Zeitschrift für Semiotik

Herausgegeben von
Ellen Fricke und Martin Siefkes

Begründet von Roland Posner†

Innovative Methods in Multimodal Comics Research

Herausgegeben von / Edited by
Janina Wildfeuer and Stephan Packard

Applying multimodal semiotics to comics and graphic novels
Corpus annotation: manual and automatic methods
Intersemiotic translation in Lobačev's "Tsar Dušan's Wedding"
Unreliable iconicity: cartoonish qualities of pictorial narration
Explaining multimodal cohesion: a Peircean approach
The semiotics of discourse relations in comics

Vorschau

Zeitschrift für Semiotik

Verantwortlich für den Inhalt: Die Herausgeber

Die Zeitschrift für Semiotik ist im
ERIH PLUS-Index gelistet.

ERIH PLUS
EUROPEAN REFERENCE INDEX FOR THE
HUMANITIES AND SOCIAL SCIENCES

Zeitschrift für Semiotik

Innovative Methods in Multimodal Comics Research

Innovative Methods in Multimodal Comics Research: Introduction

Janina Wildfeuer, University of Groningen and Stephan Packard, University of Cologne

Summary. This special issue on innovative methods in multimodal comics research brings together linguistic as well as inter- and transdisciplinary contributions engaging with the semiotic and multimodal wealth of comics, graphic novels, and other forms of visual narratives. The contributions connect to recent research with new challenges and solutions and engage in dialogue across various approaches to the multimodality of comics. In our introduction to the issue, we want to address this 'multimodality of comics' further and give some explanatory notes on our understanding of this concept and the development of the field of research connected to it.

Keywords. Multimodality, methods, linguistics, semiotics, transdisciplinarity

Zusammenfassung. Diese Sonderausgabe zu innovativen Methoden der multimodalen Comicforschung versammelt sowohl linguistische als auch inter- und transdisziplinäre Beiträge, die sich mit der semiotischen und multimodalen Komplexität von Comics, Graphic Novels und anderen Formen visueller Narrative beschäftigen. Die Beiträge knüpfen an die aktuelle Forschung an, stellen neue Herausforderungen und Lösungen vor und führen einen Dialog über verschiedene Ansätze zur Analyse der Multimodalität von Comics. In unserer Einleitung zum Heft wollen wir diese „Multimodalität des Comics" näher beleuchten und einige Erläuterungen zu unserem Verständnis dieses Begriffs und der Entwicklung des damit verbundenen Forschungsfeldes geben.

Schlüsselwörter. Multimodalität, Methoden, Linguistik, Semiotik, Transdisziplinarität

The multimodality of comics and the multimodal study of comics

Looking through the main trends of comics studies since the 1970s and at least one strand of the last two decades' explosion of research in the field,

one might easily be tempted to say that comics have always been studied as multimodal.

Two main aspects of comics as form have played a central role for this understanding: One, the interrelation of pictorial signs, i.e. images, in sequence; the other, the combination of pictorial signs with other expressive forms, most obviously and frequently by adding written words. The former is concerned with those devices and functions that are specific to the comics' aesthetics, the arrangement of pictures in panels and panel sequences. The latter points to another dimension of semiotic interrelations frequently but not necessarily encountered in comics: The confrontation of pictorial with scriptorial elements. Summarizing comics in the Italian monicker *fumetti*, denoting that art form that uses speech balloons, is no less restrictive than an insistence on the comical intention of some early and influential North American subgenres. The French *bandes dessinées* more appropriately covers what we understand as comics here: continued strips of graphical designs, which might or might not enter into additional semiotic relations with lettered elements.

These image-text relations are properly problems of multimodality, even though their "central theoretic concern for comics studies" has been questioned and labeled "a more or less resolved issue" most recently (Spanjers 2021: 95). Most importantly, the erroneous restriction of comics' aesthetics to these considerations mirrors a profound misunderstanding of multimodality: Not only need comics not entertain a multimodal relationship between script and pictures, but the underlying relationships between panels and panel elements is properly multimodal in itself even before written language appears on the page. It is then insufficient to subsume the multimodality of comics under the art form's confrontation with an established dominance of written words as hegemonic discourse (Frahm 2010), or with the rise of the flood of pictures, or a pictorial turn, in the last 150 years or so (Mitchell 1994), although both those contexts are relevant in most of the historical spaces in which comics appear (Packard 2017). Systematically, the correlation between pictures in a typical comics strip is always already engaged with several semiotic modes.

The multimodal study of comics, therefore, goes far beyond a general distinction of two main expressive forms: It fundamentally assumes that comics are significantly complex and multidimensional and that all their semiotic and meaning-making elements, i.e. semiotic resources and/or modes, should be examined, particularly with regard to their interplay. While a general focus on the combination of words and pictures, or text and images, of course pays some tribute to this kind of analysis, there is more to be said about both smaller units within these two categories, e.g. about (motion) lines, colors, certain framings or the use of fonts and/or punctuation, as well as larger units such as the overall page layout, the construction of metaphors, or the unfolding of narrative or argumentative structures.

Multimodal analyses of this kind take an integrative approach to shed light on the different roles and functions of these individual components of

a comic. At the same time, they examine comics' complex and systematic intersemiotic interplay that guides and influences readers. Comics as multimodal artefacts can therefore be seen very broadly as "communicative situations […] which rely upon combinations of different 'forms' of communication to be effective" (Bateman et al. 2017: 7) – and it is a multimodal task to examine this effectiveness from various perspectives.

From linguistics and semiotics to multimodality research proper

An interest in the complex interplay of expressive forms, or modes, in comics has certainly been present in the early fascination with semiotics that emerged with isolated but thorough accounts and encompassing theories of comics production and interpretation. Since Umberto Eco's (1964 [1972]) convincing arguments for a semiotic reading of comics that does not only focus on verbal signs, the work of Roland Barthes (1967, 1977) has played a significant role for expanding linguistic analyses to include non-linguistic units. Several other works use comics as a playing field for new insights into textual comprehension and referential movement (Krafft 1978); they speak from a creator's experience while attempting some overarching definitions and conceptual clarifications (Eisner 1985); or they continue structuralist semiotics' straddling of linguistic and aesthetic approaches into a language of comics strips (Barbieri 1990; see also the overview of linguistic approaches in Bateman and Wildfeuer 2015). The importance of semiotics for this early, fragmented research discourse is still mirrored in the formalist approach of McCloud's seminal and popular *Understanding Comics* (1994: 9), in which comics are defined as "juxtaposed pictorial and other images in deliberate sequence". It might be worth noting that the suggested helplessness of the phrase 'and other' here mirrors the same tension between categorizing the specific aesthetics of comics either by means of a confrontation between pictorial and non-pictorial signs, or by the juxtaposition of pictures alone.

 If comics studies have drastically grown in scope and ambition since McCloud, one of the results has been the relegation of a particular branch of semiotic accounts and their questions to just some of the strands of investigation. To overcome this, the debate can be broadened by asking 'what more semiotics can do for comics studies' – such as the employment of semiotics for questions of media sociology, historical differentiation, or ideological criticism (Packard and Wilde 2022). The rise of a particular multimodal semiotics and linguistics in the last three decades exactly offers potential for yet another such 'more', building from Kress and van Leeuwen's pioneering reading of images and the description of multimodal discourse (1996, 2001) and continuing through expansions and approaches towards systematization (Kress 2010; Jewitt 2014; Klug and Stöckl 2016; Bateman et al. 2017). Many recent approaches to comics studies (including Lim 2007; Forceville et al. 2010; Forceville 2011; Cohn 2013; Bateman

and Wildfeuer 2014; Cohn 2016; Dunst et al. 2018) have already followed these pioneering attempts and developed individual research foci on some of the most daunting questions from a multimodal perspective.

Looking at the developments of this field or discipline of multimodality research over time, the potential for 'more' has emerged from the ongoing diversity of and interest in the field. While multimodality research has for some time been described as a discipline of its own (e.g. Wildfeuer et al. 2019), it is now rather seen as a "stage of development within a field" (Bateman 2022: 49) that many different fields and disciplines undergo by entering their own multimodal phase with new or different interest in multimodal phenomena. As we have shown above, comics studies is surely not a discipline that only now develops an interest in multimodality as such. On the contrary, it is in fact a field or discipline that has experienced a similar development: comics have seen interest from several disciplines and these disciplines have committed to the development of comics studies with their own theoretical principles and methodological tools. Multimodal analyses of comics are just one important strand that has, however, profited significantly from both the development of comics studies in general and the further growth of multimodality research in particular.

Placing multimodality anew in the field of comics studies can now lead to even more productive work that pushes both fields forward, but also gives reason for the establishment of multimodal comics studies as a field of its own.

Multimodal comics studies

Our goal in this issue is exactly this: We aim at relating disciplinary contributions from both comics studies and multimodality research effectively in order to strengthen the research field of multimodal comics studies.

In our call for papers for this issue we originally and explicitly challenged comics studies, as an interdisciplinary field with a rich array of opportunities and challenges, to investigations of multimodality. As a result, interestingly – and thankfully!, we received contributions that go beyond the simple application of existing theories and methods from comics studies to multimodal analyses, or from multimodality research to comics studies. Instead, all articles in this issue develop their own particular combination and integration of approaches stemming from several connected fields such as semiotics, linguistics, literary theory, culture and media studies, empirical cognitive studies, and aesthetics. This shows a broad toolbox of methods and perspectives, partially complementary, that is available for the field of multimodal comics studies:

J o h n A. B a t e m a n argues that a semiotic framework built around multimodality is better suited to reflect and integrate recent conceptual and empirical insights into the cognitive and semantic properties of comics comprehension than some established literary or aesthetic approaches. In a

critical dialogue with literary and cultural critics, his contribution explores the potency of a multimodal semiotics to illuminate not only the possibility, but also the specific complexity and significance of some chosen readings of mainstream comics such as *Watchmen* and *MAUS*.

Bateman and many others proceed from a model of pictorial comprehension that assumes a segmentation of dense pictorial arrangements into graphically and functionally distinguishable elements. The empirical foundation of these ideas is widened by L a u r e n E d l i n and J o s h u a R e i s s's series of experiments measuring agreement. Through annotation tasks, general problems of panel segmentation are revisited with a specific focus on what constitutes background information, and whether readers can agree both on the recognition of such elements as opposed to other stylistic conveyances. The complementary continuation of aesthetic appreciation emphasized by M i l o š T a s i ć and D u š a n S t a m e n k o v i ć's close reading of Lobačev's comic book adaptation of the epic *Dušan's Wedding* demonstrates how a semantically rich reading can still be articulated and better understood in dialogue with particular attention to formal elements and stylistic realizations.

L u k a s W i l d e and S t e p h a n P a c k a r d consider two aesthetic particularities of comics from the point of view of a multimodal semiotic approach: Wilde discusses comics as multimodal in the sense that they combine at least two semiotic modes, a cognitive reconfiguration of lines on paper into the depicted conceptions of bodies in space, and a more elaborate interpretational mapping that revolves around the idea of an unreliable iconicity prompted by the cartoonish style prevalent in comic books. In German artists Sascha Hommer's work, Wilde finds evidence of the delineation between both modes while elucidating the artfulness of the oeuvres. Packard looks at the concatenation of images in sequence and discusses how the modes of interpretation explicated by backtracking and re-evaluating fortuitously ambiguous elements may support either a grammatical or an aesthetic model of comics comprehension, aiming to integrate both in a semiotic pragmaticist view and to point out the Romantic echoes of hermeneutic theory involved in perspectives that separate or even oppose the two to one another.

While these semiotic approaches take the fundamental observation of a multitude of semiotic modes into conceptions from other semiotic theories, J a n i n a W i l d f e u e r concludes the issue with an overview and argumentative re-perspectivation of multimodal discourse analysis in a stricter sense. Discussing the basis for such a theory of signs in philosophy and logic, she presents a set of discourse relations that reflect at least some of the grounding processes of semiosis involved in reading a comic.

We are greatly indebted and thankful to all of the contributors to this issue as well as the journal's editors and the precise and constructive work of all peer reviewers. We would also like to thank Alex Dreßen and Marian Kirwel for their assistance in preparing the various chapters for publication. Finally, thank you for your attention. We look forward to continuing each of

these and many more discussions as the field of multimodal comics studies continues to grow.

Bibliography

Barbieri, Daniele (1990). *Il Linguaggio del Fumetto*. Milano: Bompiani.

Barthes, Roland (1967). *Elements of Semiology.* Farrar, Straus & Giroux: London.

Barthes, Roland (1977). *Image – Music – Text.* Farrar, Straus & Giroux: London.

Bateman, John A. (2022). Multimodality, where next?–Some meta-methodological considerations. *Multimodality & Society* 2, 1, 41–63.

Bateman, John A. and Janina Wildfeuer (2014). A multimodal discourse theory of visual narrative. *Journal of Pragmatics* 74, 180–208.

Bateman, John A. and Janina Wildfeuer (2015). Linguistically oriented comics research in Germany. In: Neil Cohn (ed.). *The Visual Narrative Reader*. London: Bloomsbury, 19–65.

Bateman, John A., Janina Wildfeuer and Tuomo Hiippala (2017). *Multimodality. Foundations, Research and Analysis – A Problem-Oriented Introduction*. Berlin and Boston: De Gruyter Mouton.

Cohn, Neil (2013). Beyond speech balloons and thought bubbles: The integration of text and image. *Semiotica* 197, 35–63.

Cohn, Neil (2016). A multimodal parallel architecture: A cognitive framework for multimodal interactions. *Cognition* 146, 304–323.

Dunst, Alexander, Jochen Laubrock and Janina Wildfeuer (eds.) (2018). *Empirical Comics Research. Digital, Multimodal, and Cognitive Methods*. New York: Routledge.

Eco, Umberto (1964 [1972]). The myth of Superman. Review of 'The Amazing Adventures of Superman'. English translation by Natalie Chilton. *Diacritics* 2, 1, 14–22.

Eisner, Will (1985). *Comics and Sequential Art*. Tamarac: Poorhouse Press.

Forceville, Charles (2011). Pictorial runes in 'Tintin and the Picaros'. *Journal of Pragmatics* 43, 3, 875–890.

