Universität Bielefeld

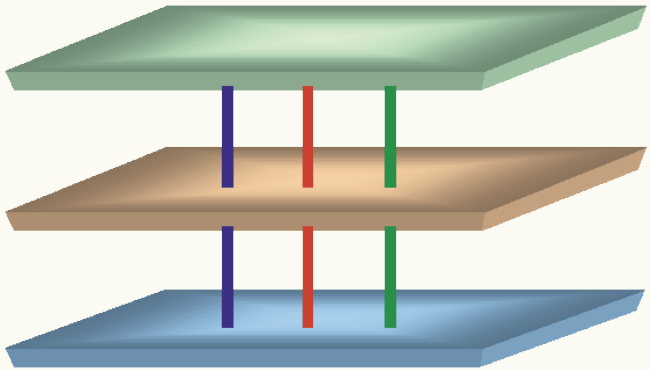
JUSTUS-LIEBIG-UNIVERSITÄT GIESSEN

Anaphoric relations as cues for rhetorical relations

Daniela Goecke, Harald Lüngen,

Maja Bärenfänger, Mirco Hilbert, Maik Stührenberg

Justus Liebig University Giessen University Bielefeld



DFG-Research Group

Text-technological modelling of information

Overview

Introduction

Background and Motivation

Corpus study

Experiment

Conclusions and Outlook

Introduction

Introduction: Projects

DFG Research Group *Text-technological Modelling of Information*

- 2nd Phase 2005-2008

Project A2 - Sekimo: *Secondary structuring of information and comparative analysis of discourse*

Project C1 - SemDok: *Generic document structures in linearly organised texts*

Introduction: Project *Sekimo* (A2): Goals

Integration of heterogeneous, XML-based resources

- e.g. Lexicon, ontology, parser

- Application domain: Resolution of definite description anaphora
- In order to resolve anaphoric relations, many pieces of linguistic and textual information are required
  - POS, agreement information, grammatical function, ontological knowledge, distance measures, semantic similiarity
- Information is built up by markup unification
- XML element and attribute information is interpreted as feature vectors
- Training / Evaluation of decision trees

Introduction: Project *SemDok* (C1): Goals

Generic document structures in linearly organised text

Goals:

- Develop a text (discourse) parser for a complex genre
  - Framework: Rhetorical Structure Theory (RST)
  - Corpus: Scientific articles – complex genre, i.e. take into account
    - logical document structure
    - genre-specific text type structure
    - thematic and anaphoric structure
- Use text-technological (XML-based) formalisms and methods

Introduction:Corpus

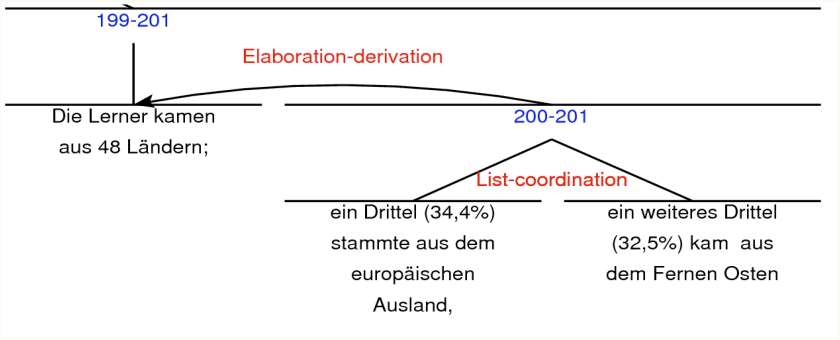
Corpus used in both Sekimo and SemDok:

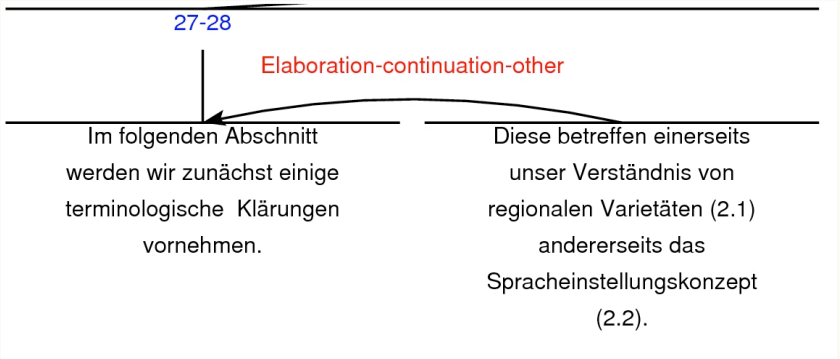
- 47 German linguistic journal articles
- + 1 newspaper article, + 1 web-published scientific article
- Annotated on several descriptive levels
- Ongoing annotation of RST structure and anaphoric structure

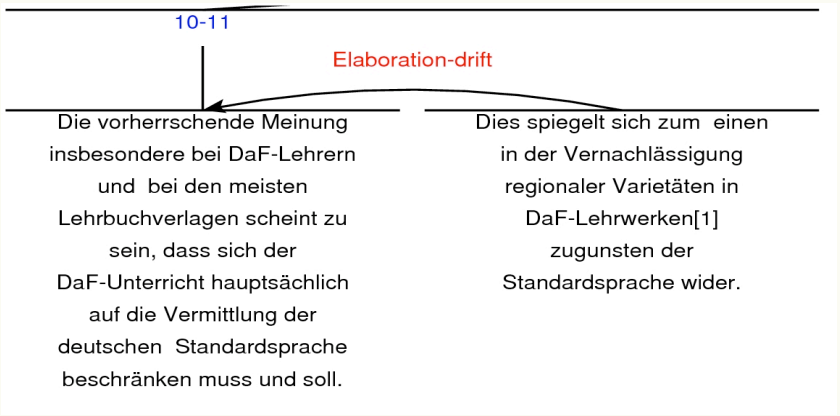
Background and Motivation

Rhetorical relations: Taxonomy for *Elaboration*

- elaboration-derivation
  - *elaboration-set-member
  - *elaboration whole-part
  - *elaboration-class-subclass
  - *elaboration-class-instance
  - *elaboration-process-step
- elaboration-integration
  - *elaboration-part-whole
- elaboration-identity
  - elaboration-continuation
    - elaboration-drift
    - elaboration-continuation-other
  - elaboration-specification
    - elaboration-assign
      - elaboration-assign-abbr
      - elaboration-assign-other
    - elaboration-specification-other
- elaboration-example
- elaboration-definition
- elaboration-restatement







Rhetorical structure: Sample Annotation

```
<para id="i734" relname="list-coordination" >
  <n>
    <hypo id="i730" relname="elaboration-continuation-other">
      <n>
        <t id="i27">Im folgenden Abschnitt werden wir zunächst einige terminologische Klärungen
          vornehmen. </t>
      </n>
      <s>
        <t id="i28">Diese betreffen einerseits unser Verständnis von regionalen Varietäten (2.1)
          andererseits das Spracheinstellungskonzept (2.2). </t>
      </s>
    </hypo>
  </n>
  <n>
    <t id="i29">Abschnitt 3 ist der Darstellung unseres methodischen Vorgehens und der
      Beschreibung der Untersuchungsgruppe gewidmet. </t>
  </n>
</para>
```

