coherence and cohesion in multimodal documents

TECHNISCHE UNIVERSITÄT DARMSTADT

intermodal relations in multimodal documents

Sabine Bartsch
Technische Universität Darmstadt
Institut für Sprach- und Literaturwissenschaft
URI: http://www.linglit.tu-darmstadt.de
E-Mail: {lastname}@linglit.tu-darmstadt.de

Wed, 18. July 2007, 14:45 – 15:30

overview

TECHNISCHE UNIVERSITÄT DARMSTADT

- multimodal communication in the target domain
- cohesion and coherence in multimodal documents
- analysis scheme
- conclusions / outlook

TECHNISCHE UNIVERSITÄT DARMSTADT

„[…] every linguistic act, spoken or written, takes place over more than one 'mode' or channel of communication“

„The study of ‘monomodal’ texts traditional in linguistics and other text-based investigations is thus revealed as an abstraction whose validity is questionable.“

(Bateman et al. 2002)

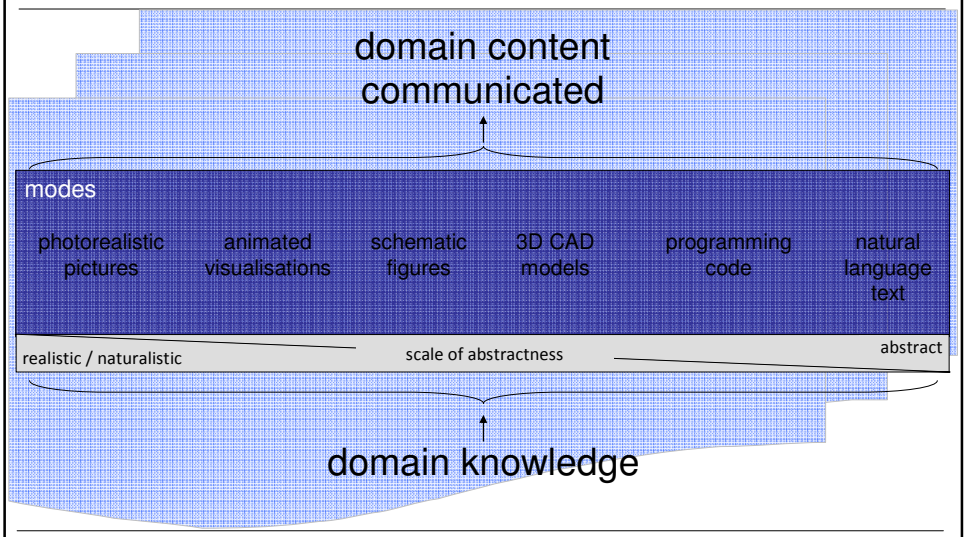
multimodal communication in the target domain

TECHNISCHE UNIVERSITÄT DARMSTADT

- target domain: data processing in construction, a subdomain of mechanical engineering
- self-characterisation: „the drawing is the language of the engineer“ (Reiner Anderl)
- semiotic resources:
  - natural language
  - visual representations such as pictures, schematic figures, sketches etc.
  - symbolic representations: formulae, diagrams, numerical data etc.
  - intermediary forms: tables, representations of data structures
  - virtual models such as CAD (Computer Aided Design) models

multimodal communication in the target domain

TECHNISCHE UNIVERSITÄT DARMSTADT



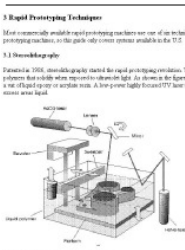
multimodal communication in the target domain

TECHNISCHE UNIVERSITÄT DARMSTADT

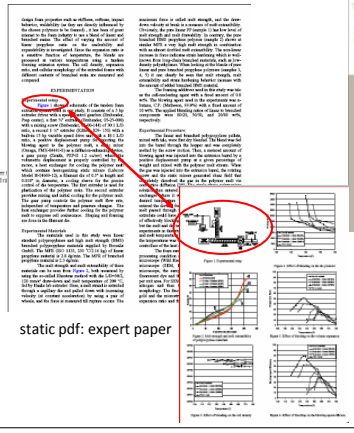
- conventional written texts:
  - printed documents (books, journals, papers, reports)
- computational communication artefacts:
  - electronic documents (webpages, online papers, pdfs)
  - visual and virtual 2D and 3D construction models
  - intermediary types (3D pdf)
  - user interfaces (databases, product data management systems, CAD construction software)
  - databases (PDM, CAD development repositories)
  - …