Forceville, Charles, Tony Veale and Kurt Feyaerts (2010). Balloonics: The visuals of balloons in comics. In: Joyce Goggin and Dan Hassler-Forest (eds.). *The Rise and Reason of Comics and Graphic Literature: Critical Essays on the Form*. Jefferson: McFarland & Co, 56–73.

Frahm, Ole (2010). *Die Sprache des Comics*. Hamburg: Philo.

Jewitt, Carey (ed.) (2014). *The Routledge Handbook of Multimodal Analysis*. Second Edition. Milton Park, Abingdon and Oxon: Routledge.

Klug, Nina-Maria and Hartmut Stöckl (eds.) (2016). *Handbuch Sprache im multimodalen Kontext*. Berlin and Boston: De Gruyter.

Krafft, Ulrich (1978). *Comics lesen. Untersuchungen zur Textualität von Comics*. Stuttgart: Klett-Cotta.

Kress, Gunther (2010). *Multimodality. A Social Semiotic Approach to Contemporary Communication*. London: Routledge.

Kress, Gunther and Theo van Leeuwen (1996). *Reading Images. The Grammar of Visual Design*. First Edition. London: Routledge.

Kress, Gunther and Theo van Leeuwen (2001). *Multimodal Discourse. The Modes and Media of Contemporary Communication*. London: Arnold.

Lim, Fei Victor (2007). The visual semantics stratum: Making meaning in sequential images. In: Terry D. Royce and Wendy L. Bowcher (eds.). *New Directions in the Analysis of Multimodal Discourse*. New Jersey: Lawrence Erlbaum Associates, 195–214.

McCloud, Scott (1994). *Understanding Comics. The Invisible Art*. New York: Harper Perennial.

Mitchell, William J. T. (1994). *Picture Theory*. Chicago: University Press.

Packard, Stephan (2017). Sagbares und Sichtbares jenseits von Schrift und Bild: Aufteilungen des Sinnlichen im Comic. In: Matthias Harbeck, Linda-Rabea Heyden and Marie Schröer (eds.). *Comics an der Grenze. Sub/versionen von Form und Inhalt*. Berlin: Bachmann, 131–143.

Packard, Stephan and Lukas R.A. Wilde (2022). What more can semiotics do for comics? Looking at their social, political, and ideological significations. *The Social, Political, and Ideological Semiotics of Comics and Cartoons. punctum: International Journal of Semiotics* 7, 2, 5–15.

Spanjers, Rik (2021). Text-image relations. In: Sebastian Domsch, Dan Hassler-Forest and Dirk Vanderbeke (eds.). *Handbook of Comics and Graphic Narratives*. Berlin and Boston: De Gruyter, 81–98.

Wildfeuer, Janina, Jana Pflaeging, John A. Bateman, Ognyan Seizov and Chiao-I Tseng (2019). *Multimodality. Disciplinary Thoughts and the Challenge of Diversity*. Berlin and New York: De Gruyter.

Asst.-Prof. Dr. Janina Wildfeuer
Rijksuniversiteit Groningen
Faculteit Letteren – CIW
Oude Kijk in 't Jatstraat 26
9712 GK Groningen
The Netherlands
E-Mail: j.wildfeuer@rug.nl

Prof. Dr. Stephan Packard
Universität zu Köln
Institut für Medienkultur und Theater
Meister-Ekkehart-Str. 11
50937 Köln
Germany
E-Mail: packard@uni-koeln.de

Katharina Staubach

Multimodale Sehflächen auf den T-Shirts von Jugendlichen

Eine semiotische Studie zu Prozessen jugendlicher Selbstinszenierung

Die Frage, wann wir durch unsere T-Shirt-Aufdrucke etwas ausdrücken wollen und wann sie rein dekorative Funktion erfüllen, ist lange Zeit ein zentrales Desiderat semiotischer Forschung geblieben. Die Studie greift diese Frage am Beispiel der Altersgruppe der Jugendlichen auf und verwendet dabei eine innovative Methodentriangulation, bei der linguistische Analysen mit konkreten Rezeptionsdaten in Beziehung gesetzt werden.

Aus linguistisch-semiotischer Perspektive wird herausgearbeitet, welche Aufdrucke sich auf der Kleidung von Jugendlichen finden, und welches Potential sie für Prozesse der Selbstinszenierung bieten. Dabei wird das Material mittels eines zuvor ausgearbeiteten Kategoriensystems sowohl inhaltlich als auch formal strukturiert. Anhand qualitativer Befragungsdaten in Form von Gruppendiskussionen und Einzelinterviews wird anschließend der Frage nachgegangen, wann diese Aufdrucke zu gezielten Kommunikations- und Selbstinszenierungszwecken eingesetzt werden und wann sie eine rein dekorative Funktion erfüllen. Dabei werden auch geschlechtsspezifische Unterschiede untersucht.

STAUFFENBURG VERLAG

Multimodal Semiotics for the Analysis of Comics and Graphic Novels

John A. Bateman, Bremen University

Summary. The broad challenge taken on in this contribution is to attempt to reconcile more literary-hermeneutic approaches to comics and graphic novels, on the one hand, and closer, more fine-grained analytic accounts, on the other. This will be done by applying a semiotic framework that takes the phenomenon of multimodality as its primary organising principle. The discussion begins by showing how several assumptions commonly made in the comics research literature concerning the nature of semiotic accounts need to be redrawn because of substantial developments in recent years directly relevant to the treatment of complex media, such as comics and graphic novels. This appears not to have been realised sufficiently in many discussions of those media. Several examples of complex narrative will be drawn on to illustrate the possibilities of a broader semiotic account that nevertheless maintains a tight connection to the details of form, thereby opening up possibilities for more focused research on a variety of phenomena previously often grouped rather loosely under Groensteen's notion of braiding.

Keywords. Multimodality, discourse semantics, comics, graphic novels, braiding, semiotic modes

Zusammenfassung. Dieser Beitrag stellt sich der umfassenden Herausforderung, primär literarisch-hermeneutische Annäherungen an Comics und Graphic Novels mit detaillierteren und präziseren Analysen in Einklang zu bringen. Die hierfür verwendeten semiotischen Konzepte orientieren sich primär an dem Phänomen der Multimodalität. In der Comicforschung haben sich eine Reihe von Annahmen über Semiotik etabliert, die aufgrund subtsanzieller Fortschritte in der Behandlung komplexer Medien wie Comics und Graphic Novels in den letzten Jahren dringend revisionsbedürftig sind, was in vielen Diskussionen über diese Medien bisher nicht ausreichend berücksichtigt worden ist. Anhand mehrerer Beispiele aus komplexen Erzählungen sollen die Möglichkeiten einer allgemeineren semiotischen Darstellung veranschaulicht werden, die den formalen Details ihres Gegenstands dennoch eng verbunden bleibt und damit Möglichkeiten für eine gezieltere Erforschung einer Vielzahl von Phänomenen eröffnet, die bisher oft eher lose unter Groensteens Begriff des „braiding" zusammengefasst wurden.

1.   Introduction: the challenge

One of the properties of comics and graphic novels that is accepted, and even celebrated, across the board is the extreme variability and range of the forms of expression that they employ – that is, their m u l t i m o d a l i t y. Indeed, one particularly prominent, almost definitional, aspect of comics is their positioning between, or across, some of the most basic distinctions traditionally drawn between medial forms. This raises substantial semiotic challenges. On the one hand, they (most commonly) rely on static pictorial depictions aligned with histories of visual representation; and, on the other hand, they simultaneously rely (most commonly) on the essential temporalities of verbal language and sequence. Although also sometimes reduced to the microcosm of the division between 'words' and 'images', the resources available are considerably broader. For example, even when focusing specifically on just those devices available for constructing 'character's subjectivity', a core facet of narrative, Mikkonen (2015) includes such diverse technical resources as:

> facial expressions, gesture, body language, gaze, and the character's position in the image in relation to other visible objects […] metaphorical images and pictograms (emanata, symbolia) […] spatial articulation, such as framing, sequencing, breakdown, page layout, and tabulation […] visual style, for instance, blurry images or changes on a scale between graphic realism and a simpler cartoon style […] (Mikkonen 2015: 101–102).

Although this richness and variety of resources is very much taken for granted within more interpretative or literary traditions to studying comics and graphic novels, traditional semiotic accounts and theories of communication are often stretched well beyond their limits when confronted with such diversity. Indeed, substantial questions remain, concerning how best to characterise the sheer range of distinct contributions and their combinations in producing coherent unfolding wholes. It is then understandable that work within more of a literary or hermeneutic orientation has with considerable justification criticised traditional semiotics for being overly restrictive, reductive, language-oriented, and structural (e.g., Postema 2013; Miodrag 2013; Horstkotte 2015).

The main goal of the current article will be to show how more finely articulated accounts of multimodal semiosis open up new possibilities for dealing with the complexity and diversity of sophisticated visual storytelling of the kind increasingly found in comics and graphic novels. In many respects, this is to echo Groensteen's (2007 [1999]) call for a 'neo-semio-

tics' but, as we shall see, in a way that maintains a far tighter theoretical (and practical) hold on the selected objects of analysis. Achieving a more integrative account will then itself demand refinements to some core constructs of semiotic inquiry, which will be provided by the specific approach that we will build on – that introduced in Bateman, Wildfeuer and Hiippala (2017).

The structure of the article is as follows. We begin by setting out some of the problematic relationships discussed between comics analysis and semiotics in order to locate more precisely where developments have been required. We then introduce the approach to multimodality and multimodal semiotics that we draw on, briefly mentioning some critiques that have been made of its application to sequential visual narrative previously. During this introduction, the article illustrates the concepts provided with respect to some examples of multimodality occurring within single comics and graphic novel panels discussed in the literature. Following this, the discussion moves to an analysis of rather more complex narrative trajectories and visual design suggested previously to be problematic for semiotic or 'linguistically'-oriented accounts. We show in each case how a multimodally more aware framework renders the examples more straightforward to analyse as well as encouraging more detailed and revealing characterisations. Finally, we summarise the points that have been made through the discussion and set out some directions for future development.

2. Some traditional misunderstandings concerning semiotics

The core argument pursued here will be that adopting a more contemporary multimodal semiotics provides a scaffold suitable for supporting the kind of complexity observed even in sophisticated comics and graphic novels. This needs to be argued because the relevance and ability of approaches rooted in semiotics and related extended linguistic approaches has been widely rejected in several discussions pursued in the field. Unfortunately, as set out with particular force by Cohn (2014), many of these discussions appear to target a semiotics and a linguistics that would be more at home in the 1970s than the 2020s. Our first task must therefore be to refocus attention on what is now available from semiotic accounts and how this substantially differs from the positions critiqued in the comics and graphic novels literature.

Many of the problems with the positions articulated can be demonstrated by means of a brief consideration of the kinds of concepts that are bundled together when discussing semiotics and linguistic approaches. This is generally carried out in a manner unmotivated by, and incompatible with, the current state of the art. The core dimensions of this positioning may be summarised thus:

i. an idea that semiotics operates solely in terms of 'semiotic codes' whose use presumes 'rigid meanings' and the exclusion of 'inference';

ii. a notion that any mention of 'grammar' commits to both (a) questions of 'grammaticality' in the Chomskyan sense that emerged in the late 1950s and early 1960s (cf. Chomsky 1957) and (b) treatments of phenomena in terms of a strict and relatively straightforward linearity;

iii. the presupposition that semiotics entails a focus on verbal-visual conventions or codes with an accompanying lack of contact with materiality and embodiment, i.e., the role of the body in perception and meaning-making as a complement to, or grounding for, representations; as well as;

iv. a reliance on artificially strict separations between semiotic systems working with signs classified as 'iconic' and 'symbolic', or 'arbitrary', 'natural', and 'conventional', and similar.

Although there may still exist approaches to semiotics reiterating these earlier organisational features, they have little bearing on the current discussion precisely because they can now generally be considered inadequate for any complex media, including treatments of language just as much as comics and graphic novels. Several prominent approaches in the discussion of comics and graphic novels then spend time arguing against 'straw-person' positions in a way that simultaneously restricts access to analytical tools crucial for engaging successfully with complex media usage.

Specific illustrations of this tendency are readily found among some of the leading scholars in the field. For example, partly drawing from and extending cases considered in Cohn's (2014) critique, Postema and Hick both find it self-evident that any notion of grammar or 'rules' is symptomatic of fixed meanings and rigidity and so should be considered singularly out of place for comics and graphic novels:

> […] images communicate largely without rules […] the smallest elements of images have no set meanings, and the way these elements are combined or even repeated are not governed by rules like grammar (Postema 2013: xvi).

> [T]he notion of a *syntax* of comics is a difficult concept to even wrap one's head around […] it is not at all clear how (if at all) systematized concatenation rules might even be described – and if there *are* such formalizable rules, we certainly don't know them (Hick 2012: 140, original emphasis).

Horstkotte also finds the application of any such notions as 'universal grammar' (however this may be defined) as misleading and inappropriate for comics and graphic novels, even though, apparently, now (somehow) acceptable for language and for (at least Hollywood) films – itself a curious position upheld by very few outside perhaps a particularly narrow Chomskyan tradition:

> the style of each graphic narrative is much more variable and distinctive than is the case in other narrative media […] there is no universal grammar for this decod-

ing as there is in verbal narrative in a natural language, or in the established nar-
rative format of the Hollywood movie (Horstkotte 2015: 32).

Further authors could be cited as similarly adopting at best questionable
views on what semiotics does and does not include (e.g., Miodrag 2013),
but the general point here should already be clear: it has become such a
standard trope to suggest that communication via 'sequential images' can-
not be considered similar to communication via language that more focused
engagement with the issues is deemed unnecessary. Approaches to com-
ics and graphic novels, particularly from more literary and cultural studies
perspectives, then reject any comparison of language properties and those
of comics often on little more basis than holding such comparisons to appear
wedded to outdated structuralist principles and so simply not appropriate
for 'modern' accounts.