Anaphoric relations: Taxonomy

- cospecLink
  - ident
  - propName
  - synonym
  - hyperonym
  - hyponym
  - paraphrase
  - addInfo
  - isA
- bridgingLink
  - poss
  - setMember
  - hasMember
  - meronym
  - holonym
  - bridging

- cospecLink:
  referential identity between anaphoric expression and antecedent
- bridgingLink:
  no referential identity; the antecendent has to be inferred from the context (`bridge the gap‘; Clark 1977)
- Anaphoric relations hold between discourse entities (discourse referents, Karttunen 1976, Kamp & Reyle 1993)

Anaphoric relations: Sample CHS annotation

```
<chs:chs>
    <chs:text>
    Im folgenden Abschnitt werden wir zunächst
        <chs:de deID="de114" deType="nom">
            einige terminologische
            <chs:token pos="N" syntax="@NH" lemma="klärung"
            morpho="FEM PL ACC" id="w415">Klärungen</chs:token>
        </chs:de> vornehmen.
        <chs:de deID="de115" deType="nom">
            <chs:token pos="PRON" syntax="@NH" lemma="dieser“
            dependValue="subj" morpho="Dem FEM PL NOM" >Diese</chs:token>
        </chs:de>
    betreffen einerseits Verständnis von regionalen Varietäten, andererseits das
    Spracheinstellungskonzept.
    </chs:text>

    <chs:semRel>
        <chs:cospecLink id="sr37" relType="ident"
             phorIDRef="de115"  antecedentIDRefs="de114" />
     </chs:semRel>
</chs:chs>
```

Anaphoric relations: Sample CHS annotation

```
<chs:chs>
    <chs:text>
    Im folgenden Abschnitt werden wir zunächst
        <chs:de deID="de114" deType="nom">
            einige terminologische
            <chs:token pos="N" syntax="@NH" lemma="klärung"
            morpho="FEM PL ACC" id="w415">Klärungen</chs:token>
        </chs:de> vornehmen.
        <chs:de deID="de115" deType="nom">
            <chs:token pos="PRON" syntax="@NH" lemma="dieser“
            dependValue="subj" morpho="Dem FEM PL NOM" >Diese</chs:token>
        </chs:de>
    betreffen einerseits Verständnis von regionalen Varietäten, andererseits das
    Spracheinstellungskonzept.
    </chs:text>

    <chs:semRel>
        <chs:cospecLink id="sr37" relType="ident"
            phorIDRef="de115"  antecedentIDRefs="de114" />
     </chs:semRel>
</chs:chs>
```

Background and Motivation: Discourse Parsing

- Automated analysis of discourse structure usually based on lexical discourse markers and grammar
  - → not sufficient for discourse relations like *Elaboration*
- *Elaboration*:
  - most frequent relation in corpus (25% of all relation instances of subcorpus)
  - common at all levels of the discourse structure
  - "default relation": unspecific; overridden by more specific relations
  - holds between adjacent discourse segments with a common topic
    - Indicated by "subject continuity" (Corston-Oliver 1998)
    - Indicated by lexical cohesion (Marcu 2000)
  - mirrors the focus structure of a text (Knott et al. 2001)

Background and Motivation: Anaphora Resolution

Various studies suggest that discourse structure constrains the structure of anaphoric relations

- ... in terms of modelling the accessibility of an antecedent candidate, e.g. in terms of the Right Frontier Constraint (Polanyi 1988, Asher & Lascarides 2003)
- ... in order to optimise the set of candidates: "[…] when hierarchical adjacency is considered, an anaphor may be resolved to a referent that is not the closest in a linear interpretation of a text“ (Cristea et al. 1999)

Rhetorical structure vs. Anaphoric structure

- Besides modelling discourse structure, the projects *Sekimo* and *SemDok* focus on automatic discourse analysis
- I.e. the availability of resources is crucial
- Therefore:
  - Logical document structure is easily available for an automatic analysis of anaphoric relations (Goecke & Witt 2006)
  - Anaphoric relations should serve as a resource for an automatic analysis of rhetorical structures

Corpus Study

Corpus Study: Hypotheses and Research Questions

Hypotheses:

- Anaphoric relations holding between discourse entities are cues for the rhetorical relation *Elaboration* to hold between text segments
- Specific anaphoric relations are cues for specific subrelations of *Elaboration*

RST-Relations	CHS-Relations
Elaboration-derivation	cospec:hyperonym bridging:holonym bridging:setMember
Elaboration-integration	cospec:hyponym bridging:meronym bridging:hasMember
Elaboration-continuation	cospec:ident cospec:synonym cospec:paraphrase
Elaboration-drift	– bridging:bridging bridging:poss bridging:abstrProp bridging:abstrCluster

Corpus Study: Hypotheses and Research Questions

- Research Questions:
  - **A:** Is an anaphoric relation between discourse entities in two discourse segments a sufficient condition for *Elaboration?*
  - **B:** Is an anaphoric relation between discourse entities in two discourse segments a necessary condition for *Elaboration* to hold?

Corpus Study: Method

For a subcorpus of 4 texts: [Σ 15622 words]

- Queries for relations between elements on three XML annotation layers of the same textual base
  - SEG (simple and complex discourse segments)
  - RST-HP (RST discourse structure)
  - CHS (Cohesion, i.e. discourse entities and anaphoric relations)

  Using *Markup Unification* and *Sekimo Inference Tools*, cf. Goecke et al. 2005

→ relationInstances.xml (Analysis A)
→ anaphoricInstances.xml (Analysis B)

