
- because the Single European Market ... (P12)
- because Saarbrücken was the navel of (the former) Germany (P14)

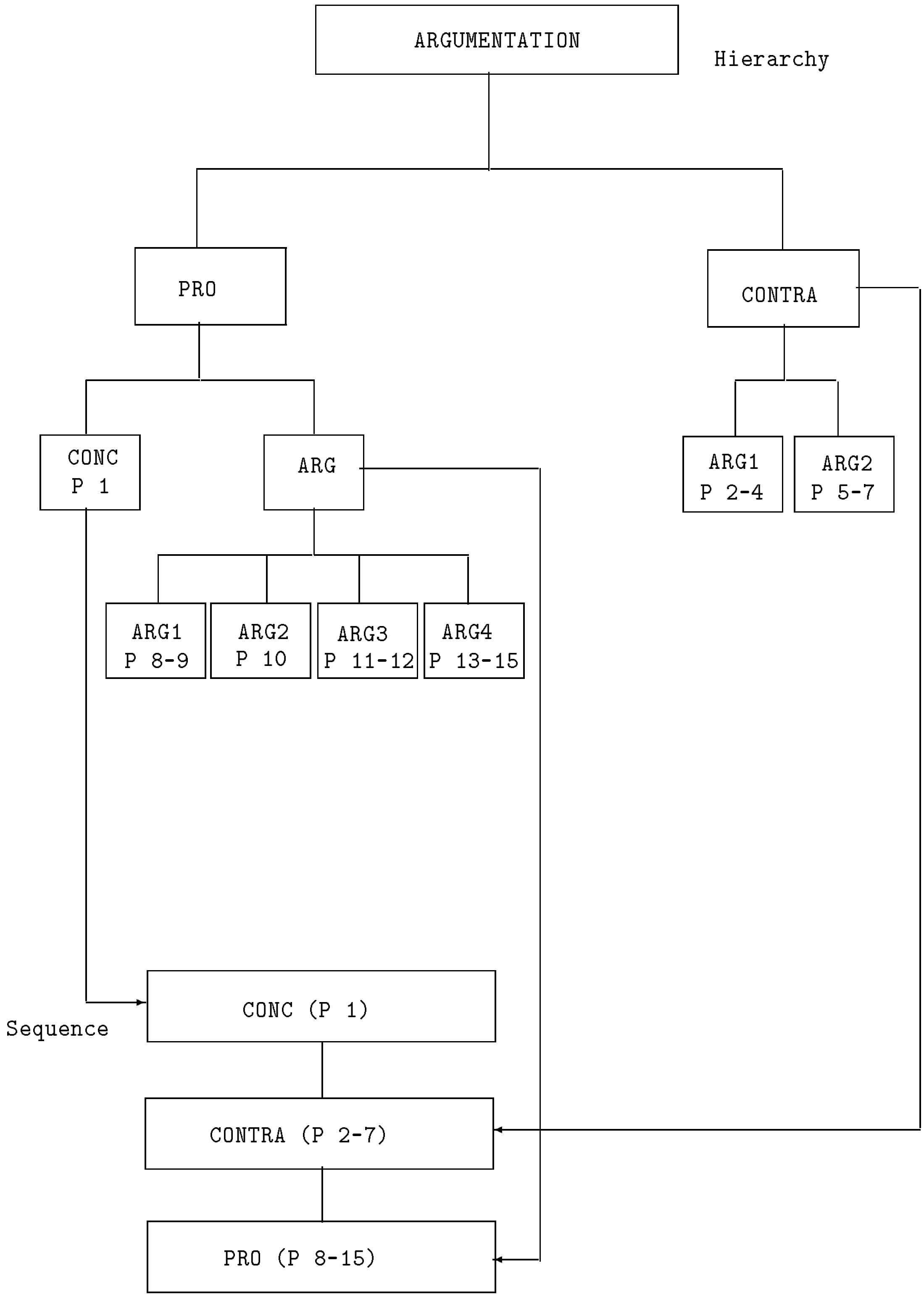
Now the representation of the global structure of the text can be completed:

CONCLUSION: The presented map should be evaluated negatively.
CONTRA-POSITION: the map reflected the geographical situation of Germany (P02 - P04)
for the savings bank Saarbrücken it is true that: the map in question reflects the geographical situation of Germany (P05 - P07)

PRO-POSITION: The map in question does not reflect the geographical situation of Germany (& set of arguments)

In figure 4 the hierarchical structure of the text and its sequential realization are presented.