The many presuppositions at work in such positions, combining ideas
that 'rules' require 'set meanings' for minimal image elements, that 'gram-
mars' are essentially linear 'concatenation rules', that talking of grammar
commits to universalist claims at odds with cultural diversity and stylistic
creativity, and so on, are then freely extended to analytic accounts drawn
from semiotics as well. This readily culminates in striking admonitions such
as the following, also from Horstkotte:

> A responsible comics hermeneutics would do well to move away from the linguis-
> tic-structuralist idea that comics narrative has a 'grammar' (Eisner 2008: 2) and
> that this grammar entails a linear reading. […] An understanding of comics in terms
> of signs as it is proposed, for instance, by Ole Frahm (2010) is reductive (Horst-
> kotte 2015: 34–35).

In these few sentences Horstkotte conflates a host of theoretically ques-
tionable assumptions as if, for comics and graphic novels, they were self-ev-
idently the case. 'Grammar' is equated with a requirement of 'linear read-
ing', something allegedly inappropriate for readers' engagements with com-
ics and graphic novels, while notions of 'signs' are definitely to be avoided.
Suggesting that the consideration of such constructs would even constitute
a lack of 'responsibility' is clearly a rather strong position and, as we will
see as we proceed, demands substantial revision.

Indeed, in the case of Frahm, it appears that the simple use of terms
such as 'sign' and 'semiotics' was sufficient to attract Horstkotte's critique,
even though the position developed by Frahm rejects precisely the kind of
broadly 'Saussurean' traditional linguistic-structuralist approaches that Horst-
kotte is opposing. Frahm's proposals are actually aligned far more closely
with pragmatically-oriented accounts of meaning-making that argue that
multimodality should not be seen in terms of systematic relations holding
between, most typically, images and texts, but instead as a specific achieve-
ment of socioculturally situated recipients striving in particular contexts to
make sense of what they are seeing (and reading). In short, multimodal

meaning-making is to be considered, in Saussurean terms, as a matter of *parole* and not of *langue* (Frahm 2010: 14–15), as situated action rather than systems of pre-established relationships (cf. Bucher 2011).

But both perspectives, Horstkotte's view of just what should and should not be considered a 'responsible' approach to comics and graphic novels and Frahm's claims of nonsystematicity among intermodal relationships and meaning-making, are equally inappropriate. Adherence to these kinds of presuppositions raises profound problems for the analysis of comics and graphic novels because they make it difficult e v e n t o i n v e s t i g a t e s y s t e m a t i c a l l y how interpretation of such media is possible. All would agree that there are considerable regularities to be uncovered and discussed, but these become marginalised in favour of claims for uniqueness, distinctiveness, and individual subjectivities – all else threatening to be, and threatened with being, 'reductive'.

As comics and graphic novels researchers attempt to refine their analyses, however, such avoidance strategies give rise to growing theoretical and methodological tensions. Kukkonen (2013), for example, proposes a broadly literary analytic position that also quite explicitly seeks to be a 'cognitively' based account of comics and comics interpretation as well. As she suggests:

> [a]s readers move from panel to panel […] they connect the clues (both verbal and visual) into a common mental model (Kukkonen 2013: 32).

Such positions adapt to comics an approach increasingly found in literary analyses that claim a cognitive foundation (cf. Sweetser 2012), which is itself a logical continuation of earlier reader-response frameworks (Iser 1978), in which textual interpretation is viewed as a process of finding textual 'gaps' which a reader then fills inferentially, drawing on any knowledge necessary. The need for some notion of 'inference' when addressing the interpretation of media products, particularly aesthetically challenging media products, can probably now be taken as uncontroversial. The configurations relevant for Kukkonen's primarily literary concerns are consequently seen as

> textual effects that emerge from a combination of clues and gaps in the text triggering particular processes in the reader's minds (Kukkonen 2013: 178).

Analysis of this kind is typically couched in terms of descriptions that (a) propose what must or could have been taken as a clue by readers, and (b) plausible interpretations suggested for those clues. But, for such analyses to move beyond hypothesis and conjecture, however insightful such conjecture may occasionally be, one needs to be able to state in detail just what constitutes clues and gaps in any 'text' under investigation. This is clearly one of Kukkonen's aims as well, as she goes so far as to suggest that her 'middle range' inquiries into literarily relevant aspects of comics, such as

self-reflexivity, subversion, voice, gender, fictional minds and characters, etc., "should be, by and large, testable" empirically (Kukkonen 2013: 178). Unfortunately, the distance between testable hypotheses and middle range descriptions of the kind pursued by Kukkonen and others remains very large. It is by no means straightforward to approach this task in a principled fashion, which is one reason why Cohn, building with a firm empirical anchoring, tends to see such proposals as little more than promissory notes unlikely to be cashed out in the foreseeable future (Cohn 2018). Symptomatic here is then, how rarely empirical results from actual cognitive studies of media such as comics and graphic novels directly influence the literary side — connections drawn generally remain generalised, suggestive, or metaphorical. Conversely, attempts to probe the literary interpretative descriptions offered empirically are equally rare. It is precisely in mediating between these domains that a more developed semiotics can provide critical support.

To prepare the ground for this, it is crucial to defuse the situation described above in which semiotics is characterised in terms of 'semiotic codes' that commit to a rigidity in interpretation not found in comics and graphic novels (and most other media as well). Much traditional semiotics has simply failed to deliver useful tools here and it is on this basis that Kukkonen can, with some justification, assert that:

> the approach to comics that will serve as the framework of my analysis […] is based on clues and gaps, readers' inferences and the mental models and fictional minds they construct, *rather than on semiotic codes* (Kukkonen 2013: 50; emphasis added).

For Kukkonen and many other researchers in the area it appears clear that 'semiotic codes' and inference somehow stand in opposition.[1] Consequently, since comics and graphic novels evidently require inference, one must exclude semiotic codes from the discussion. Although there have been proposals in semiotics that go beyond the provision of more or less straightforward bundles of signifiers and signifieds deployed in a coding-decoding model (e.g., Greimas 1987; Lotman 1990), such developments have to date offered little benefits for detailed comics and graphic novels analysis. There has consequently been little motivation among comics researchers to recast often already insightful, if informal, analyses in more traditional semiotic terms.

This, then, is the impasse that this article seeks to redress. On the one hand, while there can be little doubt that Kukkonen is absolutely correct in arguing that any simple 'code'-based view of signs would be inappropriate as support for analysing comics and graphic novels, such positions are far removed from what is now available semiotically. In fact, taken strictly, it is less clear whether any real analysts actually assumed a simple model of semiotic codes as commonly characterised – even some of the most well-known developers of code-based models, such as Eco (1976), included discussion of inferences. The problem was that these processes, and means for modelling them, were still poorly articulated at that time. On the other

hand, notions of genres, of interpretative frames, and the like, all central to Kukkonen's account, are all results of conventionalised practice that are already essentially semiotic. Indeed, most constructs found in narratological-cal approaches to comics, including focalisation, narration, discourse/story, storyworlds, and many more, are already s e m i o t i c in any useful sense of the word. For such proposals to consider themselves 'beyond' semiotics or to render semiotics unnecessary, as often proclaimed, is consequently particularly damaging as it deprives the study of complex semiotic artefacts and performances of foundations and methodologies critical for driving analysis further.

The opportunity now to be seized, therefore, is to draw on more finely developed semiotics that make such kinds of description natural targets both cognitively and semiotically. The major developments that make this possible in accounts of semiotics, and particularly in the view from multi-modal semiotics adopted in this paper, can be divided into three broad areas, two of which will be taken up below:

- First and foremost, the division between code-based, non-inferential accounts, broadly labelled as Saussurean, and inference-based accounts must be rejected. Following a view more compatible with Peirce (cf. Bateman 2018), inference is always present. Consequently, the import-ant question for further analyses and frameworks becomes not 'if', but just what kinds of inference are necessary and when do they occur, oper-ating on what kinds of premises. Some of these will be near to percep-tion (building on iconicity); others will be quite distant from perception, involving larger narrative trajectories and discourse as we shall show.
- And second, one essential property of t e x t u a l i t y, namely that of g u i d i n g i n f e r e n c e, has to be properly incorporated. It is only then that it becomes possible to formulate mechanisms for focused analysis that both move beyond more subjective, post-hoc interpretations and establish systematically investigable bridges between concrete phenom-ena in any artefacts under analyses and broader interpretations. This contribution of textuality will be a component of all the examples dis-cussed below.

The third area, which we will not have the space to discuss, concerns a rejection of the old 'disembodied' views of signs implicit in Saussure and raised to a theoretical principle by Hjelmslev's 'algebraic' perspective on semiotic systems (Hjemlsev 1961: 105). Embodiment is now finding increas-ing application in analyses of comics and graphic novels (e.g., Kukkonen 2015) and is also taken as an inalienable component of the model of semi-osis drawn upon here (cf. van Leeuwen 1999; Bateman 2019); further dis-cussion is, however, beyond the current scope.

These developments substantially change the kinds of analysis that are possible and so, in the next section, we introduce a semiotics of this kind and begin to show how it applies to comics and graphic novels quite directly.

3. The move beyond 'semiotic codes': semiotic modes and multimodal semiosis

We have suggested that there is an urgent need to refurbish our semiotic foundations if the study of complex artefacts such as comics and graphic novels is to receive adequate support. One revitalised account of semiosis responding to the requirements given above is introduced by Bateman et al. (2017). Drawing primarily on formal and functional approaches to discourse interpretation (cf. Martin 1992; Asher and Lascarides 2003), embedded within a stratified view of semiotic systems (Halliday 1978), and incorporating insights from Peirce (Bateman 2018) and social semiotics (Kress 2010), the model seeks to help shape investigations of complex multimodal communication no matter what forms of expression and materials are employed. At the same time, the approach is also strongly oriented to the demands of empirical studies and corpus analyses (Bateman 2022a).

Since the model has been set out at length elsewhere, only a brief overview of its main features will be given here, focusing on what is relevant for the discussion and analysis of the examples below. The model takes as its starting point the material-ontological conditions necessary for communication to take place at all; this is then more aligned with semiotics and the philosophy of communication and of meaning rather than literary or cognitive traditions, although broad compatibility of results is always to be pursued. Compatibility, here, is understood in the sense of triangulation, rather than reduction, whereby alternative descriptions are placed in formal correspondence relationships to one another so as to support explorations of cross-domain predictions (cf. Smith 2012, 2022; Bateman 2022b). The general orientation provided by the framework offers increasing benefits the more complex a communicative form becomes. Consequently, for media such as comics and graphic novels, we consider it essential precisely because of the well documented sophistication of the meaning-making practices that these media mobilise.

The model's core construct is the s e m i o t i c   m o d e . Each semiotic mode is a bundle of semiotic mechanisms articulated at three distinct levels of abstraction, or s t r a t a . Least abstract, i.e., closest to materiality and direct perceptual access, are the particular formations of material established for meaningful expressions within a semiotic mode; this provides what we term the c a n v a s of that semiotic mode. The basis for any approach to multimodality within this framework consequently echoes Kress's notion of 'materiality socially-shaped' for the communicative purposes of some community of users (Kress 2010: 79). Next, to be usable for meaning-making, a semiotic mode imposes qualitative groupings and organisations on material regularities, re-described in terms of the structural configurations provided by the semiotic mode. This second level of abstraction is characterised following the organisational principles established for language in systemic functional linguistics, whereby structural configurations are modelled along the dimensions of a x i a l i t y   and i n s t a n t i a t i o n (cf. Martin

2014). Under 'axis' is understood systems of interrelated paradigmatic choices realised by syntagmatic, i.e., structural, configurations. Orthogonally to this, 'instantiality' then captures the fact that resources provide a description of what is possible, the potential, which may be actualised in use to a greater or lesser degree. As common in many current linguistic models, no division is drawn between the lexicon and 'grammar' – 'lexical' entries, regardless of their material make-up, are simply more or less completely specified instantiations of the expressive resources available. These 'lexicogrammatical' structural configurations are then analogous to Cohn's use of 'grammar' mentioned above and taken up further below. Finally, at the third level of abstraction, each semiotic mode offers a set of discourse semantic resources that serves to relate the structural configurations of the second level to their contextualised interpretations. Discourse semantics are essentially dynamic and non-monotonic, i.e., abductive in the Peircean sense (Bateman 2020), and so probably constitute the most important extension beyond previous semiotic descriptions as we shall see below.

The test for the existence, or not, of some particular semiotic mode is the extent to which material distinctions appear to require, or admit, characterisations both at the level of 'technical features' (i.e., the 'lexicogrammatical' semiotic stratum) a n d the level of discourse semantics. This may fall out differently for different communities of practice and, even within communities of practice, there will usually be differences in terms of the degrees to which individual community members have control over, or access to, the potentials available. No assumption is made that it is already straightforward to work with established categories such as 'word' or 'image' as there may well be practices that blur such boundaries. Indeed, superficially similar visual marks, such as lines, forms, shadings, and so on, all operate quite differently in written language, maps, graphs, diagrams or, indeed, many of the visual components of comics and graphic novels; semiotic modes enable this diversity to be formally captured. Any semiotic mode is accordingly a 'current best hypothesis' concerning how observable material regularities are to be explained as instances of communication for some community of practice. As we shall see, the considerable additional theoretical structuring provided by the multimodal semiotic view supports a tighter theoretical grasp of comics and graphic novels without restricting analytic attention to specific facets such as sequentiality or word-image combinations, while still remaining open to empirical results achieved within other models.

One further useful conceptual distinction given by the model is a clear separation of the semiotic resources that are used to 'communicate' (which, in the style of Peirce, includes aesthetic effects as well as more message-based components) and the institutionalised 'sites' where that communication takes place. A very similar distinction is drawn by Cohn, who has always argued that one needs to clearly differentiate between what he characterises in terms of 'Visual Language' and the comics or graphic novels 'themselves' (Cohn 2013). The analogy usually drawn to explain this dis-

tinction is that between verbal language and novels: novels are not language themselves but 'contain' uses of language. Comics and graphic novels are quite similarly sites within which varied forms of expression are mobilised. In the multimodal semiotic model, this relationship is formalised in terms of a new definition of the traditionally difficult term "medium". A medium is consequently defined as a socio-historically conventionalised combination of semiotic modes used for the achievement of some collection of communicative genres (cf. Bateman 2016). Genres bring with them sets of communicative goals and conventionalised solutions for their achievement. Those conventionalised solutions can then range over any of the semiotic modes available to the medium at hand.