→ XPath expressions to obtain statistics

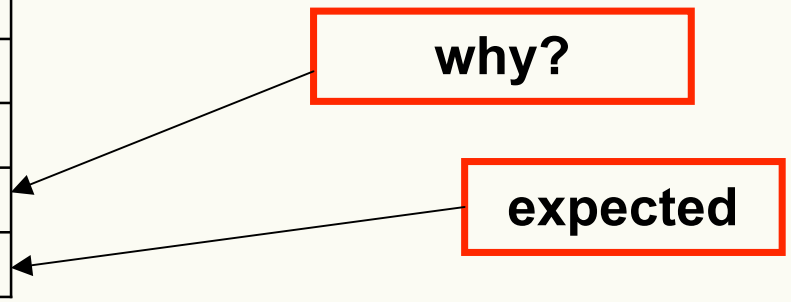
Corpus Study: Example of an *anaphoricInstance*

```
<anaphoricInstance count="ling_deu_003-4" rtype="elaboration-drift">
    <segment1>Die vorherrschende Meinung insbesondere bei DaF-Lehrern und bei den meisten
      Lehrbuchverlagen scheint zu sein , dass sich der DaF-Unterricht hauptsächlich auf die
      Vermittlung der deutschen Standardsprache beschränken muss und soll.</segment1>
    <segment2>Dies spiegelt sich zum einen in der Vernachlässigung regionaler Varietäten in
      DaF-Lehrwerken[ 1 ] zugunsten der Standardsprache wider . Zum anderen äußern Lehrer
      häufig Vorbehalte, da sie es aufgrund des hohen zeitlichen Aufwands für unrealistisch
      halten, sich neben der Standardsprache noch mit weiteren sprachlichen Varietäten zu
      beschäftigen, zumal die Standardsprache die überregionale Kommunikationsfähigkeit
      garantieren soll.</segment2>
    <anaphora atype="cospec:abstrProp">
      <antecedent>Die vorherrschende Meinung insbesondere bei DaF-Lehrern und bei den
  meisten
      Lehrbuchverlagen scheint zu sein, dass sich der DaF-Unterricht hauptsächlich auf die
      Vermittlung der deutschen Standardsprache beschränken muss und soll.</antecedent>
      <anaphor>Dies</anaphor>
    </anaphora>
  </anaphoricInstance>
```

Corpus Study: Results for A (sufficient condition)

Question: Is the existence of an anaphoric relation a **sufficient** condition for *Elaboration* to hold between discourse segments?

	Total No.
anaphoricInstance	662
@rtype=elaboration	176
@rtype=no-RST-relation	301
@rtype=other-RST-relation-than-elaboration	185



Corpus Study: Results for A - Example

```
<anaphoricInstance count="luckhardt-1" rtype="elaboration-drift">
    <segment1>Das World-Wide Web, in dem Sie sich gerade bewegen , ist ein Hypertext:ein Text auf einem
    elektronischen Medium mit Querverweisen zu anderen Texten. Dahinter steckt eine Idee, deren Ursprünge auf
    Vannevar Bush zurückgehen , der von einer Vernetzung des gesamten Wissens der Menschheit geträumt
    hat.</segment1>
    <segment2>Soweit sind wir zwar noch nicht , immerhin gibt es das WWW erst seit wenigen Jahren, aber das Konzept
    "Hypertext" scheint eine gute Grundlage für die Idee Bushs zu sein .</segment2>
    <anaphora atype="bridging:poss">
        <antecedent>Vannevar Bush</antecedent>
        <anaphor>die Idee Bushs</anaphor>
    </anaphora>
    <anaphora atype="cospec:paraphrase">
        <antecedent>Das World-Wide Web</antecedent>
        <anaphor>das WWW</anaphor>
    </anaphora>
</anaphoricInstance>
```

Segment 1 = different

Segment 2 = identical

```
<anaphoricInstance count="luckhardt-2" rtype="NO_RHETORICAL_RELATION">
    <segment1>Dahinter steckt eine Idee, deren Ursprünge auf Vannevar Bush zurückgehen , der von einer Vernetzung
    des gesamten Wissens der Menschheit geträumt hat .</segment1>
    <segment2>Soweit sind wir zwar noch nicht , immerhin gibt es das WWW erst seit wenigen Jahren, aber das Konzept
    "Hypertext" scheint eine gute Grundlage für die Idee Bushs zu sein.</segment2>
    <anaphora atype="bridging:poss">
        <antecedent>Vannevar Bush</antecedent>
        <anaphor>die Idee Bushs</anaphor>
    </anaphora>
</anaphoricInstance>
```

Corpus Study: Results for A (sufficient condition)

- *Question:* Is the existence of an anaphoric relation a **sufficient** condition for *Elaboration* to hold between discourse segments?

	Total No.
anaphoricInstance	662
@rtype=elaboration	176
@rtype=no-RST-relation	301
@rtype=other-RST-relation-than-elaboration	185

- *Results:*
  - not all discourse segment pairs with anaphoric relations are RST discourse trees at the same time
  - Occurrence of an anaphoric relation is not a sufficient condition for *Elaboration*
  - Interaction with other cues and constraints

Corpus Study: Results for B (necessary condition)

Is an anaphoric relation a **necessary** condition for *Elaboration* to hold between discourse segments?

	all	With anaphoric relations	Without anaphoric relations
Elaboration-Trees	**298**	**191**	**107**
elaboration-drift		67	44
elaboration-continuation-other		52	5
elaboration-derivation		30	8
elaboration-specification-other		20	23
elaboration-definition		6	7
elaboration		4	1
elaboration-identity		4	3
elaboration-example		4	6
elaboration-integration		4	3
elaboration-restatement		1	2
elaboration-assign-oth		1	5

NO

Corpus Study: Results for B (necessary condition)

Is an anaphoric relation a **necessary** condition for *Elaboration* to hold between discourse segments?

	all	With anaphoric relations	Without anaphoric relations
Elaboration-Trees	**298**	**191**	**107**
elaboration-drift		67	44
elaboration-continuation-other		52	5
elaboration-derivation		30	8
elaboration-specification-other		20	23
elaboration-definition		6	7
elaboration		4	1
elaboration-identity		4	3
elaboration-example		4	6
elaboration-integration		4	3
elaboration-restatement		1	2
elaboration-assign-oth		1	5

NO

Almost always accompanied by anaphoric relations

Corpus Study: Results for B (necessary condition)

	No. of	With anaphoric relations	Without anaphoric relations
Elaboration-Trees	**298**	**191**	**107**
elaboration-drift		67	44
elaboration-continuation-other		52	5
elaboration-derivation		30	8
elaboration-specification-other		20	23
elaboration-definition		6	7
elaboration		4	1
elaboration-identity		4	3
elaboration-example		4	6
elaboration-integration		4	3
elaboration-restatement		1	2
elaboration-assign-oth		1	5

no referential continuation expected; instead: lexical and grammatical cues for exemplification, specification or definition

Corpus Study: Results for B (necessary condition)

	No. of	With anaphoric relations	Without anaphoric relations
Elaboration-Trees	**298**	**191**	**107**
elaboration-drift		67	44
elaboration-continuation-other		52	5
elaboration-derivation		30	8
elaboration-specification-other		20	23
elaboration-definition		6	7
elaboration		4	1
elaboration-identity		4	3
elaboration-example		4	6
elaboration-integration		4	3
elaboration-restatement		1	2
elaboration-assign-oth		1	5

SHOULD almost always be indicated by anaphoric relations – why is this not the case?