Figure 4: Argumentative schema of the text



List of abbreviations:

ARG: argument(s)
CONC: conclusion
CONTRA: contra-position
P: proposition
PRO: pro-position

Now, a further look should be taken at the internal structure of the PRO-POSITION part of the text. All arguments give answers to the thematic question presented above and they all deal in some way with the central text-objects presented in the CONTRA-POSITION of the text: the political geography of Germany in relation to the picture of Germany drawn on the map. The consequence of this common reference is, that the arguments can be realized as an additive chain. For this purpose additive connectors are used, such as *zudem* (P10), *außerdem* (P11) and *nicht zuletzt* (P13). The specifications of these connections look very similar: they relate two units by addition and the connector is topicalized. The last argument in propositions P13 and P14 is attached by the connector *nicht zuletzt*. By using this connector some more details concerning the structure of this part of the text are given:

- this is the last argument of the argumentation (in case of the text *Nostalgie-Karte Jahrgang 1992* the whole text ends with this argument)
- this argument has to be valued at the same level as the other arguments.

The starting point of the additive chain is the predicate in proposition P08 *noch nicht mitbekommen*. Furthermore, the additive character of the argumentative part of the text is stressed, because all these global connectors (CONg) are topicalized. On the other hand, a topicalization of the local connector (CONl) *aber* in proposition P15 would destroy the additive chain, because no new argument, but a comment on the preceding argument is mentioned after the connector. Thus, we are able to distinguish between connectors which organize the global structure (the connectors which are used within the additive chain) of the text and connectors which organize information at a local level. A further example of connectors organizing the structure at a local level is *zum Beispiel* in proposition P10. We conclude that realizing the argumentative part of the text by means of additive chaining has great influence on:

- the choice of connectors which can be used and
- the possibility of their topicalization.

In our example there is no need to topicalize the additive connectors, but if they were the connectors organizing a substructure of the global structure, they could not be topicalized without destroying the chain that has been built up. Moreover, the topicalization of the additive connectors in this example emphasizes the argumentation of the PRO-POSITION.

The following figure shows the application of structure [5] in section 2.6 to this text:

```
[6] (CON =
        (GLOBAL =
                (SCHEMA = [argumentative])
                (RELATION = ADDITION)
                (CONNEC = {list (zudem, außerdem, nicht zuletzt) }
                          [Order: topicalized])
                (TEGIE = [Position: part PRO-POSITION of the text, P10-P15]))
        (LOCAL =
                (RELATION = ADVERSATIVE)
                (CONNEC = {list (aber) }
                          [Order: no topicalization])
                (TEGIE = [Position: part PRO-POSITION of the text, P15])
                (RELATION = [EXEMPLIFICATION])
                (CONNEC = { list (zum Beispiel)}
                          [Order: not topicalized])
                (TEGIE = [position: part PRO-POSITION of the text, P10])))
```

The analysis of the text Nostalgiekarte Jahrgang 1992 has shown that the semantic text structure provides an appropriate base for construing the global structure of a text. Further work has to be done for specifying the function of the text-roles in the various parts of the text.

Further texts (and texts of other text-types as well) will have to be analysed in order to identify various options for CHAINING in texts. According to the different ways of chaining, various classes of connectors should be distinguished and examined (for classification of connectors in German cf. e.g.[Heringer, 1989]).

4 Summary

In this paper we have shown that CHAINING represents a specific way to connect parts of a text on the basis of global coherence. We gave an outline of the specifications of CHAINING at the semantic and the pragmatic text level and related these specifications to schemata. On this basis we have demonstrated how the transformation of these schemata into sequences of particular sequential orders can be achieved by means of CHAINING. Finally, we have exemplarily shown how these principles of CHAINING can be applied to the analysis of a newspaper text.

References

[Brandt and Rosengren, 1991] M. Brandt and I. Rosengren. Handlungsstruktur und Informationsstruktur - zwei Seiten einer Muenze. Technical report, Sprache und Pragmatik, Arbeitsberichte, Lund, 1991.