Figure 1 presents a graphical overview of all the major components of the model. The diagram picks the vantage point of some specific medium, be that spoken language, comics, graphic novels, or some other institutionalised form of communication. Each such medium is constituted by some collection of semiotic modes, indicated by 'containment' in the diagram and represented by the three repeated ovals running down the left of the figure; any number of semiotic modes might be conventionalised as being relevant for a medium and that might itself vary over time as social needs and technological possibilities change. Individual semiotic modes are structured internally as described above and, as a consequence, are represented here using a notation for stratally organised semiotic systems based on co-tangential circles originally attributed to Halliday and presented in Martin and Matthiessen (1991).

On the extreme left of the figure, we see that the collected material strata of the semiotic modes then constitute the 'canvas' of the medium as a whole. This is a way of expressing the general premise of the model that, in any medium, material regularities need to be 'claimed' by some semiotic mode in order to be meaningful. Those meanings are then created and interpreted primarily by the operation of the discourse semantics strata of the semiotic modes, indicated in the figure on the right of the medium 'box'. All discourse semantic strata abductively construct discourse configurations and hypothesised relations holding among those configurations; in the multimodal case, these configurations then serve as interfaces across semiotic modes. Intersemiotic linkages between discourse configurations originating in distinct semiotic modes are abductively hypothesised in a process similar to a formal notion of conceptual blending (cf. Kutz et al. 2015) that we illustrate informally below.

Finally, on the extreme right of the figure, we see that the overall motivations for pursuing particular directions of interpretation or production in an extended discourse are assumed to be given by genre: different genres raise distinct communicative goals and those goals may then prioritise particular lines of discoursal development rather than others. Thus, when analysing any multimodal communication, two broad 'points of access' are defined methodologically: we might proceed focusing on material regular-
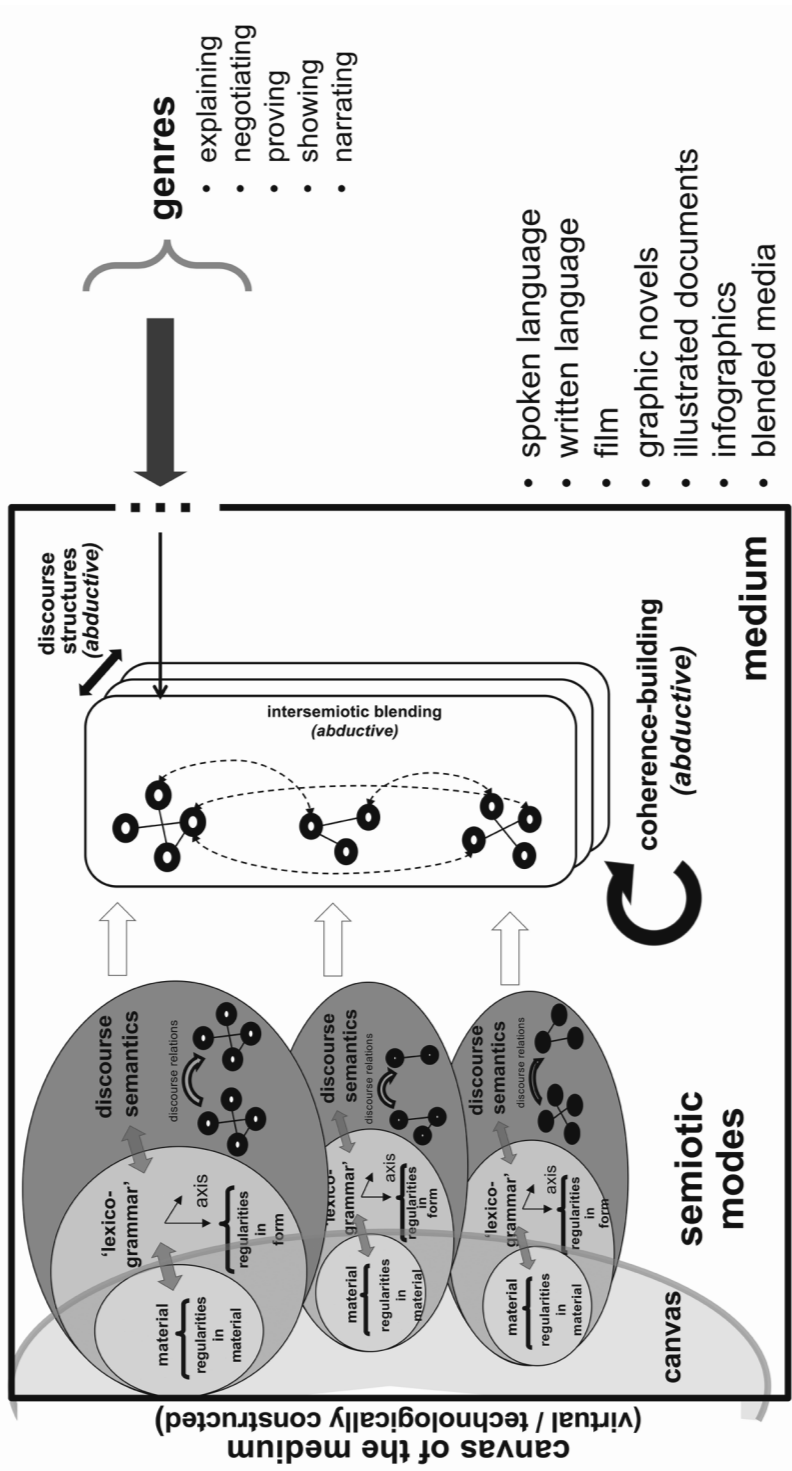
Fig. 1. Graphical summary of the model of multimodal communication defined by Bateman, Wildfeuer and Hiippala (Bateman et al. 2017: 124, Bateman 2016).

ities, i.e., by moving from left to right in the figure, and we might proceed from the communicative goals assumed to be active, as given by the selection of genre(s) upper right. Usually, both directions would be pursued together. In contrast, then, to the multimodal model of Cohn (2016) in which certain 'modes' are already assumed on broadly neuro-cognitive grounds (e.g., visual, bodily, verbal), the model adopted here always considers the question of the modes active in a medium to be an empirical issue where answers may fall out differently across times and cultures.

It is in this sense that the current model provides a stronger framework for engaging with comics and graphic novels scholarship more broadly, precisely because there is no requirement, or indeed expectation, that the number of semiotic modes active within these media is limited to some small pre-given collection, such as 'written language' and 'drawn pictorial image', or similar. By these means, the model opens the door to inputs from a variety of disciplines and traditions, from the fine arts to information design, from typography to press photography, and many more.

Different uses of material of this kind are then brought together formally by the inter-medial relationship of d e p i c t i o n (Bateman et al. 2017: 126–128). Depiction is similar to what in several research traditions is discussed in terms of "intermediality" or "remediation" (cf., e.g., Rajewsky 2005; Elleström 2010), but focuses more on the specific case when some medium shapes the material available to it in order to give the 'impression' of, or to depict, another medium. The use of a visual metaphor is thus here quite intentional. Examples in comics and graphic novels would be the inclusion of photographs, infographics, diagrams, paintings, newspapers, musical scores, and so on. These are all very different media, each with their own principles of organisation (semiotic modes), histories of development (participation in media), and corresponding literacies and their ways of assigning meaning to observed material traces are all quite distinct. In all cases, however, it is their dynamically constructed discourse configurations that are seen as providing the 'glue' by which inter-semiotic relations can be strategically generated. Discourse semantic configurations thus offer the minimally necessary 'interfaces' for allowing contributions couched within any of these traditions to communicate with one another.

It will be helpful at this point to relate this abstract schema to an example from our target media more directly, at the same time anchoring this into active literary discussions of those media as well. Rajewksy (2010), for example, argues compellingly that despite a fashion to consider medial boundaries old fashioned, the existence of individual media that are placed in various relationships to one another remains an important source of aesthetic meanings and effect. The distinction now introduced between media and semiotic modes allows us to refine this intuition considerably. It is certainly the case that comics offer substantial illustrations of productive intermediality at work, but this is argued here to be a possibility open to any medial form, simply by virtue of what it is to be a medial form at all. One of the consequences of accepting the role of semiotic modes as

resources for shaping the material on offer in some media artefact is that it then becomes more straightforward to show that many common equivocations are both unnecessary and misleading for analysis.

We illustrate this first with one of the examples discussed by Rippl and Etter (2015: 209), the panel from Fred's (2011) *Philémon: L'Intégral* shown on the left of figure 2. The observed feature here is that protagonists approach an island from the air but the island, otherwise drawn naturalistically, clearly exhibits the shape of the letter 'N'. This phenomenon is well-known but the descriptions offered are generally less than adequate and introduce confusion without need. Specifically, Rippl and Etter (2015) conclude with respect to this panel that:

> The fact that textual elements are turned into iconic elements and forms that are primarily looked at and not read demonstrates that graphic narratives question the clear division between words and pictures (Rippl and Etter 2015: 208).

Although this formulation is quite frequent, particularly in literary interpretations, its presuppositions are potentially quite misleading. There is not so much a 'questioning' of a division than rather a clear manifestation of the division's existence – there is 'simply' the far more sophisticated juxtaposition, or co-deployment, of multiple semiotic modes with respect to the same, shared materiality. This corresponds well with Rajewsky's point above about the importance of distinct media for aesthetic effects but also gives us the explicit analytic means for tracking both their distinctions and combinations.

This common intersemiotic relation, labelled h o m o s p a t i a l i t y by Lim (2004: 240–241), is depicted graphically on the right-hand side of figure 2. The phenomenon in general plays with the fact that two or more distinct semiotic modes may be materially (here: spatially) co-ordinated: that is, there is no requirement in the model that the regularities found in any material will be exhausted by descriptions from a s i n g l e semiotic mode. We see this occurring here in several rather distinct ways. For example, the combination of a pictorial representation and a frame to form a panel similarly uses the same materiality, but is less often remarked upon as a combination of 'modes' because of the prevalence of framing conventions for pictorial content. This relates to Smith's (2015) useful characterisations of the difference between the use of 'frames' in photographic arts and drawn media such as (but not restricted to) comics: whereas in both media some content is being 'framed' the fact that there is no m a t e r i a l d i s t i n c t i o n between the drawn comic panel contents and the drawn panel frame shows that a single material is being used for very different (semiotic) purposes. This is equally the case for the island and the letter 'N', although this is less strictly conventionalised and so stands out more as a distinctive design choice, even if conventionalised enough to form a recognisable trope. In both cases, however, combinations of contributions are being effected by means of material anchoring and blended discourse interpretations. This does not then question a division – indeed, it is only b e c a u s e of the divi-
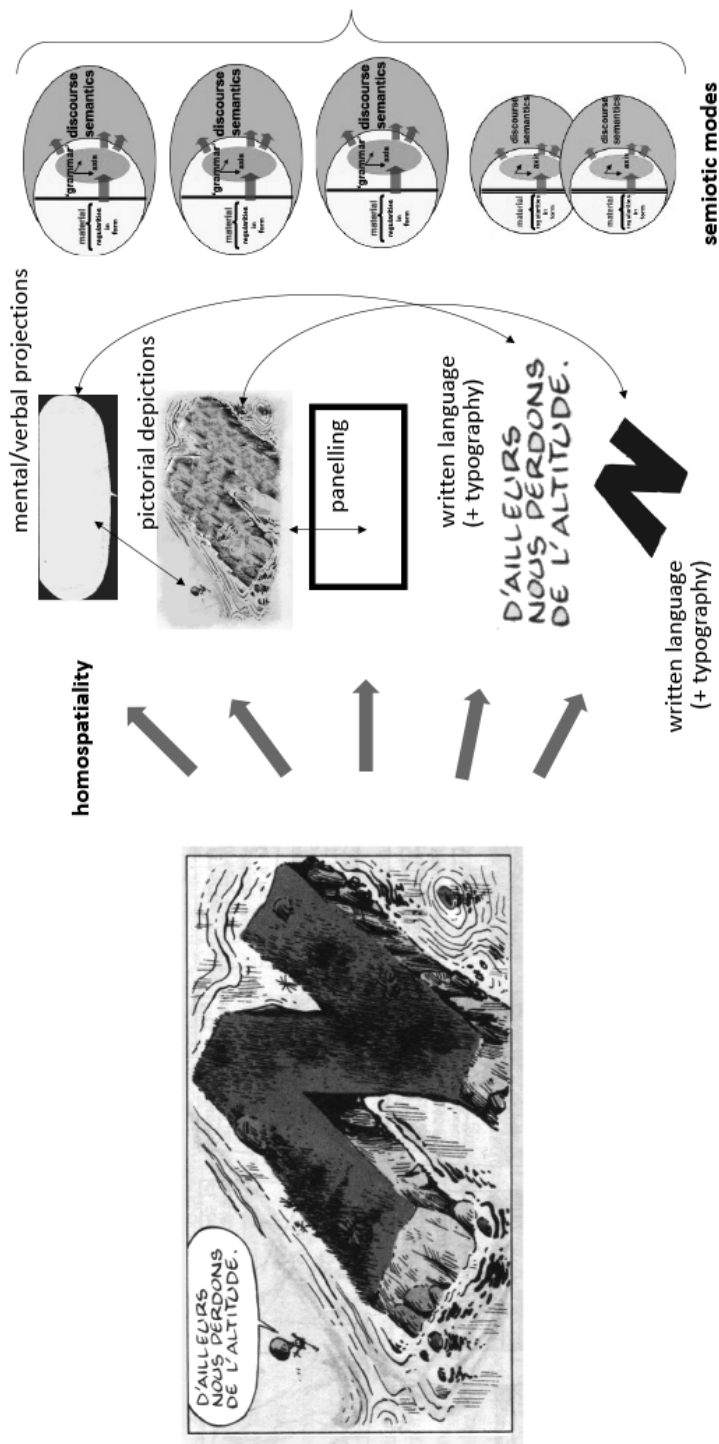
Fig. 2. Left: *Philémon: L'Intégral.* Fred (2011: 152). Right: decomposition according to contributing semiotic modes.

sion, i.e., that distinct semiotic modes are in play, that the panel has the aesthetic effect and appeal that it does.