Corpus Study: Results for B (necessary condition)

- existing anaphoric annotation: only anaphoric relations between nominal discourse entities
- Elaboration-drift: often based on a relation between a nominal/ adverbial anaphora and a propositional antecedent
- Peformed post-analysis of those Elaboration-trees without anaphoric relations between discourse segments (107)
- Results:

	Propositional Antecedent	Punctuation	Parenthesis	Other-Ana	No-Ana
107	37	11	20	31	8
100%	34.58%	10.28%	18.69%	28.93%	7.48%

Cue for Elaboration-drift

Corpus Study: Results for B (necessary condition)

- existing anaphoric annotation: only anaphoric relations between nominal discourse entities
- Elaboration-drift: per definition based on a relation between a nominal/ adverbial anaphora and a propositional antecedent
- Need for a post-analysis of those Elaboration-trees without anaphoric relations between discourse segments (107)
- Results:

	Propositional Antecedent	Punctuation	Parenthesis	Other-Ana	No-Ana
107	37	11	20	31	8
100%	34.58%	10.28%	18.69%	28.93%	7.48%

Discourse marker for elaboration-specification

Corpus Study: Results for B (necessary condition)

- Result:

  The existence of an anaphoric relation between discourse entities in two discourse segments is (approximately) a necessary condition for an *Elaboration relation* not based on other discourse markers to hold between these discourse segments

Corpus Study: Results for B (necessary condition)

- Result II:

  Elaboration relations like continuation, derivation and integration correspond to specific anaphoric relations

RST-Relations	CHS-Relations
Elaboration-continuation (52)	46 (+25) x cospec:ident 6 x cospec:paraphrase 3 x cospec:synonym 1 x cospec:addInfo
Elaboration-derivation (30)	19 (+26) x bridging:setMember 3 x cospec:isA 1 x bridging:meronym 1 x bridging:poss
Elaboration-integration (4)	3 x bridging:hasMember 1 x cospec:hyponym

Experiment

Experiment: Discourse Parser in Project SemDok

- Method
  - Cascaded bottom-up passive chart parser GAP (Generalised Annotation Parser), based on a segmentation of the input text; is called in each cascade for each “containing segment”
  - Different annotation layers provide cues and contraints to select the appropriate Reduce Rules
  - The resulting chart forms a parse forest which can be traversed to export alternative RST structures
- Status
  - Set of Reduce Rules for discourse parsing on CDS [block] level (paragraphs)
- Future:
  - Integration of history lists of discourse markers used
  - Integration of scoring scheme for rules

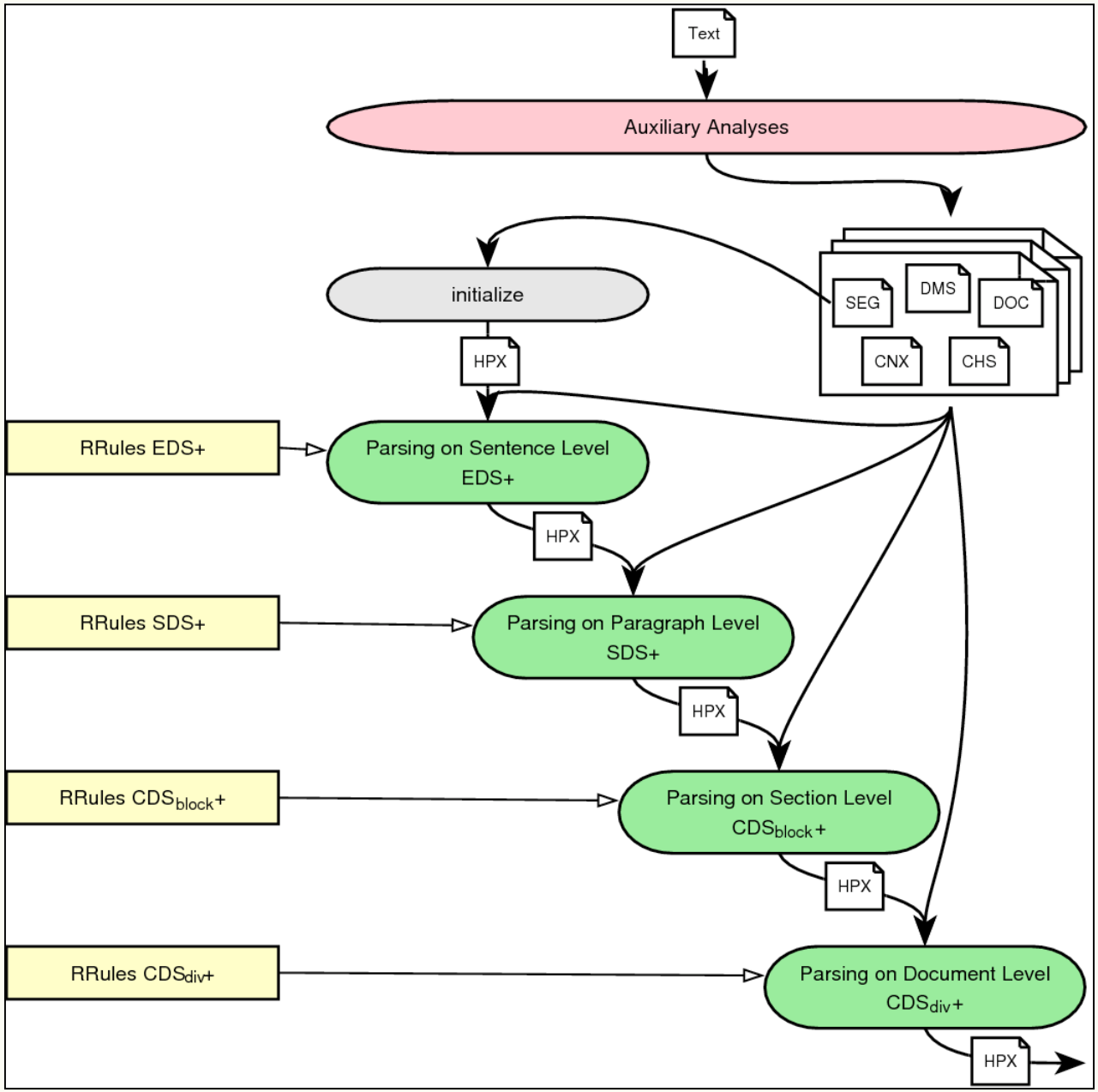
Experiment: Discourse Parser: Cascade architecture

Elementary discourse segments

Sentential discourse segments

Complex (block level) discourse segments

Complex (division level) discourse segments



Experiment: Introduction of Reduce Rules for CHS

Introduced two Reduce Rules that operate on the anaphoric relations annotation layer

Introduced a ranking of rule sets:

- Segment pair will only be tested against reduce rules of higher rank if no rules of lower rank have matched previously
- For the rule set for sentential discourse segments (SDS) in the Evaluation, the following ranking scheme was adopted:
  - Rules based on lexical discourse markers
  - Rules based on occurrences of anaphora (newly introduced)
  - Default rule

