Corpus Analyses of Multimodal Narrative: The Example of Graphic Narrative

Alexander Dunst, Rita Hartel, Sven Hohenstein, Jochen Laubrock
Aims of Research Group

- XML dialect for all aspects of graphic narrative relevant to research
- Annotation Tool (Open Access Software)
- Corpus of English-language graphic narrative (+ reference corpora)
- Eye movement corpus (+ analysis tools, R-package)
- Quantitative cultural history of English-language graphic narrative
- Development of empirically founded narratological concepts
- Psychological experiments on reception of graphic narrative
Part I: GNML Editor

Tools for (semi-)automatic annotation
a) Graphic Narrative Markup Language (GNML)

b) Graphic GNML-Editor
Editor offers:

- Automatic Panel Detection
- Semi-Automatic Balloon & Caption Detection
- Auto-Complete Function & Spell Check
Editor-Download:

http://blogs.uni-paderborn.de/graphic-literature/editor/
http://blogs.uni-paderborn.de/graphic-literature/faq/
Part II: Graphic Narrative
Quantitative Analyses and Visualizations
c) Graphic Narrative Corpus

Time period: 1970 to 2015

Corpus of English-speaking graphic narrative (graphic novels, memoirs, and non-fiction)

Reference Corpora: historical comics + German, Franco-Belgian and Japanese comic books

Sources: international comic awards, MLA and JSTOR citations, media coverage, Amazon bestseller lists

220 titles: digitalised by end of 2016, currently 13 titles annotated
Internal Differentiation: Titles, Page Numbers, and Book Covers
Internal Differentiation: Titles, Page Numbers, and Book Covers
Space vs. Text in 10 Graphic Narratives

Relative Visual Importance of Background, Characters, Balloons and Captions
Comparative Visual Importance of Main Characters in 10 Graphic Narratives

1. Cleveland
2. City of Glass
3. Our Cancer Year
4. Fun Home
5. Maus
6. Epileptic
7. Gemma Bovery
8. V for Vendetta
9. A Contract With God
10. Watchmen
Social Networks I: Interaction Types

1. A Contract with God
2. Cleveland
3. City of Glass
4. Epileptic
5. Fun Home
6. Our Cancer Year
7. V for Vendetta

edge colors:
- CoPresenceSpace
- Dialogue
- EyeContact
- Monologue
- PhysicalContact

- Harveys Pekars Fan
- Girl at Train Station

Fannie (tenant of 55 Dropsie Avenue)
Joyce Brabner (Harveys 3rd Wife)
Willie (Fannies son)
Maralyn Minks
John T. Zubal (Bookshop Owner)
Sam (Fannies husband)
Deaf Man
Peter Stillman Jr.
Mr. Pinkus (owner of Pinkus Furs)
Virginia Stillman
Goldie (secretary at Pinkus Furs)
Benny (cutter at Mr. Cohen's shop)
Peter Stillman Sr.
Rabbi at the synagogue
Mrs. Farfell (tenant of 55 Dropsie Avenue)
Daniel Quinn's Wife
Frimme Hersh
Lee (Harveys Friend)
Daniel Quinn's Wife
Receptionist at Hotel Harmony
Harveys Pekars Mother
David's imaginary confidant
Rabbi at the synagogue
Eddie (The Street Singer)
Girl (Train Station/Apartment)
David's imaginary confidant
Social Networks II: Gender Topologies

1. A Contract with God  
2. Cleveland  
3. City of Glass  
4. Epileptic  
5. Fun Home  
6. Our Cancer Year  
7. V for Vendetta
Part III: Graphic Literature
Eyetracking Analysis of Multimodal Narrative
Studies

- Eye-tracking Corpus
- City of Glass
Corpus study goals

- Empirical description of reader behavior
- What aspects of the material capture the reader’s attention?

100 participants, eye-tracking, chapters from graphic narratives
Ob nun Absicht oder Zufall dahintersteckte, es war in jedem Fall eine Mordimbezille.

Da unten lag also mein Vater, sagte ich mir.

Jetzt steckte er aber wirklich in der Patsche.
Relevant objects
Hack... Hack... Uaaah... Noch... Knirsch... Knirsch... Suf Suf Schmatz... Knirsch... Donp... Umm... M... Mir wird schwindig... Waaaaah! Waaaaah!
EIN NAGELNEUER PONTIAC SIESTA

EIN ALL-INCLUSIVE URLAUB IN DISNEYLAND

ODER ZWEI... NEIN, ZEHNTAUSEND DOLLAR FÜR SIE!

