Multimodal Text Interpretation: Modelling the Whole Process

Martin Siefkes

BreMM15: Second Bremen Conference on Multimodality
21–22 Sept 2015, Bremen, Germany

Multimodal Text Interpretation
Modelling the Whole Process

Martin Siefkes
Chemnitz University of Technology

www.siefkes.de

Licensed under the Creative Commons Attribution 4.0 License (CC BY 4.0):
https://creativecommons.org/licenses/by/4.0
Overview

- Different Perspectives
- Modelling the Whole Process
- Intermodal Patterns
Different Perspectives on Multimodality
Different Approaches – Different Perspectives

- Semiotic modes (semiotics; e.g. van Leeuwen 2011) versus
- Perceptual modes (psychology, e.g. Calvert et al. 2004)
- Integration and compositionality of meaning (e.g. Kress/van Leeuwen 2001, Stöckl 2004)
- Reading paths (e.g. Bucher 2011, Hiippala 2012)
- Genre and discursive context (Bateman 2008, O'Halloran/Smith 2011, Norris/Maier 2014)
- Relations between modes (e.g. Wildfeuer 2011, Elleström 2011, Fricke 2012, Siefkes 2015)
- Rhetorical relations and narrative structure (e.g. Bateman/Wildfeuer 2011, Wildfeuer 2014)
- Multimodal metaphors (e.g. Forceville/Urios-Aparisi 2009)
- Social embedding of discourses (e.g. Machin/van Leeuwen 2007, Kress 2009)
Is an Integrative Perspective Possible?

- Different approaches are complementary in their perspectives
- Can we model the different aspects as steps in a process model?
- The perception and interpretation of multimodal texts can then be understood as a complex, multi-stage sign process
- The order of the steps is NOT temporal, but analytical
- On this basis, results from different approaches can be integrated
Modelling the Whole Process
A Multimodal Text Interpretation Model

- Can we construct a process model that considers all these aspects?
- After each stage of the analysis, a text / discourse representation is enriched with the results.
- Detailed analysis schemata exist, but mostly result in static, not process-oriented representations.
- MuTIM aims to provide a framework for modelling the various steps, and provide representations for the outcome of all analysis steps (including representations before semantic analysis).
- Long-term goal: integration of results of computational approaches, that exist already for various modes: language (NLP), image analysis, facial recognition (incl. facial expression), gesture recognition (all by trainable statistical models).
- Risk of computational approaches (i.e. information extraction): return to monomodal approaches.
- What about intermodality?
Steps of Analysis (1)

1. **Delimitation** of the text from other texts, signs, and non-semiotic context

2. **General categorization**
   a) Identification of the multimodal text type (film, website, brochure, …)
   b) Identification of the genre (action film, political website, tourist brochure, …)
   c) Identification of the semiotic modes present (language, image, gesture, …)
   d) Which of these modes are embedded in others? (gesture in film or comics; music in film; …)

3. **Syntactic parsing**: separation of each mode into units that determine or distinguish meaning

4. **Analysis (parsing)** of intermodal expression relations
   ➢ spatial & temporal relations, overlap, colours, brightness, …
Steps of Analysis (2)

5. Determination and empirical verification of reading path(s)

6. **Semantic analysis** (interpretation) of the parsed expressions
   - Determining reference (to objects and situations), assigning propositions

7. **Disambiguation** of polysemous expressions
   - through context, background knowledge, semantic frames

8. **Intermodal reference unification** (e.g. to persons, objects, places)

9. Determination of **intermodal content relations**
Steps of Analysis (3)

10. Analysis of Intermodal Patterns
   - through context, background knowledge, semantic frames

11. Wider inferences (based on background knowledge)

12. Large-scale textual structures (e.g. thematic and narrative organisation, suspense structure, etc.)

13. Situation and context

14. Stylistic features that relate this text to other texts (by the same or other authors)

15. Aesthetic description and evaluation of the text
Intermodal Patterns
Intermodality – The Missing Link(s)

- **Intermodality** research thinks about the contributions of modes to the overall textual ensemble
  - concentrates on the configurations across modes
- Intermodality can be described as looking for patterns across modes
- Patterns come in two types
  - configurations / collocations: statistically more frequent
  - holistic patterns: with additional meaning(s) or perceptual qualities
Three Types of Patterns

- **Process** patterns
- **Functional** patterns
- **Semiotic** patterns
Process Patterns

• Patterns in the process of interpretation (described in MuTIM)

• It is possible to perceive multimodal texts in different ways:
  
  – for an illustrated text: to read the text and look at images afterwards; to look at each image when the text reaches it; to look at an image only when the text explicitly refers to it (“cf. image x”); to quickly glance over the images and look closer when the text starts to talk about the topic; etc. The number of patterns multiplies when we include captions.

  – for exhibited artworks: description then artwork; other way around; glance at artist's name, artwork, then description; etc.

• Patterns of complete and incomplete perception
Functional Patterns

- Describe different ways semiotic functions can be distributed over modes
- General textual functions (SFL: ideational, interpersonal, textual)
- Further separation for specific genres, e.g.
  
  a) **Interpersonal function** of social media profiles: positive presentation of a person, gaining followers, notifications about people in the network, etc.
  
  b) **Textual function** of social media profiles: demonstrating which parts belong to the profile (e.g. posts by, and general information about, a specific profile); what belongs to the timeline; how you can change to another profile; etc.
  
  c) **Ideational function**: imparting information, announcing events, etc.
Semiotic Patterns

1) Which modes are present?
   - modes can be necessary or optional for a pattern to be present

2) Which conditions apply for each separate mode?
   a) Expression conditions
   b) Content conditions

3) Which conditions apply for the relations between modes?
   a) Expression relations
   b) Content relations

4) What holistic qualities does the pattern have?
   a) Holistic qualities of expression
   b) Holistic qualities of content
Example 1: Posts & Tweets

• Post / tweets form the content part of a social network
• Similarities: which modes are involved and what they do
• Together with other patterns, this patterns defines the genre “profile on a social network website”
Martin Siefkes
@stilomet

Research assistant at @tuchemnitz. Researching #semiotics, #multimodality, linguistics, gesture in multimodal interaction, and digital humanities.