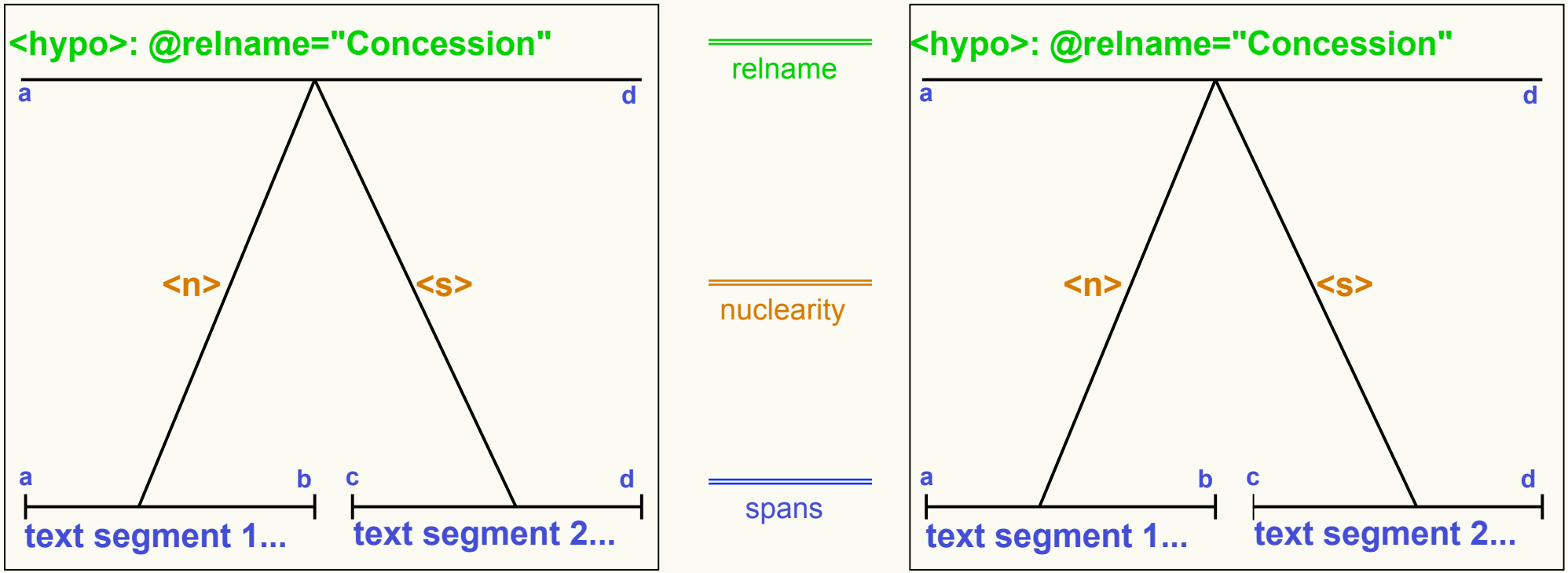
Experiment: Reduce Rules in GAP

For the experiment, two new rules were introduced based on the findings of the corpus study

Rule	*Reduce target*	*Cues*
Rule A1	Elaboration-continuation N-S	cospec:ident *OR* (corpus: 46 of 52) cospec:paraphrase *OR* cospec:synonym *OR* cospec:addInfo *AND* anaphoric expression in vorfeld or subject of 1st sentence (42 of 52)
Rule A2	Elaboration-derivation N-S	bridging:setMember *OR* (19 of 30) bridging:meronym *OR* bridging:poss *AND* anaphoric expression in vorfeld or subject of 1st sentence (16 of 30)

Experiment: Evaluation method

- Comparison with manual reference annotation of RST structure
- Precision and Recall, by matching RST SubTrees



MATCH

Experiment: Evaluation results

- Parsing on CDS [block]-level, i.e. max. segments = paragraphs
- 1 article, Reference annotation contains 151 RST-Subtrees

		Anaphora processing	Default Relation	Precision	Recall
Baseline	Ia	No	List-coordination [1)]		33,72
Baseline	Ib	No	Elaboration-drift [2)]		36,05
Informed	II	Any, anywhere in segment-2 -> Elaboration-drift	List-coordination	–	34,88
Informed	IIIa	Anaphoric types in rules for E-cont und E-derivation	List-coordination	+	40,69
Informed	IIIb	Anaphoric types in rules for E-cont and E-derivation + vorfeld constraint [3)]	List-coordination	+	38,37

- 2nd most frequent relation in corpus
- Most frequently occurring subtype of Elaboration in corpus
- Anaphor must occur in vorfeld of 1st sentence of segment-2

Experiment: Evaluation results

S = Supertype: only *Elaboration*, no subtypes

		Anaphora processing	Default Relation	Precision	Recall
Baseline	S-Ia	No	List-coordination		37,21
Baseline	S-Ib	No	Elaboration		37,21
Informed	S-II	Any, anywhere in Segment 2	List-coordination	+	41,86
Informed	S-IIIa	Anaphoric types in rules for Elaboration	List-coordination	+	46,51
Informed	S-IIIb	Anaphoric types in rules for Elaboration + vorfeld constraint	List-coordination	+	44,18

Conclusions and Outlook

Conclusions and Outlook

Conclusions

- Anaphoric relations are (approximately) necessary conditions for Elaboration relations (other signals absent)
- Certain subtypes of anaphoric relations are cues for certain subtypes of Elaboration relations
  - Especially the major subtypes Continuation, Derivation, and Drift can be related to (subtypes of) anaphoric relations
- Initial experiments suggest that it might be worth while to integrate anaphora resolution as an auxiliary analysis component into a discourse parser

Conclusions and Outlook

Outlook

- Completion of discourse parser (C1)
- Completion of anaphora resolution (A2)
- Integrating more anaphora-based rules from corpus study
- Evaluation on bigger corpus
- Perform similar comparative corpus analyses with *thematic structure* as defined in HyTex project (U Dortmund)