examples

TECHNISCHE UNIVERSITÄT DARMSTADT



html-page: teaching materials



static pdf: expert paper

interactive pdf: model of a turbine

corpus

TECHNISCHE UNIVERSITÄT DARMSTADT

- 140 multimodal documents
- genres:
  - research articles
  - lab reports
  - book chapters
  - instructional web pages
  - technology presentations
- tenor: expert-to-expert, expert-to-learner, learner-to-teacher, expert-to-user/customer
- mode: written-to-be-read

corpus

TECHNISCHE UNIVERSITÄT DARMSTADT

- 140 multimodal documents
- semiotic resources overall:
  - text (~1.7 mio. running words)
  - visual representations (~ 780)
  - tables (~ 300)
  - diagrams
  - virtual 2D and 3D CAD models (~ 20)
  - formulae , numerical data (unknown, yet)

corpus: overview

TECHNISCHE UNIVERSITÄT DARMSTADT

	expert-to-expert	expert-to-user	expert-to-learner	student-to-teacher
research articles	X			
lab reports	X			X
book chapters	X		X	
instructional web pages	X	X	X	
technology presentations	X	X		

corpus: semiotic resources

TECHNISCHE UNIVERSITÄT DARMSTADT

	text	visual representation	symbolic representation	intermediary types	virtual models
research articles	X	X	X	X	
lab reports	X	X	X	X	X
book chapters	X	X	X	X	
instructional web pages	X	X			
technology presentations	X	X		X	X

motivation

TECHNISCHE UNIVERSITÄT DARMSTADT

- how does the target discipline construe meaning multimodally?
- how is domain knowledge construed in different semiotic resources?
- how is multimodal semiosis instantiated in different types of documents in the target domain?

question: intersemiosis

TECHNISCHE UNIVERSITÄT DARMSTADT

- how do multiple modalities interact in multisemiosis?
- how is coherence achieved?

focus on intersemiosis:

- do they jointly contribute to the coherence of a multimodal document?
- **how** do they jointly contribute to the coherence of a multimodal document?

detour: some features of coherent (multimodal) text

TECHNISCHE UNIVERSITÄT DARMSTADT

- linearity
- rhetorical structure
- logical continuity
- multidimensional relations between sense units within semiotic resources
- surface cohesive devices
  (recurrence / repetition; substitution; ellipsis …)
- interactivity

detour: coherence in multimodal documents

TECHNISCHE UNIVERSITÄT DARMSTADT

- extension of the notions of coherence and cohesion to multimodal communication artefacts
- how do the concepts transfer?
- what are the necessary extensions / constraints?

methodological issue

TECHNISCHE UNIVERSITÄT DARMSTADT

- the lack of a rich and well established descriptive inventory – such as that available for natural language – for the description of other modalities
- models for the interaction between different modalities in multimodal documents
- the lack of well-organised data resources for the computer-aided study of multimodal documents

additional issues

TECHNISCHE UNIVERSITÄT DARMSTADT

- resources for the study of multisemiosis
  - established methodology
  - multimodal corpora
  - tools for annotation
  - annotation standards
  - query

research context

TECHNISCHE UNIVERSITÄT DARMSTADT

- genre and multimodality (GeM; Bateman et al. 2001)
- grammar of visual design (Kress & van Leeuwen 1996)
- visual transitivity frames (Baldry & Thibault 2006)
- visual message elements (Royce 2007*a*)
- transmedia design (Lemke 1998)
- intersemiosis in mathematical discourse (O'Halloran 1999a,b; 2004a)

- method: corpus and computational linguistics

annotation: status quo

TECHNISCHE UNIVERSITÄT DARMSTADT

- ideational features in language
  - PoS-tagging
  - syntactic parsing
  - chunking
  - lexical analysis
  - transitivity: ideational features (Systemic Coder)