The binding of the distinct material contributions is characterised as a discourse semantic blend as suggested above. In the present case, one of the semiotic modes provides a discourse entity linked back through its material to the shape of a letter 'N' from the written alphabet, while another of the semiotic modes provides a discourse entity linked back through its material to a pictorial representation of an island. The blend brings these two discourse entities 'together' as a merged element having specifically selected properties of both. Kutz et al. (2015) give further formal details and examples of how this mechanism operates. Particularly interesting here are the diverse scales at which the introduced discourse elements may play roles in interpretation; many of these are 'non-structural' in the sense of multimodal cohesion (Tseng 2013) or braiding (Groensteen 2007 [1999]), to which we return below, but remain nevertheless 'guided' by discourse development (textuality).

To begin this development, within the single panel as shown, the possibilities for resolving the properties introduced by the blend are very limited, but this in itself is an aesthetic feature: either the island just happens to be shaped like an 'N' or someone has formed it in that way for some, as yet, unknown reason. Both are essentially 'diegetic' in that the observed property is anchored within the storyworld. Already, however, discourse hypotheses may be formed that will remain pending until further evidence is gathered and so these stretch beyond the confines of the single panel – for example, letters of the alphabet have a defined sequentiality and so there may be both 'preceding' and 'following' islands shaped like 'M' and 'O'; or perhaps all islands in this storyworld are shaped like 'N', and so on. Regardless of specific hypotheses, however, one is dealing properly with a textually-cued blend in that the resulting discourse elements have properties imported both from the realm of islands and from that of written language. The final resolution of the puzzle – that is, the discourse hypothesis that offers the most coherent characterisation of the information provided – is in this case only reached after considerable further input.

Homospatiality of this kind is by no means new, of course, as early illuminated manuscripts from the 12th century onwards document. Elliot (2003) discusses a further related variety of homospatiality in 19th century novels and its critical reception. But none of these really q u e s t i o n the 'word-image' division; for this, one would need to go back considerably further in history to times and cultures where, arguably, written language and pictorial representations had not yet separated as semiotic systems (cf., e.g., Kammerzell 2009; Burdick 2010). Within the multimodal framework adopted here even situations such as these remain unproblematic precisely because there is no presupposition of division prior to beginning analysis. That is, the semiotic modes appearing along the right-hand side edge of figure 2 are not assumed to exist a p r i o r i but must themselves be the results of empirical study of concrete socio-historically anchored semiotic practices. This

methodology for decomposing the semiotic contributions at work in any artefact under analysis is described in detail in Bateman et al. (2017) and applies to all media, although comics and graphic novels present a host of particularly interesting cases. These show well the benefits of adopting a more systematic approach – as the examples we discuss below will demonstrate.

4. Discourse semantics: the missing ingredient of former semiotic approaches

We have seen that one of the primary reasons that literary-hermeneutic approaches to comics and graphic novels consider linguistics, and often semiotic approaches in general, to be inadequate is an idea that semiotics conceptualises meaning in terms of rigid codes relating Saussurean signifier-signified pairs or Hjelmslevian form-expression ensembles. Semiotic systems of these kinds are generally insufficient for characterising communication, however, precisely because of their lack of provision for interpretation. The semiotic mode construct introduced in the previous section addresses this concern by explicitly complementing the relating of material regularities to qualitative categories with the further inferential component now labelled discourse semantics. This is the concrete analytic mechanism by which the "clues and gaps" and "readers' inferences" called upon by Kukkonen (2013: 50) and others are incorporated explicitly into the model.

Crucially, however, just as with the other levels of semiotic abstraction drawn upon by the model, quite specific properties and mechanisms are defined for discourse semantics and these are central for all discourse interpretation:

> Discourse cannot be understood without paying attention to the inferences that readers, hearers and viewers must perform; but these inferences may well also need to draw on more discourse-specific kinds of organisation that need description in their own right (Bateman 2014: 206).

Although "discourse" is another one of those terms that are multiply defined, sometimes in quite incompatible ways, the definition used here is strictly that indicated above, i.e., as the third level of abstraction within a semiotic mode that is responsible for linking characterisations of forms, marks, technical features of some semiotic mode with interpretations that serve as the basis for further inferences concerning what forms are communicating.

Discourse semantics makes it far more straightforward to relate the multimodal semiotic model to questions raised in areas such as narrative studies or transmedial narratology. As Steiner (2004) discusses in some detail, the basic challenge in using static visual materials for narrative purposes is that one needs to articulate episodes over which a story can unfold: i.e., temporal relations (at least) need to be signalled in a material canvas which does not support dynamic traces (cf. Bateman 2021);

Smith (2015) places some important aspects of this development in an interesting broader historical and transmedial context. Artists over the centuries have risen to this challenge in a variety of ways – restrictions in material possibilities rarely prove insurmountable for the communicative uses made of those materials. Thus, forming stories from panels in particular sequences, as taken as the norm in comics and graphic novels, is one logical solution to the task well anchored in a variety of historical forebears. But this is just one of several possibilities – and, as often noted in more literary discussions, other possibilities are readily found e v e n i n c o m i c s a n d g r a p h i c n o v e l s (cf. Horstkotte 2015). This should be seen as an important counterbalance to any focus on physical sequentiality to the exclusion of other semiotic techniques. From the multimodal semiotics perspective, sequentiality is just one of the many ways available to semiotic modes for organising their material traces or marks to support interpretation. Some semiotic modes make use of this possibility, while others do not.

Within this multimodal semiotic view, most concepts of narratological interest, such as points of view, focalisation, and so on, are necessarily placed as contributions to discourse semantics. This follows directly from their generally being matters of interpretation: they cannot be directly 'read off' of combinations of material clues. As, for example, Horstkotte and Pedri clearly recognise:

> focalization operates at the discourse level, since it is here that textual signals cue the reader to reconstruct the storyworld under the aspectuality of a specific fictional mind (Horstkotte and Pedri 2011: 335).

In fact, this is equally true for the overarching term "narrative" itself, which, following Wolf (2003), is also to be seen as a matter of degree following from the interpretation of cues locatable within some object of analysis. Properties such as narrative focalisation are then c o h e r e n c e - c r e a t - i n g d e v i c e s – that is, the assumption of an interpretation in terms of some form of focalisation rather than another m a k e s s e n s e of the clues found in a 'text' by formally capturing how they contribute to a coherent reading. They may, given further clues, be found not to have been the best choice, and authors can work with this uncertainty deliberately to provide conflicting guidance for interpretation. All of these kinds of operation are typical for dynamically unfolding discourse semantics but are not generally relevant when addressing syntax.

Our notion of discourse semantics has nevertheless sometimes been criticised in its application to comics and graphic novels from the empirical-analytic side. Most specifically, the early description of the application of discourse semantics offered by Bateman and Wildfeuer (2014) is critiqued at length by Cohn (2018) on the grounds that the notion is not helpful and that his own account in terms of 'grammar' is both preferable theoretically and more empirically motivated. Setting out the details of this dis-

cussion would be beyond the scope of the present paper, particularly because Cohn's view of 'grammar' has very different properties to those generally assumed in the critiques of the notion of 'grammar' reported on above and is, in any case, already far more oriented to 'discourse' than critiques of Cohn often assume:

> the combination of images may be closer to the structure used between whole sentences: a *narrative structure*. Indeed, the structure used to understand sequential images may be the same as that for understanding sequences of sentences in discourse and sequences of shots in film (Cohn 2013: 65).

Cohn's insistence on the presence of 'grammar' does not then require that certain sequences of panels be ruled out as 'ungrammatical' in any naïve syntactic sense. For Cohn, it is entirely sufficient for there to be differences in the ease with which readers can find interpretations to qualify as evidence for a 'grammar' being at work. As a consequence, he considers a suggestion made by Bateman and Wildfeuer (2014) that a level of 'grammatical' description (in Cohn's sense) is unnecessary for comics and graphic novels to be incorrect.

Cohn's argument offers an important corrective in that Bateman and Wildfeuer's suggestion may well have gone too far. One result of our earlier investigations into a broad range of media deploying visual and verbal forms of expression had been to observe the tendency that the verbal semiotic modes have highly developed lexicogrammatical semiotic strata while visual semiotic modes tend to rely far more on discourse semantics to operate; this is the basis of the intuition voiced by researchers above concerning a lack of 'rules' and also aligns suggestively with Wittenberg and Jackendoff's (forthcoming) recent proposals for a trade-off between the communicative work taken on by pragmatics and by syntax depending on the formal complexity of the grammatical systems available. Bateman and Wildfeuer's (2014) position was then to take this general tendency and to assert it more categorically for the case of comics and graphic novels, thereby excluding potential contributions from a stratum of grammar.

Strictly speaking, however, such a position is not compatible with our own model of semiotic modes as introduced above and, indeed, as should have been employed throughout Bateman and Wildfeuer (2014) as well. As we have seen, this model insists on b o t h a discourse semantics level of organisational principles a n d a stratum of 'lexicogrammatical' configurations for e a c h semiotic mode. The task of the former is to provide constraints on interpretation guided dynamically by textuality, while the task of the latter is to capture structural constraints on well-formedness and to provide the basis for compositional semantics (cf. Bateman and Wildfeuer 2014: 186); both can be specified by means of 'rules', although those rules are seen as having quite distinct formal properties at the two strata.

Cohn's studies in a variety of experimental settings have now gathered considerable empirical evidence that grammar-like configurations play an

important role in the comprehension of comics and graphic novels and so their exclusion in Bateman and Wildfeuer (2014) was premature. Indeed, from a broader semiotic perspective, this should not be surprising. As known from a long tradition of studies of language change, it is common for patterns of use anchored in discourse to become 'solidified', 'grammaticalised', or 'entrenched' (each from a different perspective) so as to constitute g r a m - m a t i c a l formations and lexicalised units. This is a fundamental component of usage-based approaches to language (cf., e.g., Tomasello 2005) and broadly construction-oriented accounts (Goldberg 1995; Jackendoff 2002), to which Cohn explicitly subscribes. Given the extensive use of visual sequences for communication over the last century and more, it would then be rather unlikely for grammaticalisation processes not to have occurred in corresponding media.

Nevertheless, even though there are suggestive similarities with respect to the role of 'discourse semantics' in our account and that of 'narrative grammar' in Cohn's account, the recognition of a lexicogrammatical stratum of organisation for visual sequences does not obviate the need for a discourse semantics stratum within this semiotic mode as well. Specifically, even though both approaches establish a strong orientation to discourse, the means by which discourse is modelled remain quite distinct.

On the one hand, Cohn's model builds on earlier (broadly cognitive) linguistic work such as that of Jackendoff (1990) and Langacker (2001), in which discourse comprehension, when addressed at all, was modelled primarily in terms of the incremental growth of conceptual structures, 'mental models', or situation models as consecutive utterances bring new semantic information to bear. Cohn's account goes substantially further by (i) having his visual narrative grammar predict and interpret particular sequences of elements, and (ii) relating those sequences in parallel to conceptual structures by making use of the general grammatical notion of 'construction', i.e., partially instantiated usage patterns added to a stock of communicative strategies maintained for a resource, to link levels of description. The earlier cognitive models extending conceptual structures incrementally are then argued to be inadequate: "semantic processing alone cannot account for various relations between panels beyond image-to-image juxtapositions" (Cohn 2020: 363). Indeed, the diversity of presentations in narrative contexts, which we would generalise here to include all communicative situations, warrants

> a system separate from meaning to allow such differences in presentation. Such phenomena require more than just monitoring perceptuo-semantic changes (Cohn 2020: 363).

Cohn (2019) then outlines a collection of constructions involving visual narrative grammar sequences that trigger s p e c i f i c inferences for filling in the conceptual models being constructed. This also suggests how more extended narrative sequences might become available to a community of

users by means of conventionalisation, offering another important building block in the treatment of discourse.

On the other hand, the discourse semantic stratum of the present multimodal semiotic model focuses more on how contributions to a 'text' may be meaningfully (i.e., coherently) combined even when triggering constructions are not present. As described above, combination in this case operates by means of the attribution of discourse relations, which formally generate discourse structures, which then in turn further constrain subsequent combinations. The selection of particular connections rather than others is formalised within a non-monotonic, i.e., abductive, framework of discourse 'rules' in line with the principles of dynamic semantics (cf., e.g., Kamp 1981; Kamp and Reyle 1993; Asher and Lascarides 1994). This provides direct support for flexible discourse interpretation as a c o h e r e n c e - s e e k i n g m e c h a n i s m ; we suggest below that this view is particularly beneficial when we turn to longer, and more complex, narrative trajectories. As Bateman and Wildfeuer set out:

> A detailed discourse semantics expressed within a dynamic logic identifies structurally determined gaps in knowledge of very specific kinds that must then be filled abductively from context. Discourse semantic principles then control when and how world knowledge may be accessed in this interpretation process […] discourse semantics thus seeks to characterize in a manner that is multimodally viable more precisely just what kind of 'gaps' are created in a work, how they are created, and how they may be filled (Bateman and Wildfeuer 2014: 185).

It is likely, in any real analysis, that these two aspects of discourse will combine: that is, there will be both particular discourse expectations triggered by conventionalised patterns and general coherence-seeking inferences that abductively construct over-arching discourse structures – i.e., 'making sense' of what is on offer. In both cases, the processes of interpretation are to be seen as highly guided and constrained, and so are far removed from allowing 'general' inference as, for example, might appear relevant to an analyst in some particular case; interpretation is always tightly tied to the formal details of the objects of analysis.

For present purposes, however, we see the different formal mechanisms involved in the two accounts as good reasons for locating them at distinct semiotic strata within a single semiotic mode of sequential static images being used for narrative effect. The representational level of Cohn's visual narrative grammar is described as operating "using similar architectural principles as a syntactic structure in sentences" (Cohn 2020: 363), albeit at a higher level of abstraction, but involving unification of partial structures as a primary mechanism; in the multimodal semiotics model, these are mechanisms typical of lexicogrammatical strata. In contrast, the representational level of discourse semantics involves abduction and non-monotonic reasoning over structures, to systematically add further information not present in the starting materials. The use of discourse semantics as an

opportunity for aesthetic design and interpretation manipulation may then also receive explicit representation in its own right. Further work is clearly required both empirically and theoretically to see how these views may complement one another.