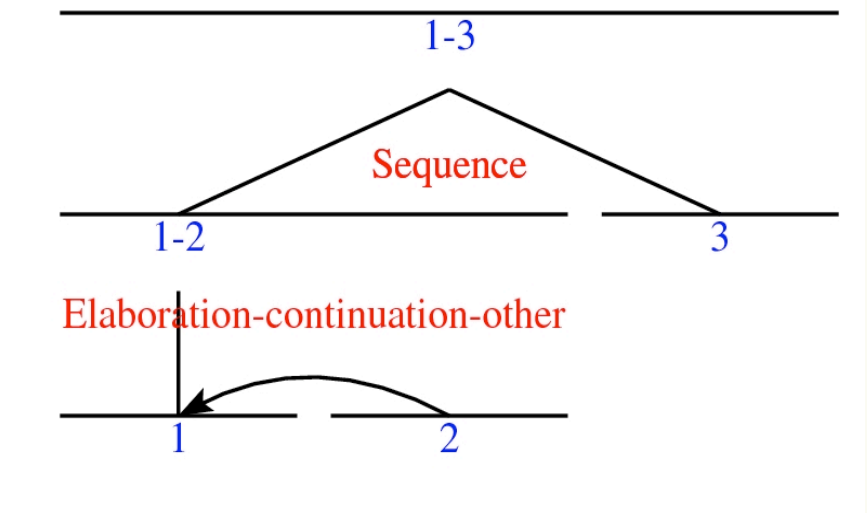
Additional slides

Corpus Study: Results for B (necessary condition)

```
<anaphoricInstance count="ling_deu_003-4" rtype='elaboration-drift'>
    <segment1>Die vorherrschende Meinung insbesondere bei DaF-Lehrern und bei den meisten
      Lehrbuchverlagen scheint zu sein , dass sich der DaF-Unterricht hauptsächlich auf die
      Vermittlung der deutschen Standardsprache beschränken muss und soll.</segment1>
    <segment2>Dies spiegelt sich zum einen in der Vernachlässigung regionaler Varietäten in
      DaF-Lehrwerken[ 1 ] zugunsten der Standardsprache wider . Zum anderen äußern Lehrer
      häufig Vorbehalte, da sie es aufgrund des hohen zeitlichen Aufwands für unrealistisch
      halten, sich neben der Standardsprache noch mit weiteren sprachlichen Varietäten zu
      beschäftigen, zumal die Standardsprache die überregionale Kommunikationsfähigkeit
      garantieren soll.</segment2>
    <anaphora atype='cospec:abstrProp'>
      <antecedent>Die vorherrschende Meinung insbesondere bei DaF-Lehrern und bei den meister
      Lehrbuchverlagen scheint zu sein, dass sich der DaF-Unterricht hauptsächlich auf die
      Vermittlung der deutschen Standardsprache beschränken muss und soll.</antecedent>
      <anaphor>Dies</anaphor>
    </anaphora>
  </anaphoricInstance>
```

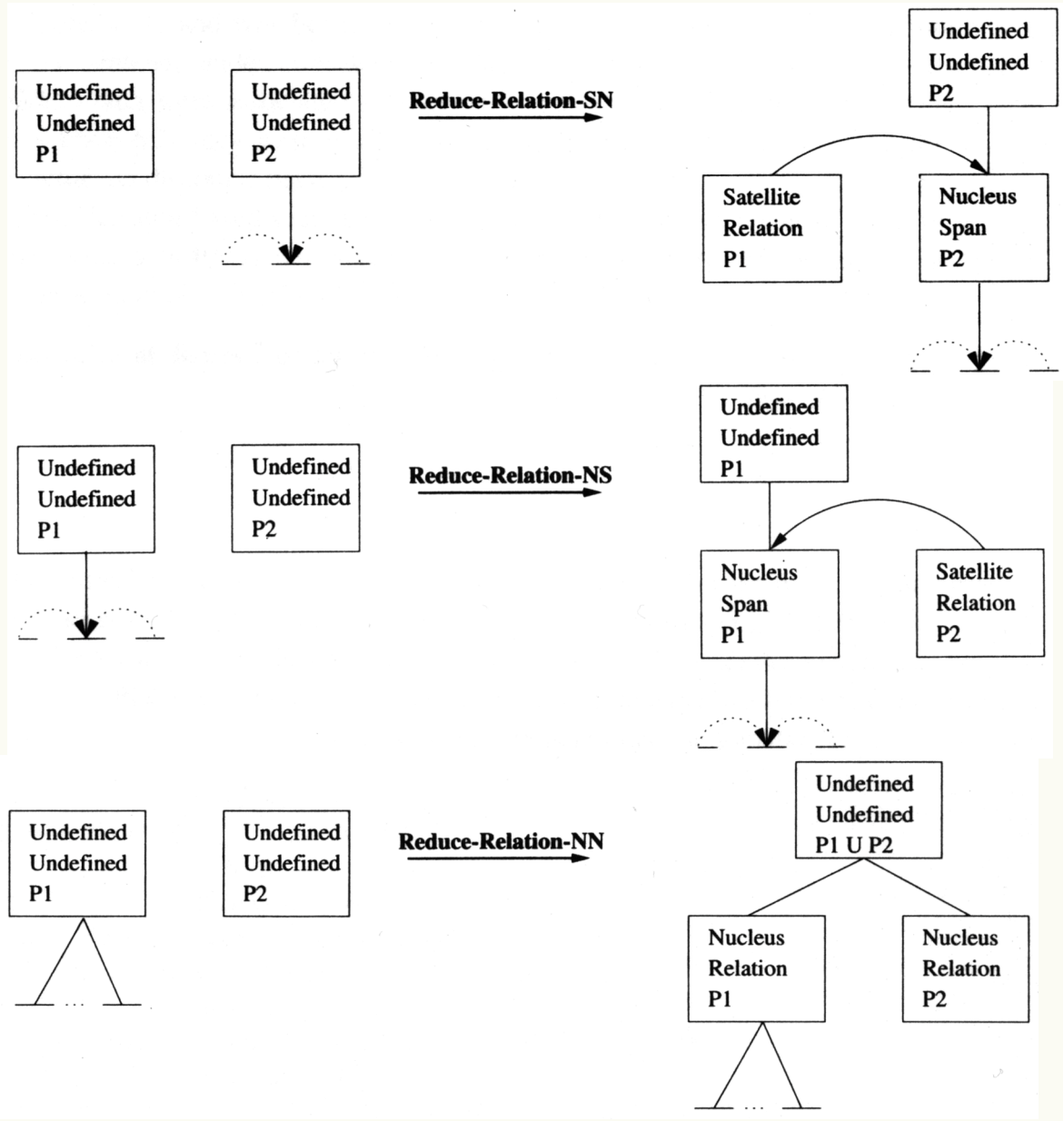
Rhetorical structure: Sample Analysis

- Im folgenden Abschnitt werden wir zunächst einige terminologische Klärungen vornehmen.
- Diese betreffen einerseits unser Verständnis von regionalen Varietäten (2.1) andererseits das Spracheinstellungskonzept (2.2).
- Abschnitt 3 ist der Darstellung unseres methodischen Vorgehens und der Beschreibung der Untersuchungsgruppe gewidmet



Experiment: Reduce Rules in GAP

- 5 rule types according to Marcu (2000) and others
- Rule acquisition and generation from discourse marker lexicon is ongoing
- Status in April 2007
  - 93 generated rules based on lexical discourse markers
  - 8 manually added rules referring to annotation of grammatical or document structure