  } automatic annotation

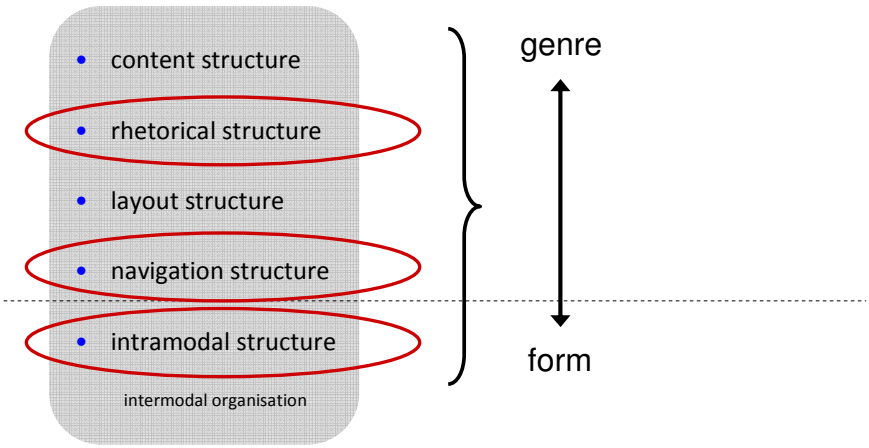
- textual features in language
  - lexical cohesion
  - Rhetorical structure (RSTTool)

  } semi-automatic + manual annotation

⇨ some of these annotations are extensible to other semiotic resources: e.g. transitivity, lexical cohesion, RST

levels of multimodal analysis

TECHNISCHE UNIVERSITÄT DARMSTADT



adapted from the GeM model; Bateman et al. 2001

linguistic text analysis

TECHNISCHE UNIVERSITÄT DARMSTADT

situational parameters	field tenor mode		
lexis	terminology collocations		
grammar	transitivity	process types	material, relational, mental, existential, behavioural, verbal
		circumstantial elements	time, place, manner
textual structure	cohesion thematic structure rhetorical structure generic structure		logico-semantic relations process type distribution

multimodal analysis of science texts

TECHNISCHE UNIVERSITÄT DARMSTADT

	verbal	intersemiosis	visual
ideational	terminology, collocations processes, participants, (attributes), circumstances	inter-modal phoric links	visual transitivity: objects, actions actions objects, (attributes)
interpersonal	modality, modulation	interactional features: intermodal references, hyperlinks, sequencing, menues, buttons, clickable areas	color, shading, etc.; interactivity features
textual	lexical cohesion, grammatical cohesion, conjunctive relations	inter-modal recurrence, repetition, substitution, ellipsis, rhetorical structure	visual cohesion, (explicit) sequencing,

metafunctional analysis: intersemiosis

TECHNISCHE UNIVERSITÄT DARMSTADT

- ideational metafunction
  - experiential
    - terminology, collocations
    - transitivity (process types, participant configurations)
  - logico-semantic relations
- interpersonal metafunction
  - interactional features / interactive semiosis
    - navigation structure
    - „reading“ path, traversals through the document
    - user interaction with the document
  - color, shading etc.
- textual
  - cohesion
  - rhetorical structure

cohesion in mm documents

TECHNISCHE UNIVERSITÄT DARMSTADT

- verbal cohesion
  - repetition (most frequent type, ≅ 40 %)
  - paraphrase
  - substitution
  - ellipsis
  - conjunction
  - collocation
- visual cohesion
  - repetition
  - visual paraphrases
  - visual collocation
- inter-modal cohesion

inter-modal cohesion

TECHNISCHE UNIVERSITÄT DARMSTADT

- simple repetition
- complex lexical repetition

- simple mutual paraphrase
- simple partial paraphrase
- other complex paraphrase
- [antonymous complex paraphrase]