Chemnitz, Sachsen
siefkes.de

Tweets
1.165
FOLGE ICH
177
FOLLOWER
109
FAVORITEN
190

Tweets
Cathy Davidson @CathyNDavidson · 1 Std.
Exciting! Call for papers (due Nov 1), HASTAC 2016: Impact! from DH to digital activism, impacts in the world. hastac.org/opportunities/

CfP: @europeanisa Workshop on 'Living the "new normal": Post-crisis politics of money, debt and time' in Tübingen bit.ly/1W4ENRz
Killing you
Like a cat in a box
Or not

gtwilson  Tom Wilson
http://www.shortbusacademy.com

@gtwilson

246

@larand We got off the boat in the 1850s from Ireland. @akulbe @tomcat @rabryst
Hope to see some of you in Bremen next week! #Bremm15 #DNC1

Then, there's this year's DiscourseNet Congress (24 to 26 Sept).
<table>
<thead>
<tr>
<th>Mode</th>
<th>Expression</th>
<th>Content &amp; Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>image</td>
<td>type: image or symbol height: 2 or 3 font rows position: left</td>
<td>implied connection with the account (direct, e.g. face of user, or indirect, e.g. topic of interest)</td>
</tr>
<tr>
<td>text</td>
<td>element_1: name, style: bold [element 2: handle, style: no spaces] element 3: time stamp element 4: body of post</td>
<td>element_1: real name of account owner OR alias element_2: alias for account owner element 3: time of writing relative to page load OR absolute time (for older entries) element 4:</td>
</tr>
<tr>
<td>symbols</td>
<td>element_1: answer symbol [element 2: share / retweet symbol] element 3: like symbol</td>
<td>element_1: answering a post or tweet element_2: resend from own account element_3: mark preference / bookmark</td>
</tr>
<tr>
<td>fields</td>
<td>[field for text entry] position: below post OR below profile</td>
<td>entering text to answer a post / tweet</td>
</tr>
<tr>
<td>graphic_elements</td>
<td>element_1: white background element_2: grey background</td>
<td>element_1: background for posts element_2: separates posts</td>
</tr>
</tbody>
</table>
Example 2: Header Images

- Header images are used in profiles of social networks as well as on other websites (the example shows the website of a university library).
- Always placed above the general content, sometimes with overlaid text (e.g. profile or site name).
- This example is more widely used than example 1, across various more specific genres (such as the genre “profile on a social network website”).
**Karma:** 125.13

- **Wie erhält man mehr Karma?**
- **Profil-Ansichten:** 5628
- **Eingeladene Freunde:** 0
- **Plurks:** 8727
- **Plurk-Antworten:** 191031

**Mitglied seit:** 2011-3-14
**Letzter Login:** 2015-9-14

---

**STATISTIKEN**

**FREUNDE**

- mutedtempest als Freund hinzufügen
- Alle Freunde (125)

**FANS**

- Plurks von mutedtem... folgen
- Alle Fans (11)

---

**mutedtempest @mutedtempest**

32 Jahre alt, weiblich
Uppsala, Sweden

Privaten Plurk senden

A BUTT WITHIN A BUTT WITHIN A BUTT

Kyouraku Shunsui at Eachdraidh and Town of Salem. I'm a grad student and complete bum. Hi.

Beziehung:
Single

---

**mutedtempest [boy venting] Okay, im trying to**

**mutedtempest [rt/mute! i'm honestly not as u**

**mutedtempest [Places in Which to Settle in A**

**mutedtempest [interstellar] and then i did a t!**
<table>
<thead>
<tr>
<th>Mode</th>
<th>Expression</th>
<th>Content &amp; Function</th>
</tr>
</thead>
<tbody>
<tr>
<td>image</td>
<td>element_1: type: image</td>
<td>implied connection with the account; general or typical semantic content; visualisation that is relevant to the institution described in the text</td>
</tr>
<tr>
<td></td>
<td>height: max. half of visible page</td>
<td></td>
</tr>
<tr>
<td></td>
<td>position top of page</td>
<td></td>
</tr>
<tr>
<td>text</td>
<td>element_1: title of page, style: larger than average</td>
<td>element_1: name of institution/user OR topic of site</td>
</tr>
<tr>
<td></td>
<td>element_2: body of page</td>
<td>element_2: content of the site</td>
</tr>
<tr>
<td>[menu (interactive)]</td>
<td>element_1: position: top of page</td>
<td>element_1: opening subsites or tabs</td>
</tr>
<tr>
<td></td>
<td>expression_relation (image: element_1, menu: element_1): overlap OR next to</td>
<td></td>
</tr>
</tbody>
</table>
Conclusions

1. Multimodal text interpretation is a multi-stage sign process
   - Signs in all present modes are parsed, decoded or interpreted, relations between them are analysed, and integrated
   - MuTIM: analysis steps that need to be taken for a full model of the process

2. Intermodality: additional aspects of multimodal (in comparison to monomodal) sign processes
   - Pattern-based approach: consists of defining a set of patterns for each text type (e.g. film) and genre
Selected Bibliography