Corpus Study: Hypotheses and Research Questions

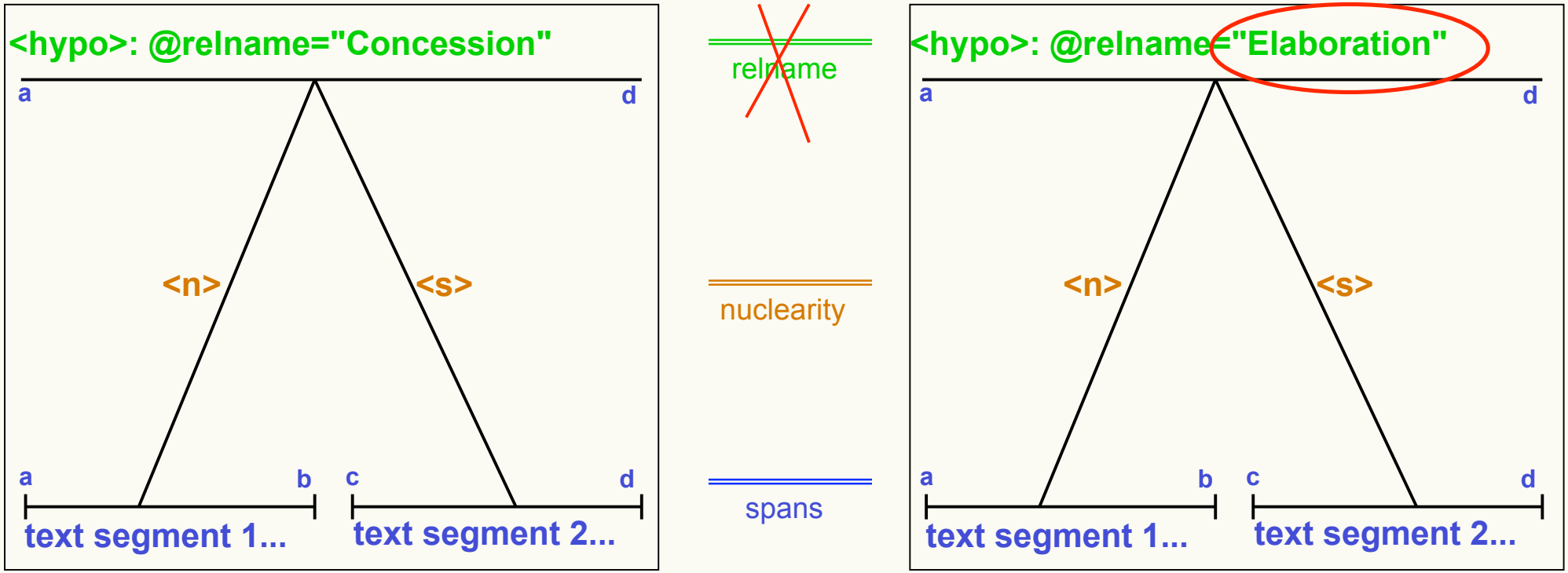
- Research Questions:
  - **A:** Is an anaphoric relation between discourse entities in two discourse segments a **sufficient** condition for *Elaboration?*
  - **B:** Is an anaphoric relation between discourse entities in two discourse segments a **necessary** condition for *Elaboration* to hold*?*
- Query formulation:
  - **A:**
    Given: Adjacent discourse segments containing discourse entities between which an anaphoric relation holds
    - Does a rhetorical relation hold? Yes/ No
    - Which one? Elaboration / Not Elaboration
  - **B:**
    Given: RST-Subtrees (only binary ones considered)
    - Which kind of rhetorical relation? Elaboration / Not Elaboration
    - Does an anaphoric relation hold between the two child segments? Yes/ No

Corpus Study: Results for B (necessary condition)

	No. of	With anaphoric relations	Without anaphoric relations
Elaboration-Trees	**298**	**191**	**107**
elaboration-drift		67	44
elaboration-continuation-other		52	5
elaboration-derivation		30	8
elaboration-specification-other		20	23
elaboration-definition		6	7
elaboration		4	1
elaboration-identity		4	3
elaboration-example		4	6
elaboration-integration		4	3
elaboration-restatement		1	2
elaboration-assign-oth		1	5

Experiment: Evaluation method

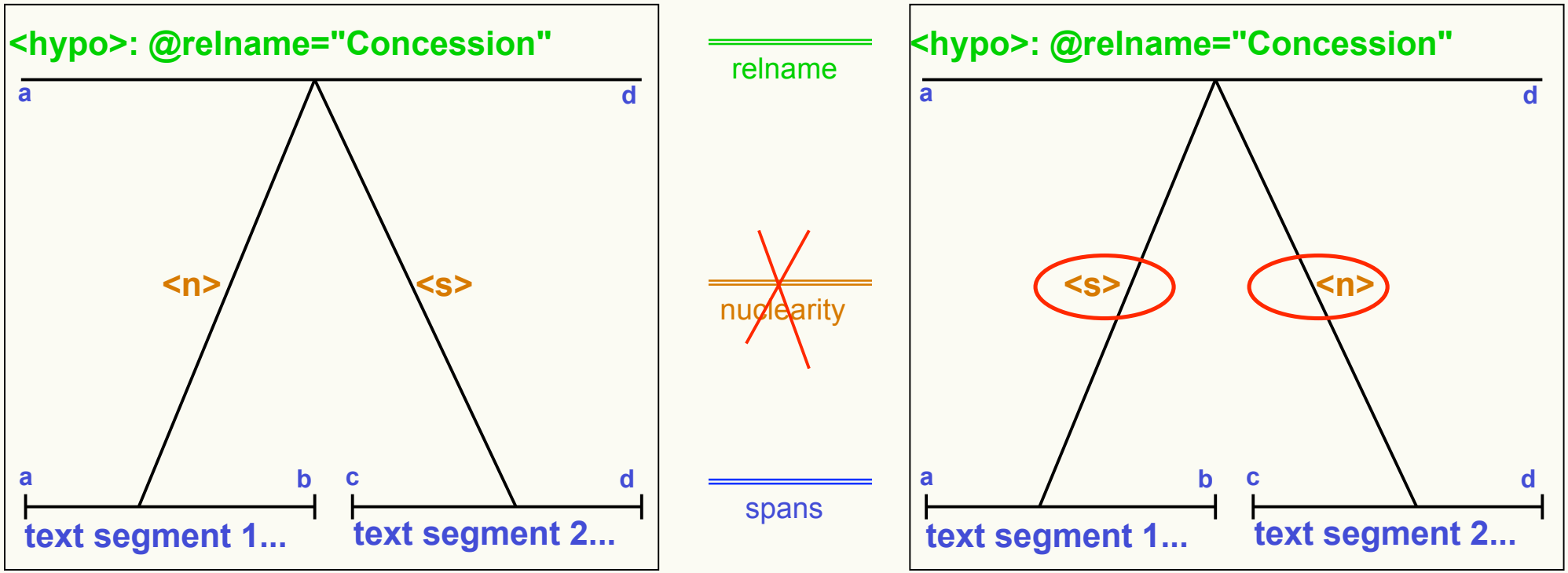
- Comparison with manual reference annotation of RST structure
- Precision and Recall, by matching RST SubTrees