- conjunction

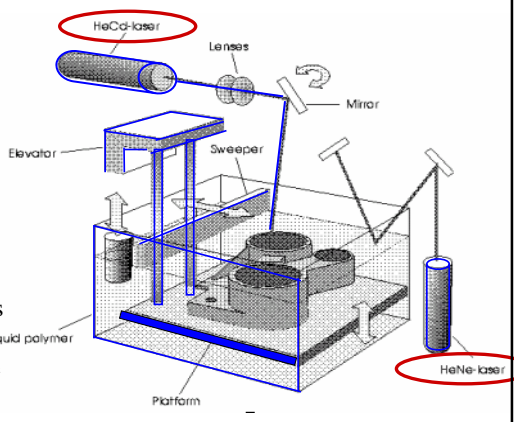
Hoey 1991

inter-modal cohesive chains and clusters: objects in different modes

TECHNISCHE UNIVERSITÄT DARMSTADT

The technique [stereolithography] builds three-dimensional models from liquid photosensitive polymers that solidify when exposed to ultraviolet light. As shown in the figure below, the model is built upon a platform situated just below the surface in a vat of liquid epoxy or acrylate resin.
A low-power highly focused UV laser traces out the first layer, solidifying the model's cross section while leaving excess areas liquid. Next, an elevator incrementally lowers the platform into the liquid polymer. A sweeper re-coats the solidified layer with liquid, and the laser traces the second layer atop the first.

repetition

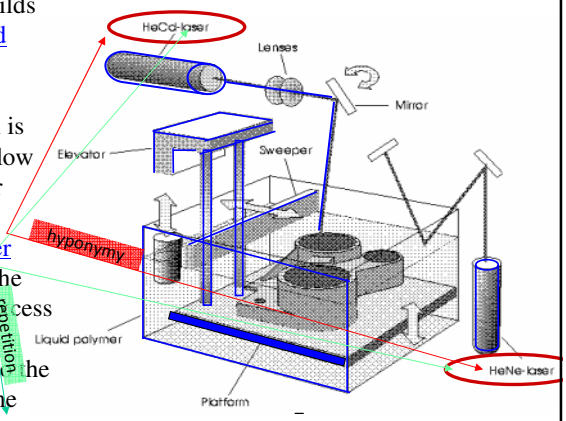


text more general than image

inter-modal cohesive chains and clusters: objects in different modes

TECHNISCHE UNIVERSITÄT DARMSTADT

The technique [stereolithography] builds three-dimensional models from liquid photosensitive polymers that solidify when exposed to ultraviolet light. As shown in the figure below, the model is built upon a platform situated just below the surface in a vat of liquid epoxy or acrylate resin.
A low-power highly focused UV laser traces out the first layer, solidifying the model's cross section while leaving excess areas liquid. Next, an elevator incrementally lowers the platform into the liquid polymer. A sweeper re-coats the solidified layer with liquid, and the laser traces the second layer atop the first.



text more general than image

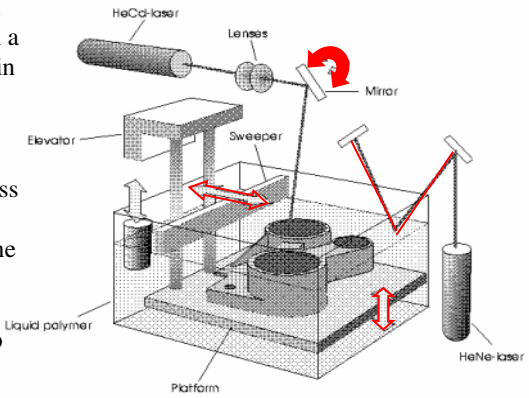
inter-modal cohesive chains and clusters: processes in different modes

TECHNISCHE UNIVERSITÄT DARMSTADT

The technique [**stereolithography**] builds three-dimensional models from liquid photosensitive polymers that solidify when exposed to ultraviolet light. As shown in the figure below, the model is built upon a platform sit[illegible] the surface in a vat of liquid epoxy or acrylate resin.
A low-power highly focused UV laser traces out the first layer, solidifying the model's cross section while leaving excess areas liquid. Next, an elevator incrementally lowers the platform into the liquid polymer. A sweeper re-coats the solidified layer with liquid, and the laser traces the second layer atop the first.