HEY! ICH HAB WAS GEWONNEN!

HA HA, BLOSS 'N WITZ

DA GEWINNT DOCH EIN NICHT JENDEM WAS

ABER NACH DISNEYLAND WOLLTE ICH SCHON IMMER MAL...

DIE GANZEN FAHRSCHULEND ALLES, WAS MAN SO AUS DEM FERNSEHEN KENNT...
A picture says more than a thousand words: Text dominates processing time
A picture says more than a thousand words: Text dominates processing time
Effects of lexical variables

Similar effects as in reading of text
City of Glass: The Graphic Novel

Questionnaires

• Comics and text expertise (adapted from Cohn, 2014)
• Do you enjoy reading comics

• Comprehension
  • *literal*: How old is Quinn?
  • *inferential*: What does the author depict as the New Babylon?
Is there a specific comics expertise?

Yes! Independent types of expertise: no significant correlation of comics and text expertise

Strong correlation of comics (but not text) expertise and comics enjoyment

$(r = .71, p < .001)$

$(r = .21, p = .172)$
Does comics expertise aid in understanding the visual language?

Mean inferential comprehension as a function of text expertise (n.s.)

Mean inferential comprehension as a function of comics expertise
Does comics expertise aid in understanding the visual language?

Mean inferential comprehension as a function of text expertise (n.s.)

Mean inferential comprehension as a function of comics expertise

Yes! Comics expertise supports drawing inferences, not simple reproduction
Is comics expertise reflected in viewing behavior?

<table>
<thead>
<tr>
<th>Random Effects</th>
<th>Variance</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>0.039</td>
<td>0.198</td>
</tr>
<tr>
<td>Trial</td>
<td>0.027</td>
<td>0.166</td>
</tr>
<tr>
<td>Residual</td>
<td>0.189</td>
<td>0.435</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>Coefficient</th>
<th>SE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept (text and image)</td>
<td>7.439</td>
<td>0.040</td>
<td>183.91</td>
<td>&lt; .001***</td>
</tr>
<tr>
<td>Type: text only</td>
<td>-0.178</td>
<td>0.023</td>
<td>-7.89</td>
<td>&lt; .001***</td>
</tr>
<tr>
<td>Type: image only</td>
<td>-0.568</td>
<td>0.018</td>
<td>-31.48</td>
<td>&lt; .001***</td>
</tr>
<tr>
<td>Number of words/ panel</td>
<td>0.054</td>
<td>0.001</td>
<td>78.11</td>
<td>&lt; .001***</td>
</tr>
<tr>
<td>Comic fluency (standardized)</td>
<td>0.054</td>
<td>0.031</td>
<td>1.77</td>
<td>.084</td>
</tr>
<tr>
<td>Text fluency (standardized)</td>
<td>-0.046</td>
<td>0.031</td>
<td>-1.49</td>
<td>.144</td>
</tr>
<tr>
<td>Panel number</td>
<td>-0.016</td>
<td>0.002</td>
<td>-9.63</td>
<td>&lt; .001***</td>
</tr>
<tr>
<td>Panel number^2</td>
<td>0.004</td>
<td>0.001</td>
<td>6.38</td>
<td>&lt; .001***</td>
</tr>
<tr>
<td>Text only × comic fluency</td>
<td>-0.021</td>
<td>0.021</td>
<td>-1.01</td>
<td>.312</td>
</tr>
<tr>
<td>Image only × comic fluency</td>
<td>0.101</td>
<td>0.012</td>
<td>8.72</td>
<td>&lt; .001***</td>
</tr>
<tr>
<td>Text only × text fluency</td>
<td>-0.008</td>
<td>0.021</td>
<td>-0.37</td>
<td>.711</td>
</tr>
<tr>
<td>Image only × text fluency</td>
<td>-0.002</td>
<td>0.012</td>
<td>-0.15</td>
<td>.880</td>
</tr>
</tbody>
</table>

Note. *** p < .001, ** p < .01, * p < .05; AIC = 16697, BIC = 16810
Is comics expertise reflected in viewing behavior?