NO MATCH

Experiment: Evaluation method

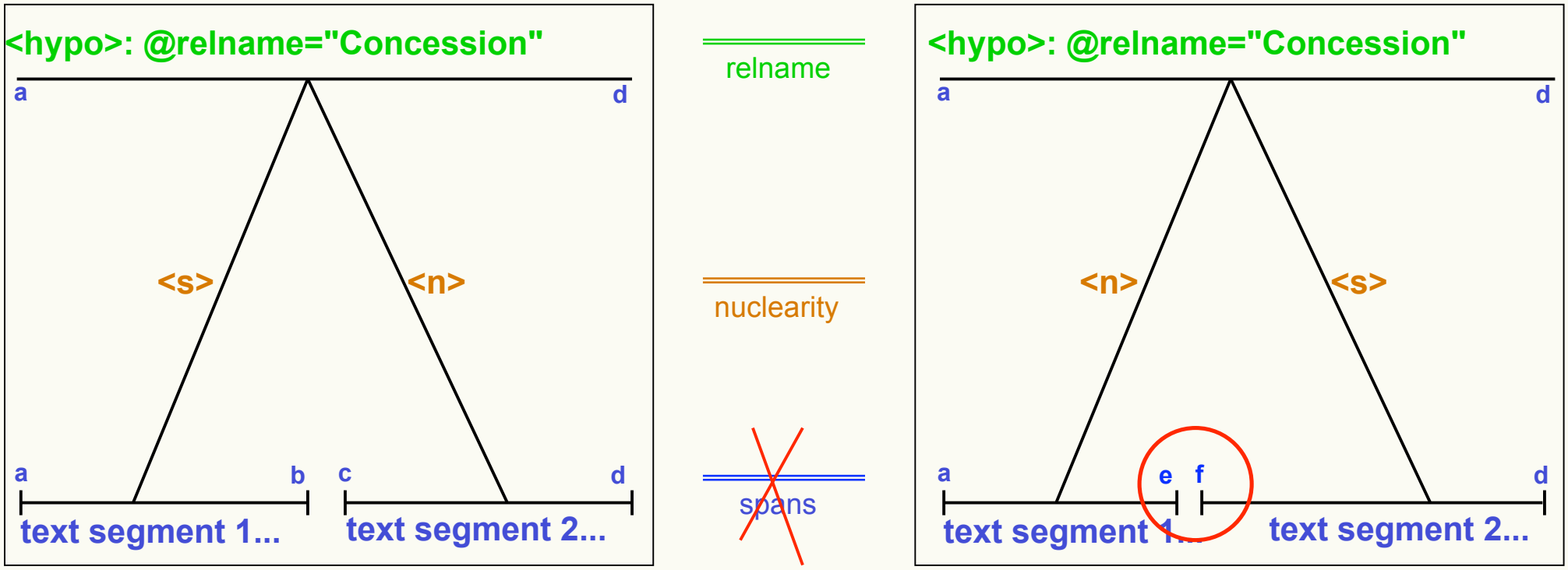
- Comparison with manual reference annotation of RST structure
- Precision and Recall, by matching RST SubTrees



NO MATCH

Experiment: Evaluation method

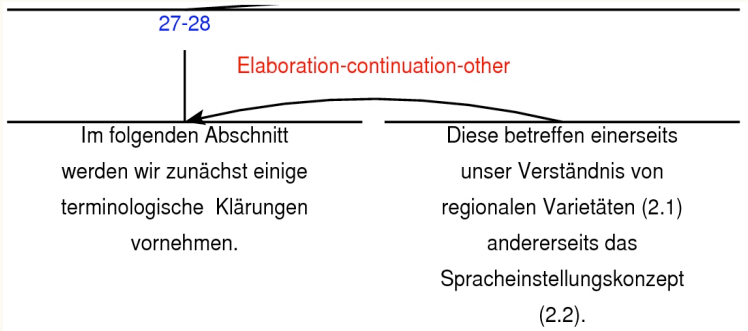
- Comparison with manual reference annotation of RST structure
- Precision and Recall, by matching RST SubTrees



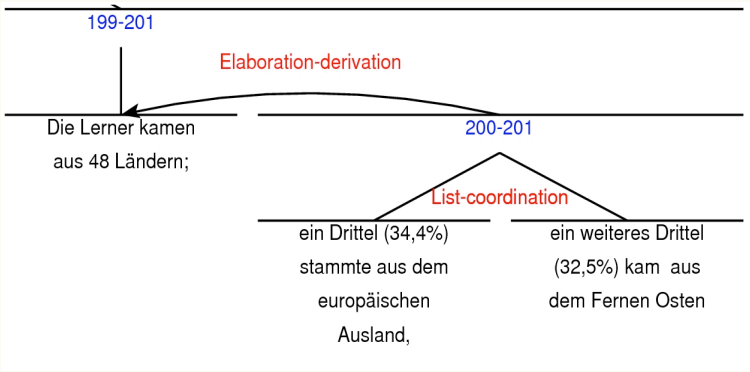
NO MATCH

Corpus Study: Results for B - Examples

```
<anaphoricInstance count="ling_deu_003-17" rtype='elaboration-continuation-oth'>
    <segment1>Im folgenden Abschnitt werden wir zunächst einige
    terminologische Klärungen vornehmen.</segment1>
    <segment2>Diese betreffen einerseits unser Verständnis von regionalen
    Varietäten, andererseits das Spracheinstellungskonzept.</segment2>
    <anaphora atype="cospec:ident">
      <antecedent>einige terminologische Klärungen</antecedent>
      <anaphor>Diese</anaphor>
    </anaphora>
</anaphoricInstance>
```




```
<anaphoricInstance count="ling_deu_003-109" rtype="elaboration-derivation">
    <segment1>Die Lerner kamen aus 48 Ländern;</segment1>
    <segment2>ein Drittel ( 34,4% ) stammte aus dem europäischen Ausland, ein
    weiteres Drittel (32,5% ) kam aus dem Fernen Osten (China, Japan, Korea,
    Taiwan) .</segment2>
    <anaphora atype="bridging:setMember">
      <antecedent>Die Lerner</antecedent>
      <anaphor>ein Drittel</anaphor>
    </anaphora>
    <anaphora atype="bridging:setMember">
      <antecedent>Die Lerner</antecedent>
      <anaphor>ein weiteres Drittel</anaphor>
    </anaphora>
  </anaphoricInstance>
```



Corpus Study: Results for A (sufficient condition)

	Instances	Elaboration	No-RST-Relation	other-than-elab
bridging:set-member	102	38	41	24
bridging:identSense	12	3	2	5
bridging:meronym	6	2	2	1
bridging:bridging	67	29	24	23
bridging:hasMember	37	5	23	9
bridging:poss	36	8	15	10
bridging:holonym	0	0	0	0
cospec:ident	403	99	180	119
cospec:paraphrase	107	30	41	36
cospec:addInfo	10	6	2	2
cospec:synonym	64	14	34	16
cospec:isA	47	14	25	8
cospec:properName	4	1	3	0
cospec:hyperonym	3	2	0	1
cospec:hyponym	2	0	0	2