But, to move on, we bring these points of discussion together in our next concrete example, applying the notion of discourse semantics to a more complex case involving coherence-seeking interpretation. The example, involving the sequence of panels from Spiegelman's (1991) *Maus* shown in figure 3, has received attention in the literature from several scholars, including useful discussion by Pedri from the perspective of subjectivities and narrative voice in a series of articles (cf. Horstkotte and Pedri 2011: 340–341; Pedri 2015a: 135, and Pedri 2015b). The segment concerns the attempt of one prisoner in a Nazi concentration camp from the Second World War to convince the German guards that he is not Jewish but actually a German, and so should be set free. According to the conventions that Spiegelman sets up for his graphic novel, Jews are generally depicted graphically as mice and Germans, particularly those actively involved in genocide, as cats.

Fig. 3. Art Spiegelman (1991: 50). *Maus: A Survivor's Tale. II: And here my troubles began.*

The panel of interest for our analysis is the second in the sequence as this uses the material available in the frame in multiple simultaneous ways that together communicate – i.e., guide interpretation to – uncertainty in knowledge. Such examples are also very important for countering the common suggestion that is made concerning the 'concreteness', and hence semiotic limitation, of pictorial representations. In the present case, we show how a more explicit orientation to the w o r k i n g s of discourse – i.e., discourse semantics – help unravel many of the issues that are problematic for less articulated analytic frameworks. Although the interpretation of the sequence is, in all likelihood, not in doubt, a closer

analysis shows how informal proposals for interpretations can be lifted to formally derivable discourse configurations shorn of much empirically untested theoretical baggage.

The first panel is relatively straightforward in that it shows a verbal statement of the depicted figure and deploys a caption box for comments extra-diegetic in relation to the visually depicted scene: "Only they hit him and they laughed". At this point in the narrative, the reader knows well that much of the visual depiction is a visually mediated representation of the narrative being related by Spiegelman's father concerning his actual experiences in the Second World War. The source of the message in the caption is thus discoursally settled and unproblematic.

The second panel is in contrast strikingly complex but, again, by considering what aspects of the shaped material are being 'governed' by which semiotic modes, the interpretation is rendered relatively straightforward. If this were not the case for some account, then that would be evidence that that account is insufficient, since there is probably little doubt in the present case that most readers will come to a similar set of interpretations. The deployed semiotic modes must therefore be giving sufficient cues to guide discourse interpretation quite closely. The task of the account, then, is to show this in operation without needing to make potentially unmotivated assumptions of prior conventionalisation. The resources of pictorial depiction and speech bubbles deployed are unproblematic: the characters involved in the dialogue are shown in the panel and are clearly the sources of the shown verbal content under the assumption of an appropriately supporting semiotic mode – here it is evident that substantial conventionalisation will be at work. But the panel then mobilises further material resources available to pictorial representations concerning their graphic style and, in this case, shading. Semiotically, then, there are not one but two pictorial depictions taking place in the same panel: one in the 'foreground' and one in the 'background'.

Since these do not align naturalistically in any way – i.e., they resist discourse interpretation as a depiction of a single scene – an open discourse interpretation challenge is created. Graphically, the background image of the second panel repeats the image of the previous panel (not only in terms of its content but also including the previous panel's speech balloon, doubly indicating that this is not a naturalistic rendition), but with one difference: the head of the character which was a mouse (indicating that the character was a Jew) is now shown as a cat (indicating that the character was a German). This establishes graphically a clear c o n t r a s t relationship between the previously depicted scene and the current background scene. This is then also, for any reader who might fail to notice, underlined in the utterance of the left-hand character, Spiegelman, who asks "was he REAL-LY a German?", to which Spiegelman's father responds "who knows?".

The discourse resolution of this complex is then sophisticated, although still relatively straightforward when the contributing elements are identified sufficiently clearly. This resolution is summarised in figure 4. Particularly of note are the explicit structural and graphical cues guiding interpretation

embodied in the sequence's design. The graphic novel is unfolding as the Art Spiegelman character's visualisation of the father, Vladek's, experiences. The narration of the father is suggested in the first line of the figure: as readers we generally have no access to this (although recordings have been included in digital versions of the graphic novel). Our access is mediated by the visual composition, shown in the second line of the figure, which employs the semiotic resources of the graphic novel including panels and page layout as indicated above. In the panel at issue, there is a tension reminiscent of what in film and theatre terms would be termed breaking the Fourth Wall since there is a step one level 'up' in the diegesis: i.e., the story reaches out from its being told/shown and addresses its own creation, also sometimes referred to in the context of comics and other forms as 'meta-narration'. Here, the Art Spiegelman character is asking the father what was actually the case so that it can be drawn appropriately – i.e., as a mouse, if the character was 'actually' a Jew or as a cat, if the character was 'actually' German.

The visualisation and the response of the father then place these categories and their clear distinction of

Fig. 4. Discourse relations among the components of the panels from Art Spiegelman's (1991: 50). *Maus: A Survivor's Tale. II: And here my troubles began*.

Jews and Germans in doubt. Their questionable status is expressed graphically by backgrounding a depiction of that alternative (by placing it literally in the background of a panel and shading it). Placing it as a panel by itself would have made the precise discourse relation of contrast between it and the preceding depiction difficult to uncover. There would still be a clear contrast, but the motivation for that contrast would be lost. Instead, the relation is cued explicitly and unambiguously by breaking the visualisation and turning to the actual process of Art Spiegelman's interpreting the ongoing tale of the father. This different level of diegesis could also have been expressed visually in its own 'panel' as shown in the figure, but leaving it separate in this way would in turn require considerable retroactive discourse work to make sense of the preceding contrasting depictions of the same character as two distinct animals in two distinct panels. The graphical version of the German-variant could also have been simply omitted at that point turning directly to the 'content'-question, although this would have broken the story unfolding visually quite abruptly, changing diegetic levels.

The adopted solution of superimposing the alternative view and the questions raised about those views as layers of a single panel is then more elegant by far. It maintains the unfolding discourse of the father's story being narrated, but raises the question about the actual identity of the character doubly – both in terms of what might have 'actually' been the case and in terms of the decisions that need to be made by the illustrator to render that story graphically. As a further call back to the often assumed concreteness of visuals, the illustrator is construed here as having to decide just which categorisation applies in order to know which visualisation to employ. This is emphasised in its own right by virtue of the fact that the alternative is shown not as a simple illustration of the situation but as a r e p e t i t i o n   o f   t h e   p r e v i o u s   p a n e l, complete with speech balloon. The father's rejection of a neat resolution to the issue is then maintained graphically as both visual renditions are offered to the reader. Thus, uncertainty itself is depicted graphically as well as verbally in a mutually supportive fashion. The contrasting depictions of the character are not resolved, but stand as open questions. The only discourse relation that is resolved is that between the foreground characters, the 'now' of the unfolding events, and the depicted contrast that visualises how the scenes would have been shown given either one resolution or the other.

What we can see from this analysis is that many of the issues that, for example, Horstkotte and Pedri (2011) raise in their close narratological analysis of the sequence are also identified – but without needing to rely on the extensive theoretical superstructure within which they couch their description. Instead, the questions raised are shown as more specifically constructed by the graphic design itself. It is for this reason that the interpretation offered for the sequence is not controversial; one can go on and raise issues of focalisation, subjectivities, narrators and so on – but these, in the last resort, will be d e p e n d e n t   o n the direct discourse interpretations established by discourse semantics and the requirement of dis-

course semantics to pursue discourse relations that maximise coherence. We will see this point being made with even more force in our re-analysis of some examples from Moore and Gibbons's (1986–1987) *Watchmen* in the following section.

5.   Towards treatments of more extended visual narration

To show how all the discussion points raised so far productively combine, we turn finally to particular examples discussed by Horstkotte (2015) concerning some more extended graphic narratives. The analyses Horstkotte offers are, on the one hand, sufficiently precise that we can engage with them in detail, while, on the other hand, showing many of the problems that remain when the semiotic foundations demanded for close analysis of media are not drawn upon. Although there are other comics and graphic novels scholars whose analyses we could draw on, we select Horstkotte as an illustrative case because of the combination of detail with which her analyses are presented on the one hand, and the continuing (cf. Horstkotte and Pedri 2022) lack of contact exhibited with broader multimodal semiotic research methods on the other. Several positions, including that of Horstkotte, are compared and contrasted in similar terms in the broader context of German comics research in Wildfeuer and Bateman (2016). Above we mentioned how Horstkotte explicitly rejects certain aspects of more formal, more semiotic approaches on the assumed grounds that these restrict attention to simple questions of linearity; now we will see both that this is not the case and that the very lack of a semiotic frame for couching analyses far too readily lends credence to over-generalised statements that distract rather than elucidate. Our proposal will be that even the kinds of interpretative directions that Horstkotte wishes to pursue can be made significantly s t r o n g e r when suitably grounded in an appropriate semiotics of the medium. This can then be seen as a general message for any wishing to engage more thoroughly with the complexities of these media.

Horstkotte and others pursuing similar arguments against linearity commonly draw on examples that clearly document that there are interesting and complex relations holding among non-locally sequenced panels and that these relationships need to be grasped in order to understand the narratives being analysed. Sequential interpretations and their semiotic/linguistic models are rejected as missing such phenomena. More specifically, Horstkotte observes that each panel is "part of a sequence, narrating a self-contained sequence of events, unified by stylistic choices" (Horstkotte 2015: 45) and that a variety of scales of sequence may be relevant. Each sequence then

> has to be read in the larger context of a narrative and has to be interpreted with reference to its narratorial origin and its perspectivisation (Horstkotte 2015: 45).

This is no doubt true but provides little motivation for reducing the role of sequentiality. Indeed, turning away from semiotic analysis undermines the very tools that allow analysis to proceed most effectively.

In order to achieve some analytic hold on the phenomena being analysed, Horstkotte instead draws extensively on Groensteen's (2007 [1999]) notion of 'braiding'. Braiding allows for several distinct kinds of relationships among elements on the pages of comics and graphic novels. These include sequential relationships but also open up the possibility of distant relations suggested by graphical or other properties of design. However, much of what Horstkotte characterises in terms of braiding is precisely what a discourse semantic interpretation provides, with the difference that the discourse semantics is sufficiently closely linked with material distinctions as to show how interpretation is generally strongly guided by design – something the importance of which Horstkotte's rejection of semiotics serves to minimise. This limits what can be done, the questions that can be asked, and the methods available for pursuing investigation. Restricting the range of constraints that can be explicitly drawn upon also plays directly into Cohn's main critique of braiding, that of allowing apparently arbitrary connections to be drawn. As Cohn points out:

> An average 24-page monthly comic book with six panels per page would have 144 total panels, yielding 10,296 possible panel relationships! […] Without some sort of system to constrain these relations (i.e., a grammar!) keeping track of all these connections between panels (whether they are 'active' or not) would overwhelm human working memory (Cohn 2014: 68).

Seen in this light, accounts of braiding that fail to explicitly specify conditions under which it is sensible to seek relationships make it difficult, if not impossible, to derive predictions. Indeed, leaving open just which panels may need to be placed in relationships to others requires, in principle, that all and any may be considered, which flies in the face of what is known about human cognitive processing.

Horstkotte's examples certainly succeed in showing that relationships are built up across panels and elements that are not immediately sequential, but does not provide a systematic framework that might restrict interpreters in their task so as to stay within plausible limits – limits which are then themselves commonly relied upon for effective aesthetic design. In contrast, Horstkotte maintains her strong position against the relevance of semiotic treatments on the grounds that this reflects a particular 'dogma' concerning the centrality of 'sequence' and 'grammar' that is inappropriate:

> Despite these infinite choices, one of the most repeated dogmas of comics studies is the understanding of comics as a linear or 'sequential art' with a 'grammar' composed of panels and frames separated by gaps and gutters (Horstkotte 2015: 33).

As we saw above, and as argued at length by Cohn (2014), the consideration of 'grammar' here is multiply problematic. But the most important point for the current discussion is the fact that sequentiality does play important roles in making sense of any graphic narrative. There is no need, however, to allow this to exclude other contributions to interpretation. Conversely, the existence of other contributions to interpretation cannot be taken as an argument that sequentiality (of kinds still to be defined more precisely) does not play a role. As noted above, whether or not sequentiality plays a role is a property of the (lexicogrammatical stratum of the) semiotic modes being considered. Several of the semiotic modes relevant for the media of comics and graphic novels use this property, several do not.

Horstkotte's discussion then appears to conflate sequentiality as a s e m i - o t i c property, i.e., one of the conditions for deciding which 'marks' are relevant for interpretation at all, and the accompanying b e h a v i o u r s by which users of semiotic modes go about creating their interpretations. Opening up the range of relations into which panels may be placed during interpretation should never be considered the same as stating that reading becomes an issue of 'roaming' across what is being presented as if this were a subjective mystery – one can always pursue different aspects of a graphically presented static artefact, but interpretation is generally far more constrained. Constrained means not that behaviour is forced – it is always possible, after all, to read the last page of a novel before the first if one wishes – but rather that particular courses of engagement with an artefact will be more or less strongly cued by design and rendered 'pertinent' by the semiotic modes at work. Such principles of design and their effects can be made the object of focused research and generally lead to perhaps surprising degrees of regularity. In contrast, Horstkotte's account illustrates rather clearly how researchers can come to over-value the specificity and supposed 'uniqueness' of cases when the more general semiotic principles being deployed remain unclear.

The example we focus on to conclude this discussion is Horstkotte's treatment of a key episode from Moore and Gibbons's (1986–1987) *Watchmen*. This episode, involving the murder of 'The Comedian', one of the story's main protagonists, is presented at three quite distinct points in the story: first, at the beginning, when it is shown along with an initial police investigation of the crime scene involving two detectives; next, in the second chapter, when it, at first glance, appears to accompany thoughts of another character during the Comedian's funeral; and finally, towards the end, when it accompanies explanations being offered by the protagonist that actually committed the murder. Horstkotte anchors her approach along the lines explained above by framing the discussion as follows:

> The dynamic interaction between the visual and verbal has to be studied in the context of the entire graphic novel since it is only here that panels gain their full meaning. Although the sequence is almost identical each time it occurs, the variation in context – i.e., the alternating panels of the first narrative – and the refer-

ence of all three occurrences to each other as well as the minute variations of the
series and their various combinations with different verbal tracks call for a more
layered account than a linear understanding of sequentiality is able to provide
(Horstkotte 2015: 43).