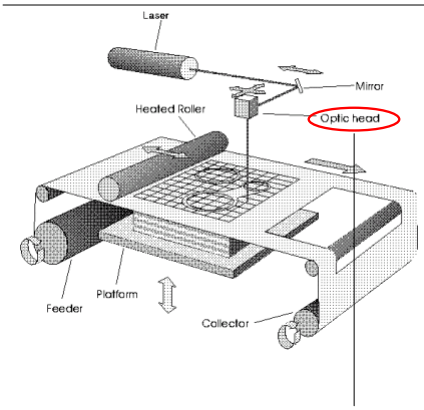
complex repetition

image more general and unspecific than text



inter-modal ellipsis

TECHNISCHE UNIVERSITÄT DARMSTADT



ellipsis in text

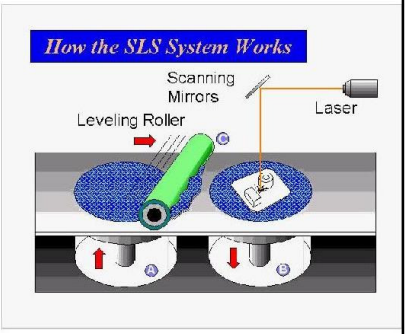
As shown in the figure below, a feeder/collector mechanism advances the sheet over the build platform, where a base has been constructed from paper and double-sided foam tape. Next, a heated roller applies pressure to bond the paper to the base.
A focused laser cuts the outline of the first layer into the paper and then cross-hatches the excess area (the negative space in the prototype). Cross-hatching breaks up the extra material, making it easier to remove during post-processing. During the build, the excess material provides excellent support for overhangs and thin-walled sections. After the first layer is cut, the platform lowers out of the way and fresh material is advanced.

inter-modal cohesion: conjunction in text and figure

TECHNISCHE UNIVERSITÄT DARMSTADT

[…] solid ground curing (SGC) is somewhat similar to stereolithography (SLA) in that both use ultraviolet light to selectively harden photosensitive polymers. Unlike SLA, SGC cures an entire layer at a time. Figure 5 depicts solid ground curing, which is also known as the solider process. **First**, photosensitive resin is sprayed on the build platform. **Next**, the machine develops a photomask (like a stencil) of the layer to be built. This photomask is printed on a glass plate above the build platform using an electrostatic process similar to that found in photocopiers. The mask is **then** exposed to UV light, which only passes through the transparent portions of the mask to selectively harden the shape of the current layer.

non-explicit parallel conjunction in text and image



inter-modal references: explicit references

TECHNISCHE UNIVERSITÄT DARMSTADT

- intermodal phoric links
  as shown in, as exemplified in, ...
  → used rather unspecifically, no good indicators of image function

… as depicted in Figure x.x:

Single-Point Cutting-Tool Geometry. Figure 13.2.3 depicts the location of various angles of interest on a single-point cutting tool.

TECHNISCHE UNIVERSITÄT DARMSTADT

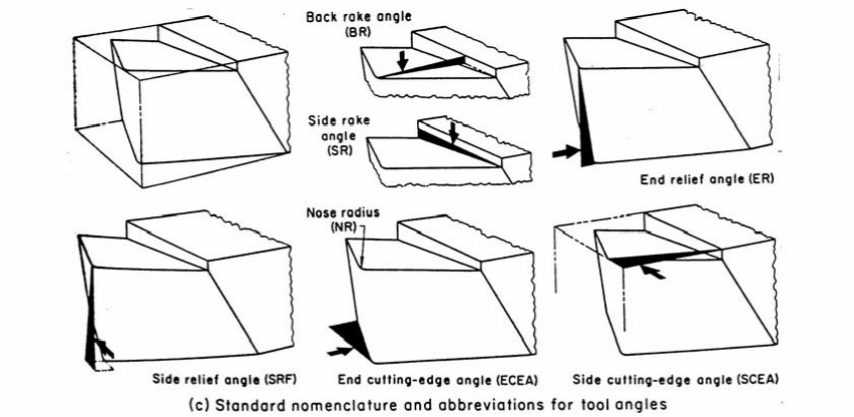


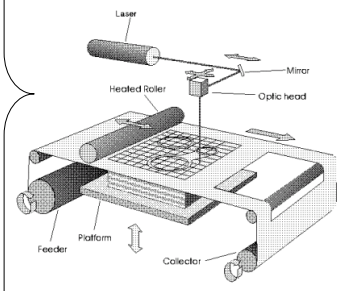
FIGURE 13.2.3 Standard nomenclature for single-point cutting-tool angles. (From *ASM Handbook, Machining*, vol. 16, 9th ed., ASM International, Metals Park, OH, 1989, 141. With permission.)