[Brinker, 1992] K. Brinker. *Linguistische Textanalyse. Eine Einfuehrung in Grundbegriffe und Methoden.* E. Schmidt-Verlag, Berlin, 1992.

[Dahlgren, 1989] K. Dahlgren. Discourse coherence and naive semantics. Technical report, L.A .U.D. Paper no. A 271, Duisburg, 1989.

[Daneš, 1974] František Daneš. *Papers on Functional Sentence Perspective.* Academia, Prague, 1974.

[de Beaugrande and Dressler, 1981] R. de Beaugrande and W.U. Dressler. *Introduction to Text Linguistics.* Longmans, London, 1981.

[de Beaugrande, 1984] Robert de Beaugrande. *Text Production. Toward a Science of Composition.* Ablex, Norwood, New Jersey, 1984.

[Heinemann and Viehweger, 1991] Wolfgang Heinemann and Dieter Viehweger. *Textlinguistik: Eine Einfuehrung.* Niemeyer, Tuebingen, 1991.

[Heringer, 1989] H. J. Heringer. *Lesen lehren lernen: Eine rezeptive Grammatik des Deutschen.* Max Niemeyer Verlag, Tuebingen, 1989.

[Hobbs, 1979] J.R. Hobbs. Coherence and coreference. *Cognitive Science*, 3(1), 1979.

[Klein and von Stutterheim, 1991] W. Klein and C. von Stutterheim. Text structure and referential movement. Technical Report 22, Sprache und Pragmatik, Lund, 1991.

[Metzing, 1980] D. Metzing. *Frame Conceptions and Text Understanding.* Walter de Gruyter, Berlin, 1980.

[Minsky, 1975] Marvin Minsky. A framework for representing knowledge. In *Theoretical Issues in Natural Language Processing 1*, 1975. See also: paper of the same name in Winston, P. H. (ed.) *The Psychology of Computer Vision*, (1975) New York: McGraw-Hill. Reprinted in Ronald J. Brachman and Hector J. Levesque (eds.) *Readings in Knowledge Representation*, (1985) Los Altos: Kaufman.

[Minsky, 1990] M. Minsky. *Mentopolis.* Klett-Cotta, Stuttgart, 1990.

[Noel, 1990] D. Noel. Coherence relations and mental models. Technical report, Universitaet Antwerpen, Antwerpen, 1990.

[Norgard-Sorensen, 1992] J. Norgard-Sorensen. *Coherence Theory. The Case of Russian.* Mouton de Gruyter, Berlin, 1992.

[Oversteegen, 1989] E. Oversteegen. *Tracking Time. A proposal for the representation of temporal expressions in Dutch.* PhD thesis, University of Utrecht, 1989.

[Rothkegel, 1993] Annely Rothkegel. *Text Knowledge and Object Knowledge*. Pinter, London, 1993.

[Schaeffner and Wiesemann, 1988] Ch. Schaeffner and U. Wiesemann. Further remarks on political argumentative texts. *Zeitschrift fuer Anglistik und Amerikanistik*, 4:339–348, 1988.

[Schaeffner *et al.*, 1987] Ch. Schaeffner, G. M. Shreve, and U. Wiesemann. A procedural analysis of argumentative political texts. *Zeitschrift fuer Anglistik und Amerikanistik*, 2:105–117, 1987.

[Thompson and Mann, 1987] S. A. Thompson and W. C. Mann. Rhetorical structure theory: a framework for the analysis of texts. *Papers in Pragmatics*, 1(1):79–105, 1987.

[Tonfoni, 1990] G. Tonfoni. *Text representation systems*. Eurographica, Helsinki, 1990.

[Toulmin, 1959] S. Toulmin. *The uses of argument*. Cambridge University Press, 1959.

[van Dijk, 1980] Teun A. van Dijk. The semantics and pragmatics of functional coherence in discourse. *Versus*, (26/27):49–65, 1980.

[Von Stutterheim, 1992] Ch. Von Stutterheim. Quaestio und Textstruktur. In H. P. Krings and G. Antos, editors, *Textproduktion. Neue Wege der Forschung*, pages 159–171. Wissenschaftlicher Verlag, Trier, 1992.