To state that panels only gain their 'full meaning' in the context of the 'entire
graphic novel' may be true but is unhelpful. It is also somewhat ingenuous
since we are generally concerned with many steps and stages in interpre-
tation prior to some 'full meaning' – even assuming that any such state of
knowledge is achievable – and it is, again, generally the case that it is pre-
cisely these earlier stages that shape any putative full meanings that might
be proposed.

Methodologically, therefore, such statements need to be treated with
considerable caution. Whereas there are media, particularly those employ-
ing pictorial semiotic modes, where it may not be possible to assign con-
text-free interpretations to individual formal 'items', such as lines, brush-
strokes, shapes, etc., without anchoring them within a whole, this is rarely
the case for those more complex units carrying the broad narrative sweep
of comics and graphic novels at higher levels of abstraction. The semiotic
modes carrying the narrative operate quite differently in that far more inter-
pretative guidance is designed into what is shown. Moreover, even when
such strategies are deployed in a narrative, this is nothing particularly spe-
cific to comics or graphic novels: an artefact in almost any medium may
provide information in a manner that requires additional information gleaned
from elsewhere in the artefact for an adequate interpretation to be achieved.
This is because the ability to work in this way is a function of the mecha-
nisms of discourse semantics, where interpretation is a l w a y s a directed
process of making hypotheses on the basis of given evidence to increase
perceived coherence. There is nothing to stop that evidence being distrib-
uted broadly across an artefact.

Even when distributed, however, there will often nevertheless be a host
of guiding clues supporting interpretation: this is the defining property of
textuality. And, as we will now see, the most predominant semiotic organi-
sation in comics and graphic novels remains strongly sequential in organ-
isation. The most damaging aspect of Horstkotte's treatment is then the
suggestion that a 'linear understanding of sequentiality' is in some way
incompatible with a 'more layered account'. Indeed, Horstkotte herself states
that sequentiality is one level of several, only to reiterate the inapplicability
of semiotic and linguistic accounts, i.e., precisely the source of theoretical
frameworks that can supply the most detailed models of how sequentiality
operates i n r e l a t i o n t o o t h e r semiotic strategies. A more layered
account in fact d e m a n d s a fuller understanding of the semiotic workings
of sequentiality in order to move beyond conjectural analyses and to pro-
vide sufficient focus to analyses that they may be interrogated empirically.

We now take up the three repeated depictions of the murder episode
in *Watchmen* in detail. In each repetition, panels showing the murder are

interleaved with panels depicting events 'ongoing' at that point in the story. Moreover, in all three instances, the panels showing the murder and preceding fight are accompanied by captions anchored in the unfolding 'ongoing' story rather than the murder-scene. There are often cross-references across the verbal and the visual of a non-referential, suggestive nature but, despite these disjunctions, both kinds of tracks generally run linearly and independently. Nevertheless, this use of sequentiality in no way speaks against, or dilutes, the additional g r a p h i c a l use of the page composition as contributing to the depiction of alternations across time-space locales.

We see this clearly in the graphical design of the first occurrence of the murder scene, shown in figure 5 for reference, where the alternation is constructed visually by means of red panels alternating with more green-blue-orange panels. This establishes non-sequential compositional support for pursuing what is clearly a sequence-based reading. One might profitably explore to what extent perception of an alternation would become harder if this visual design cue were removed. Alternation itself, however, is already a structure highly dependent on linearity, which means that Horstkotte's warnings against attention to linearity are quite out of place. Moreover, alternation is a common structure found in several media, comics and film included, and is typically associated with a limited range of discourse functions, including temporal simultaneity (as in chase scenes, telephone calls, etc.), memory, and in-story (i.e., embedded) narration (cf. Bateman 2013). Thus, although we do not have single 'set meanings', we certainly are guided strongly to particular lines of interpretation rather than others. In short, the page design signals graphically that the reader needs to consider the two sets of events in relationship to one another in some way – that is, a d i s - c o u r s e   r e l a t i o n must be selected from those provided by the semiotic mode in order to maximise the coherence of the observed materials.

The linearity of such alternations also in no way restricts narration to temporal linearity or even single storyworld lines – quite the contrary in fact: alternation is one of the prominent resources of linearly organised semiotic modes that massively expands those modes' expressive potential. It is also a natural development commonly used as a structuring device in many media, probably most prominently in music. The existence of such patterns makes the case for a proper treatment of sequentiality even stronger, for without it analyses lack a primary resource for pursuing how such media strategies are being used.

In the present case, the sequence begins as if a discourse hypothesis of mental 'projection' between the main storyline and the alternated murder scene could apply — that is, the images would show the detectives' imagination of what had occurred. This would then mean that alternating murder scene panels should 'illustrate' the detectives' dialogue. But the details given of the murder visually do not in fact always quite match the detectives' musings sufficiently to anchor this hypothesis as clearly being the most coherent. The option of an independent story-contribution, such as, for example,

Fig. 5. *Watchmen* (Moore and Gibbons 1986–1987, Chapter I: 2–4).

a flashback, therefore remains open, as do several other epistemic sources. Importantly, the basic operation of discourse semantics of seeking discourse relations to raise overall coherence does not guarantee that unique solutions will be found and the process of discourse relation resolution itself is often by no means straightforward. Indeed, it is more often precisely the challenge of discovering coherent relations that drives narrative appeal and depth. In short, the alternation means that there is a connection to be drawn, in a manner quite similar to Cohn's (2019) narrative constructions discussed above, but the question of *which* relation precisely is an issue that must be abductively pursued within the discourse semantics.

The g u i d e d   c o m p l e x i t y that such structures create can be seen even more clearly in the second occurrence of the murder scene sequence in *Watchmen*. This sequence is connected, as Horstkotte notes, within the framing episode of the funeral of the murdered protagonist. In fact, employing a fairly standardised narrative device, the funeral is used as a means of introducing aspects of the background story by running through memories and associations concerning the murdered 'Comedian' from several of the characters present. The chapter is appropriately called 'Absent Friends'. In this spirit, Horstkotte proposes that the repeated murder scene is, in this case, connected broadly to memories or reminiscences of a further character, Rorschach. But the structuring throughout this chapter is far more complex than this suggests and so it is worthwhile tracking this closely in order to show both that this alternation is again operating in a manner that invites particular interpretative paths rather than others and that this is strongly constrained by the sequential organisation of the narrative throughout.

To begin it should be noted that alternation constitutes a dominant presentational strategy for this chapter as a whole. The chapter begins with a sequence of no less than 22 panels strictly alternating between the cemetery in New York where the funeral is taking place and another character's, Sally Jupiter's, home in California. The dialogue displayed also alternates between the two locations, but is generally aligned as in the first alternation above to set up a counter-rhythm with parts of the dialogue in California additionally (and non-diegetically) commenting on events in the cemetery. This leads into a first embedded reminiscence concerning the murdered Comedian anchored to Sally Jupiter. The narrative then turns to focus on the funeral and presents memories of three other characters present at the cemetery – Veidt, Dr Manhattan, and Nite Owl.

The first three memories presented are all strictly embedded, i.e., without alternation, and are explicitly introduced visually by clear (i.e., perceptually prominent) graphic 'matches' (cf. Mikkonen 2017: 41, Gavaler 2022: 125) followed by an equally evident switch in time and location (cf. Gavaler 2022: 199). This structure is so established that the fourth memory, that of Nite Owl, can dispense with an opening graphic match altogether, using a close-up of the character followed by a very differently sized panel instead.

At this point the funeral is over but yet another character, Moloch, is shown visually but unnamed leaving the cemetery. This character returns to his apartment, followed by Rorschach, who was also shown previously standing outside the cemetery.

Rorschach then breaks into Moloch's apartment and forces him to divulge any information he may have about The Comedian's murder. This leads to a further strictly embedded sequence (introduced by a close-up panel of Moloch) running over 18 panels. The chapter maintains its overall sense of alternation during those 18 panels by alternating their dominant colour even though they are spatiotemporally continuous. This alternation could therefore also be hypothesised discoursally to be diegetic, for example, as being caused by external lighting. Following Moloch's story, the narrative returns (via a graphic match) to the present and tracks Rorschach walking along some New York streets back to the cemetery. It is only at this point that the second recurrence of the alternated murder sequence appears. In this case, the switch to the murder scene is abrupt without any of the perspectivalisation seen in the previous reminiscences shown in the chapter. Continuity across the alternation is maintained only in the accompanying captions, which are anchored by colour, shape, and content to Rorschach. We can see, therefore, that characterising this repetition as simply alternating with the funeral scene as Horstkotte (2015: 43–44) suggests would not be sufficient (nor accurate).

In fact, given the complexity of alternations unfolding so far in this chapter, it comes as no surprise that this second repetition of the murder scene chooses to raise that complexity still higher by drawing on another very general way in which structural alternation can be involved in the expansion of semiotic potential. Once established, alternations set up a strong 'backbone', or scaffold, that supports further structural development. The expectation that one is within an alternation allows a medium to explore variations and developments within the (linear) confines provided by contributions – i.e., panels in the case of comics and graphic novels – that have 'space' between them for simultaneous variations of lines of narrative development. In the present case, in these 'spaces' created by the alternating backbone of the panels of the murder scene, we do not find returns to the cemetery, but events previously depicted in the chapter from o t h e r characters' reminiscences and memories. These follow the order of the five previously embedded sequences, beginning with a panel from Sally Jupiter's memory, followed by single panels from the memories of Veidt, Dr Manhattan, Nite Owl and, finally, from Moloch's story as told to Rorschach.

The connection between the events depicted and the surrounding narrative is then even more tenuous in this second repetition, despite the fact that the captions being shown throughout the alternation depict Rorschach's evaluations of the sad state of the world. The graphically depicted events become, in this structure, more general illustrations that make clear and motivate some of the views that Rorschach is expressing as

general background information from the world which that character inhabits, rather than specific memories, and so are almost impossible to sensibly attribute to any specific character. Any restriction to issues of who is narrating and whether the events are 'hypothesised', 'subjective', 'aspectual', 'focalised', and so on that are prioritised in narratological readings (cf. Horstkotte and Pedri 2011) consequently turns out to be insufficiently responsive to the complexity of the material under analysis precisely because the vital clues necessary to derive such discourse interpretations are not receiving appropriate attention in their own right.

The overall structure of the second re-occurrence is then as follows. To begin, a short sequence of panels at the cemetery focuses in on the character Rorschach. The alternation then announces itself by showing the first panel of the murder scene as done in the previous occurrence. Again, the overall composition of the page suggests different statuses for some of its panels by clear colour contrasts with the cemetery panels. However, in this case, the panel directly following the first murder-scene panel is not a return to the cemetery but a shift to a completely different event. The reader here has a discourse challenge to resolve: there would be an expectation that an alternation is playing out but it is unclear what that alternation is alternating between. The discourse relations posited here will then depend on when the reader recognises the inserted panel. If the reader does not recognise that panel, then the alternation still remains as a structural (linear) configuration – it is simply less clear what is being said beyond nonspecific hypotheses of 'illustrations' of the comments being made by Rorschach in the captions. When the reader recognises the inserted panel as being repeated, however, discourse interpretations are likely to converge rapidly.

Thus, crucial here is this combination of encouraging recognition of events depicted earlier in the graphic novel and linear construction. The fact that connections need to be drawn across broader stretches of the graphic novel, than immediately consecutive panels, is in many respects analogous to any cohesive device in any medium operating anaphorically. This is presumably what Horstkotte is wishing to focus on in terms of non-sequentiality and subordinating to Groensteen's notion of braiding. But this is to say little more than readers interpret what they are reading. Pointing out the phenomenon is important, but then we need to articulate further how the precise design of comics and graphic novels supports and encourages particular courses of interpretation rather than others: this is fundamental to the workings of any semiotic mode.

In the present case, it is precisely the linear construction of alternations that demands that certain connections be made by virtue of the 'structural slots' created by the alternation. This can, again, not be couched as an argument against the importance of sequentiality or semiotics. Indeed, this also allows us to resolve somewhat the critique made by Cohn above concerning Groensteen's under-constrained notions of braiding: the panels

inserted in this second alternation certainly suggest some operation of braiding because they draw on previous events distributed broadly (and 'non-structurally') over the current chapter. B u t , the 'instructions' to look for those connections are structurally invoked, which means that a far tighter connection is being drawn here than free association. Although this does not always have to be the case since it is certainly possible simply to suggest connections by visual repetition, in the example discussed here there is far more structural work occurring because the 'gaps' to be filled are identified graphically and structurally as an essential component of the narrative's sequential construction.

The importance of taking this sequential contribution more seriously is then shown well in the increased complexity of the second occurrence, suggested graphically in figure 6. In this case, the repeated panels from the murder scene alternate not with a narrative 'now' but with the temporally ordered collection of panels that occurred previously in the funeral chapter. They are not narratively related amongst themselves and so could well be characterised as 'episodic' in Metz's characterisation of syntagmatic configurations in film. Episodic syntagma are where

> there is a general forward progression in time but the elements are selected according to some particular organisational feature (Bateman 2007: 22).

Although very familiar from film, this kind of narrative structuring is clearly specific neither to that medium nor to comics and graphic novels.

The specific sequence here is made up of (i) a meeting of the previous group of 'masked crimefighters', the Minutemen, from the 1940s, (ii) an attempted initial meeting of the current group of superheroes from 1966, (iii) Saigon at the end of the Vietnam war in 1975, (iv) riots in New York City in 1977 and, finally, (v) the scene immediately prior to the murder that took place in the apartment of Moloch. Only the last of these involves Rorschach directly (in that the events depicted in the panel were related to Rorschach by Moloch in Moloch's apartment after the murder); the others are drawn from the reminiscences of quite distinct characters as explained above. Employing terms from Metz, we therefore have an alternation over two tracks: one track (the murder scene) unfolding chronologically as an 'ordinary sequence', and the other track unfolding in temporal order but 'episodically' rather than constructing a simple scene. Only following this alternation on the page is the reader returned to the framing cemetery segment.