text as complex paraphrase of image

TECHNISCHE UNIVERSITÄT DARMSTADT

teaching materials / teacher to student:
[ctd. from annotation slide] As shown in the figure below, a feeder/collector mechanism advances the sheet over the build platform, where a base has been constructed from paper and doubles-sided foam tape.
Next, a heated roller applies pressure to bond the paper to the base.
A focused laser cuts the outline of the first layer into the paper and then cross-hatches the excess area (the negative space in the prototype).
Cross-hatching breaks up the extra material, making it easier to remove during post-processing.
During the build, the excess material provides excellent support for overhangs and thin-walled sections.
After the first layer is cut, the platform lowers out of the way and fresh material is advanced.
The platform rises to slightly below the previous height, the roller bonds the second layer to the first, and the laser cuts the second layer.

Scoping:
reference to larger and distributed sections, document parts



TECHNISCHE UNIVERSITÄT DARMSTADT

scientific paper / expert communication:

- Laser Sintering (SLS), and their use in the HIP process to fabricate tool steel mold cavities, Figure 1.
- The interaction of software programs necessary for Curved LOM is illustrated in Figure 2 (see [1] for an expanded description). The algorithms are incorporated in two separate packages.
- In fact, the cured shell fit almost perfectly on a freshly-built mandrel, as illustrated in Figure 5.

⇨ Figures often at the end of the text.

inter-modal relations: status of modes

TECHNISCHE UNIVERSITÄT DARMSTADT

- relative **status** of modes
  - text supporting image (*anchorage*),
  - image supporting text (*illustration*),
  - text and image equal / complementary (*relay*)

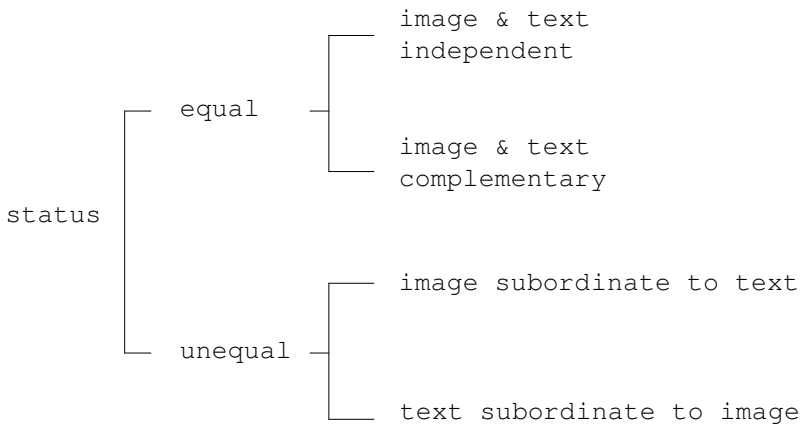
(Barthes 1977)

- equality | unequality

(cf.also Martinec, Salway 2005)