<table>
<thead>
<tr>
<th>Random Effects</th>
<th>Variance</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Subject</td>
<td>0.039</td>
<td>0.198</td>
</tr>
<tr>
<td>Trial</td>
<td>0.027</td>
<td>0.166</td>
</tr>
<tr>
<td>Residual</td>
<td>0.189</td>
<td>0.435</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Fixed Effects</th>
<th>Coefficient</th>
<th>SE</th>
<th>t</th>
<th>p</th>
</tr>
</thead>
<tbody>
<tr>
<td>Intercept (text and image)</td>
<td>7.439</td>
<td>0.040</td>
<td>183.91</td>
<td>&lt; .001 ***</td>
</tr>
<tr>
<td>Type: text only</td>
<td>-0.178</td>
<td>0.023</td>
<td>-7.89</td>
<td>&lt; .001 ***</td>
</tr>
<tr>
<td>Type: image only</td>
<td>-0.568</td>
<td>0.018</td>
<td>-31.48</td>
<td>&lt; .001 ***</td>
</tr>
<tr>
<td>Number of words/ panel</td>
<td>0.054</td>
<td>0.001</td>
<td>78.11</td>
<td>&lt; .001 ***</td>
</tr>
<tr>
<td>Comic fluency (standardized)</td>
<td>0.054</td>
<td>0.031</td>
<td>1.77</td>
<td>.084</td>
</tr>
<tr>
<td>Text fluency (standardized)</td>
<td>-0.046</td>
<td>0.031</td>
<td>-1.49</td>
<td>.144</td>
</tr>
<tr>
<td>Panel number</td>
<td>-0.016</td>
<td>0.002</td>
<td>-9.63</td>
<td>&lt; .001 ***</td>
</tr>
<tr>
<td>Panel number^2</td>
<td>0.004</td>
<td>0.001</td>
<td>6.38</td>
<td>&lt; .001 ***</td>
</tr>
<tr>
<td>Text only × comic fluency</td>
<td>-0.021</td>
<td>0.021</td>
<td>-1.01</td>
<td>.312</td>
</tr>
<tr>
<td>Image only × comic fluency</td>
<td>0.101</td>
<td>0.012</td>
<td>8.72</td>
<td>&lt; .001 ***</td>
</tr>
<tr>
<td>Text only × text fluency</td>
<td>-0.008</td>
<td>0.021</td>
<td>-0.37</td>
<td>.711</td>
</tr>
<tr>
<td>Image only × text fluency</td>
<td>-0.002</td>
<td>0.012</td>
<td>-0.15</td>
<td>.880</td>
</tr>
</tbody>
</table>

Yes! Comics experts selectively spend more time fixating the image (but not the text) content.
Heatmap of fixations of the lower comics expertise tercile (n = 15)
Heatmap of fixations of the upper comics expertise tercile (n = 15)
Conclusion and Outlook
Conclusion

• Description of under-researched medium using combination of stylometric, network, and cognitive analysis

• Integration of digital and cognitive methods to study production, circulation, and reception of popular cultural form

• Rich annotations that include visual features and reception data can help both the (digital) humanities and cognitive science
Outlook
Outlook

- Integration of OCR trained on comics fonts and automatic object recognition into editor software
Outlook

- Integration of OCR trained on comics fonts and automatic object recognition into editor software
Outlook

- Integration of OCR trained on comics fonts and automatic object recognition into editor software
- Use feature descriptors from computer vision and computer linguistics for richer annotations, object and character recognition

Classic (GIST, SIFT, HOG, HOOSC) or deep features (Krizhevsky et al., 2012) to describe image and code artistic style (Gatys et al., 2015; Elgammal & Saleh, 2015)
Outlook

• Integration of OCR trained on comics fonts and automatic object recognition into editor software

• Use feature descriptors from computer vision and computer linguistics for richer annotations, object and character recognition

• Use of machine learning for stylometric analyses and prediction of viewing behavior

Classic (GIST, SIFT, HOG, HOOSC) or deep features (Krizhevsky et al., 2012) to describe image and code artistic style (Gatys et al., 2015; Elgammal & Saleh, 2015)
Outlook

• Integration of OCR trained on comics fonts and automatic object recognition into editor software

• Use feature descriptors from computer vision and computer linguistics for richer annotations, object and character recognition

• Use of machine learning for stylometric analyses and prediction of viewing behavior

• Intermedial, multinational and historical analysis based on corpus of 200 full-length graphic narratives

Classic (GIST, SIFT, HOG, HOOSC) or deep features (Krizhevsky et al., 2012) to describe image and code artistic style (Gatys et al., 2015; Elgammal & Saleh, 2015)
Thank you for your attention!

dunst@mail.upb.de
laubrock@uni-potsdam.de
graphic-literature.upb.de