inter-modal relations: status of modes

TECHNISCHE UNIVERSITÄT DARMSTADT



inter-modal relations: logico-semantic relations

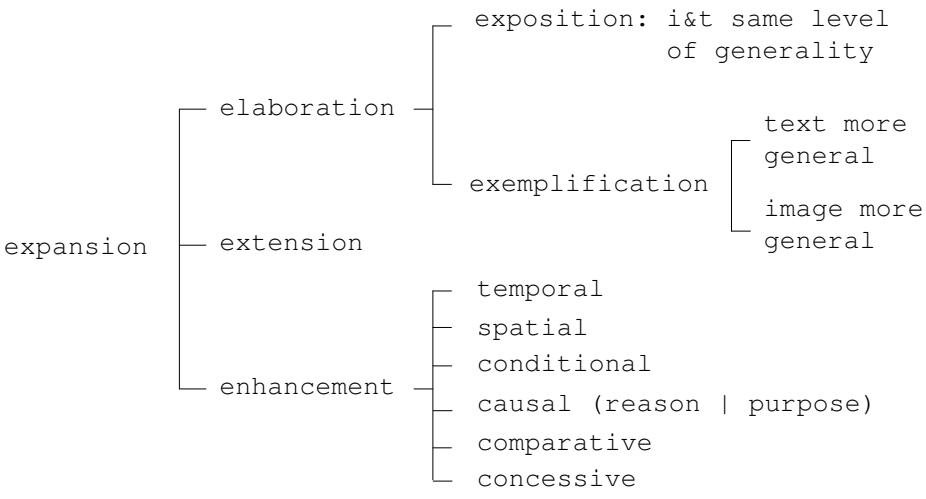
TECHNISCHE UNIVERSITÄT DARMSTADT

- logico-semantics:
  - expansion:
    - elaboration (elaboration on the meaning of another unit by a more detailed description of it)
    - extension (extension of the meaning of another unit / mode by adding further, related information)
    - enhancement (of a mode by qualifying it in terms of time, place, cause, and other circumstantial meanings)
  - projection

⇨ possibility of extension of these relations to further discourse units and other modes such as images, tables, formulae, 3D models etc.

inter-modal relations: logico-semantic relations

TECHNISCHE UNIVERSITÄT DARMSTADT



corpus annoation

TECHNISCHE UNIVERSITÄT DARMSTADT

```
<unit id="u-01.106">3.2 Laminated Object
Manufacturing</unit>
<unit id="u-01.109"><xlink:href="//gemBase/unit[@id='u-
01.118']" role="text-to-image" from=„u-01.109“ to=„u-01.118“
status=„unequal-image-subord“ logsem=„enhancement-temporal“
cohesion=„complex-paraphrase“/>As shown in the figure below,
a feeder/collector mechanism advances the sheet over the
build platform, where a base has been constructed from paper
and double-sided foam tape. [see next slide] </link></unit>
<unit id="u-01.118" alt="Bilder/Figure2“ objects=„laser
mirror platform collector …“ processes=„vertical_move
horizontal_move“/>
<unit id="u-01.119">
   <xlink:extended role="caption" title="figure caption">
   <xlink:locator href="//gemBase/unit[@id='u-
01.118']"/>Figure 2: Schematic diagram of laminated object
manufacturing.</xlink:extended></unit>
```


visual semiotic resource annotation

TECHNISCHE UNIVERSITÄT DARMSTADT

Objects	PARTICIPANT: ACTOR
Processes	CHANGE-STATE, ACTION MOVEMENT
Directionality	TRAJECTORY: vector, directionality

```
<unit id="u-01.119">
    <object function type=„participant-actor“>laser
       </s:object>
    <process function type=„change-state">
       </s:process>
    <trajectory type=„uni-directional“</s:trajectory>
</unit>
```

metadata for schematic figure: experiential description

TECHNISCHE UNIVERSITÄT DARMSTADT

- PARTICIPANT: AGENT: laser beam
  PROCESS: CHANGE-STATE fuse
  PARTICIPANT: powdered materials
  GOAL: into a solid object
- PARTICIPANT: AGENT: laser
  PROCESS: ACTION MOVEMENT: traces
- PARTICIPANT: the pattern,
  GOAL: sintering it together
- TRAJECTORY: vector, directionality

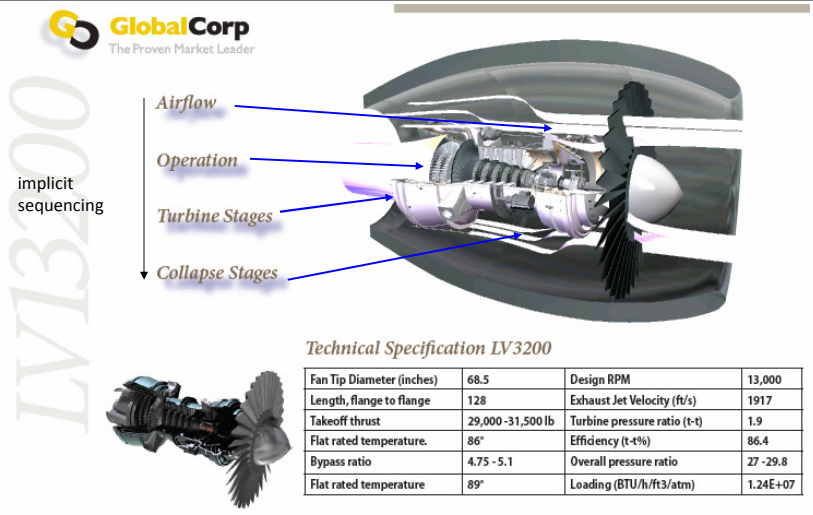
Example: interactivity in 3d-pdf

TECHNISCHE UNIVERSITÄT DARMSTADT

- cohesive devices:
  - repetition of terms
  - paraphrasing
  - text – construction data
  - interaction activated cohesion
- new forms of interactivity
  - user interaction with the document
  - manipulation of an image or model
  - integration with an ontology of the parts

Example: interactivity and cohesion in 3d-pdf

TECHNISCHE UNIVERSITÄT DARMSTADT



Technical Specification LV3200

Fan Tip Diameter (inches)	68.5	Design RPM	13,000
Length, flange to flange	128	Exhaust Jet Velocity (ft/s)	1917
Takeoff thrust	29,000 -31,500 lb	Turbine pressure ratio (t-t)	1.9
Flat rated temperature.	86°	Efficiency (t-t%)	86.4
Bypass ratio	4.75 - 5.1	Overall pressure ratio	27 -29.8
Flat rated temperature	89°	Loading (BTU/h/ft3/atm)	1.24E+07

Example: interactivity in 3d-pdf

TECHNISCHE UNIVERSITÄT DARMSTADT

- linearity: multiple paths through the document
- animated document
- interactivity: recipient can manipulate and change the document
- integration of construction data in the document

- interpersonal metafunction:
  - manipulation and adjustment of lighting, perspective, detail, colour, transparency etc.

outlook

TECHNISCHE UNIVERSITÄT DARMSTADT

- extension of reusable resources
  - multimodal corpora
  - annotation
  - shared tools
- annotation
- use automatic tools where feasible
- manual annotation
- interchangeable formats (e.g. XML)

- making use of virtual interdisciplinarity

possible outcomes or applications

TECHNISCHE UNIVERSITÄT DARMSTADT

- what kind of reception skills do recipients of such documents require in order to decode and benefit from multiple modes?

- what kind of (text) production skills do producers of such documents require in order to produce useful multimodal documents?

references

André, E., T. Rist. 1994. 'Referring to World Objects with Text and Pictures.' Coling 1994.
Bateman, J., K.J. Rondhuis. 1997. 'Coherence relations': Towards a general specification. Discourse Processes, 1997.
Bateman, J., J. Delin, R. Henschel. 2002. Multimodality and empiricism: methodological issues in the study of multimodal meaning-making. GeM Report 2002/1.
Barthes, R. 1977. *Image – Text – Music*. London: Routledge.
Halliday, MAK, R. Hasan. 1976. *Cohesion in English*. London, New York: Longman.
Hoey, M. 1991. *Patterns of Lexis in Text*. Oxford: OUP.
Martinec, R., A. Salway. 2005. A system for image-text relations in old (and new) media. Visual Communication Vol 4(3): 339–374.
Mann, W.C., S.A. Thompson. 1987. Rhetorical structure theory: a theory of text organization. Tech. rep. RS-87-190, USC/ISI, Los Angeles, CA.
Martin, J.R. 1992. *English text: systems and structure*. Amsterdam: Benjamins.
O'Donnell, Mick. 2002. *UAM Corpus Tool.* URL: http://www.wagsoft.com/CorpusTool/
O'Halloran, K. 1999. 'Towards a systemic functional analysis of multisemiotic mathematics texts', Semiotica 124–1/2 (1999), 1–29.
O'Halloran, K. 2004a. *Mathematical Discourse. Language, Symbolism and Visual Images.* London, New York: Continuum.
O'Halloran, K. 2004b. *Multimodal Discourse Analysis. Systemic Functional Perspectives.* London, New York: Continuum.

coherence in multimodal text

- sequential coherence
  - adjacency (of turns / sentences relating to one another)
  - Relevance (to the discourse topic)
- interactional coherence
  - interactivity