It is clear then that these additional episodic panels cannot readily be associated with Rorschach, which weakens considerably any grounds for Horstkotte's and Horstkotte and Pedri's suggestions that it is Rorschach's 'imagination' of the murder event that is framing the second recurrence. This conjecture even leads them to misread the graphical clues:
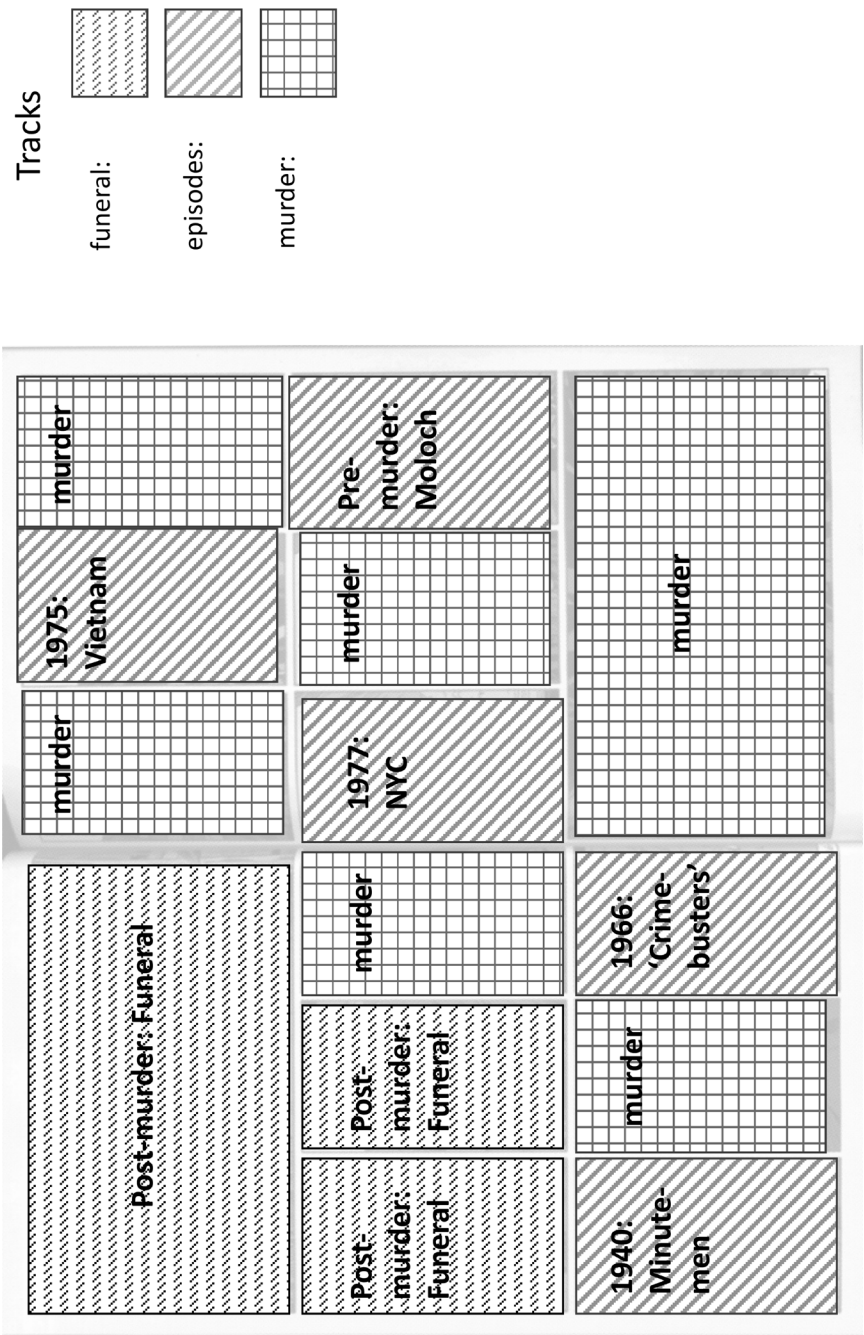
Fig. 6. Alternating track structure of the second murder depiction from *Watchmen* (Moore and Gibbons 1986–1987, Chapter II: 26–27).

> the lower degree of colour distinction between the murder sequence and its alternating frame also emphasises this filtering of information through Rorschach (Horstkotte and Pedri 2011: 346),

and so, it is suggested, the panels

> tell not what is, but constitute themselves as a subjective representation of Rorschach's memories of events in the 1940s (Horstkotte 2015: 43–44).

In rather sharp contrast to this conjecture, we have now seen that the track alternating with the murder scene is in all likelihood not to be associated with Rorschach at all. This then allows us to motivate the distinct colouring far more clearly – i.e., in a manner that increases overall coherence – as a graphical device cuing two further tracks being alternated, not the one that Horskotte and Pedri comment on. This is why there are t h r e e quite distinct levels of colouring; this is emphasised graphically in figure 6.

Moreover, none of these tracks has much to do with the embedding of subjective representations. A far more complex narrative construction is underway, weaving together multiple storylines and subjectivities s t r u c - t u r a l l y, all tightly linked to the graphical details of the design. In contrast, the discussion in Horskotte and Pedri links in a rather undifferentiated fashion close descriptions of details, reasonable hypotheses concerning some potential interpretations that would be well motivated from those details, and far looser speculations concerning possible narrative take-up. These speculations are, as suggested here, sometimes simply inaccurate with respect to how the narrative is structured. Of course, a reader might pursue such flights of fancy but to what extent this should be considered relevant for an analysis of the text is less clear.

It is in the end, then, perhaps quite symptomatic of the marginalisation of the workings of linear structure that the distinct layers of analysis required here appear to be omitted in both Horstkotte and Pedri's (2011) and Horstkotte's (2015) descriptions. While it is true that one needs to consider 'non-local' contributions to the alternations in order to fully place their constituent panels, this only unfolds within the tightly organised linearity of the composition as a whole and is otherwise no more mysterious than the use of pronouns in a verbal text. It is only the s t r u c t u r a l d e t a i l s o f t h e a l t e r - n a t i o n that make these references do substantially more interesting narrative work and which allow any robust statements to be made concerning their 'ontological and epistemological status' (Horstkotte 2015: 45). Taken together, therefore, we can see that the complexity of this narration does indeed require a 'multilayered' analysis, as Horstkotte argues, but one which crucially builds on the sequential configurations in the material rather than more simply assumed connections or assumptions of non-structural 'braiding' loosely described and freely interpreted.

Very similar points, although interestingly different in the details, can be made for the third and final recurrence of the murder scene. In this

sequence, the events of the murder are being related by the character who has now been revealed as the perpetrator, Veidt. The captions within the repeated panels are consequently now quite explicitly given as comments being made by Veidt. An additional panel not seen before occurs, explicitly showing Veidt holding the Comedian above his head, prior to the panel seen before of the Comedian being thrown through a window. The murder sequence is therefore now suggested even more strongly to be an illustration of what actually occurred. The precise alternations at work in this third occurrence take yet a further structural turn, however. Rather than remaining with a straightforward alternation of Veidt's somewhat half-hearted confession of guilt and explanation, on the one hand, and the murder scene, on the other, a third narrative line is woven into the alternation, again giving a three-track structure across these pages. The included murder scene thus looks back towards the beginning of the graphic novel; the first interleaved track anchors events in the story's now; and the third track leads the reader step-by-step to the culmination of the chapter, Veidt's destruction of New York, which, in this case, runs simultaneously (or nearly so) to the second track of the alternation.

This requires careful narrative and structural construction. By the point in the narrative when this third and final repetition of the murder scene is reached, the style of storytelling is so established that connecting the repetition of the murder sequence with Veidt needs further preparation. The captions showing Veidt's comments in relation to the depicted images are either ironic, puns, or strikingly literal throughout, but this continues the style of narrative commentary seen throughout the novel without any need to connect the events depicted with that commentary diegetically. Thus it could become unlikely that readers would see the repetition of the murder sequence as being commented on, or 'focalised', by Veidt at all. The scene would then be a simple flashback to an event not shown in its proper temporal position as preceding the narrative: realising no focalisation, no hypothesis, no memory at all.

To close down some of these potentially relevant lines of interpretation, the narrative itself takes (structural) pains to bring the two tracks more closely together, to ensure, again, particular discourse interpretations rather than others. One of the additional alternating chains running through this chapter is indeed explicitly constructed as Veidt's memories of events that he has been present at over a long period of time leading up to the present moment. This is consequently also available for structuring the segment under discussion and so the intercut repeated panels of the murder are designed to fit precisely into this already ongoing chain. At this point, then, a reader might finally (and is strongly invited to) bring the murder sequence shown into a relationship with Veidt's experiences. This is only really made completely explicit, however, with the new additional panel that joins the sequence in this third repetition: a panel clearly showing (by a graphical match with the preceding panel) that it is Veidt lifting the Comedian prior to

throwing him through the window. This is then a big 'reveal' and the previously climactic exit of the Comedian through the window becomes little more than a release.

Horstkotte and Pedri state for their analyses that they have adopted

> a constructivist focus (and not a rhetorical one) to emphasise the concrete ways in which actual readers respond to textual cues such as focalisation markers (Horstkotte and Pedri 2011: 349),

but no evidence is presented for this claim. The resulting analyses, in fact, demonstrate just how important it is to provide motivated constraints on interpretation that a notion of braiding alone does not (yet) provide. With the discussion of braiding, interpretations remain under-constrained and suggestive. Horstkotte and Pedri propose, for example, that the reduction in size of the final panel showing the Comedian going through the window on its third repetition reflects the subjectivity of Veidt and his lack of concern with human life. While this may or may not be the case, what certainly is the case is that there is no room for a full width panel on that page! The bottom full width is already (spatially) committed to the third alternation track, showing a New York City intersection frequently present in the story, and there is still a further comment from Rorschach and Nite Owl to accommodate so as to close the structural alternation with the murder. The reduced murder panel, now occurring for the third time and simply repeating what is already known, might be expected to have little additional impact, regardless of size. Production constraints of this kind are probably sufficient to motivate the design decision here and so any further interpretations may well be in danger of transgressing the boundaries of fantasy. In contrast, providing detailed structural views is one way of beginning to fill out more detailed proposals for explicit connections and contrasts across panels that can then be examined empirically, motivated in all cases by concrete design decisions in the page composition, but using the semiotic power of a collection of strongly sequentially articulated semiotic modes as guide throughout.

We can thus finally reject the style of analysis given by Horstkotte and Pedri in favour of a far closer treatment of the consequences of linearity, structure, and discourse semantics provided by a multimodal semiotic account. Whereas Horstkotte and Pedri claim:

> if readers fail to ask who focalizes each of the repetitions, then a crucial dimension of the story is lost on them. Focalization is the narrative tool that makes it possible for readers to experience what the storyworld is and feels like, thus ensuring their engagement with it (Horstkotte and Pedri 2011: 349–350).

What in fact appears to describe the take-up of these sequences more effectively is the making explicit of how particular structural configurations drive discourse interpretations. Some of these may lead to interpretations that may be glossed in terms of 'focalisation', others less so. But only when

sufficient attention has been paid to the precise structural organisation created by the various semiotic modes employed can we be confident that we are extracting as much as possible from the material design of the comics and graphic novels so as to be able to do proper justice to their often highly sophisticated organisation.

6.   Conclusion and outlook

The media of comics (and graphic novels) allow the mobilisation of a broad range of semiotic resources that together support the construction of extended narrative (and other) sequences. Consequently, one of the primary aims of this contribution has been to show how a more contemporary and finely articulated account of the semiotics of multimodality can be used to distinguish and relate more effectively the various semiotic domains at work within and across panels in these media. The units out of which such sequences emerge exhibit considerable variation, not only in their constitution (drawn images, written language occurring in diverse roles and forms, often conventionalised visual marks of movement, connection, affect, and so on), but also in their extent (patterned configurations over portions of panels, panels, collections of panels, entire pages, and so on). Although often referred to in more interpretative analyses of specific cases, ways of engaging with this diversity in more formal contexts such as those required for empirical work have been limited. Similarly limited have been approaches capable of explaining how c o h e r e n c e can emerge in the face of that diversity.

   Whereas several accounts adopting more textual or literary approaches note that coherence is a matter of interaction between a 'text' and the cognitive processing of a reader (cf. Saraceni 2016), this is only to circumscribe quite broadly something which already follows directly from any reasonable communicative account with a semiotic foundation. An interpretation cannot be 'in' a text as this would violate basic Peircean understandings of semiosis (i.e., wrongly attributing Thirdness to Secondness, i.e., what is 'present'); this is also the traditional distinction drawn, for example, between cohesion and coherence in text linguistics (cf. Halliday and Hasan 1976; de Beaugrande and Dressler 1981). What is then crucial is to characterise how this interaction itself is to be modelled, particularly paying sufficient attention to the extreme multimodality of the 'marks' deployed. As a generalised solution to this, we have proposed in this article a thorough reorientation to semiotic modes as defined within the theory of multimodal semiotics adopted. Through several examples, we have shown how a more semiotically-based account of this kind provides much needed guidance for organising our analytic access to highly complex graphic materials such as comics and graphic novels. In particular, tracking the precise development of the discourse structures called for by textual design forces materials to be incorporated in analysis in ways maximally supportive of their interpretation.

Current models of multimodality therefore provide both a broader foundation for relating very different kinds of semiotic resources and one which is open to a broad range of inputs, from narrowly empirically motivated to more hermeneutic, discourse-interpretative proposals for what is going on. Moreover, methodological principles for supporting the move from the interpretative to the more empirically supportable are an essential part of the account. Only with such a broad methodological basis is it likely that the study of rich and complex media such as comics and graphic novels can grow to address questions of relevance for the entire community of concerned scholars, while still remaining firmly anchored in the material details of the artefacts being analysed.

Notes

1   This limitation of the scope of semiotics has a long history of its own. Indeed, it is relevant to note how this kind of argument runs through engagements with several 'visual' media – perhaps most well articulated in accounts of film, where early semiotically-derived proposals from Metz (1974) led to broad-brush rejections that prevented productive exploration for several years; Bateman (2007) discusses this development in some detail. Mitchell (1992) has also raised similar concerns with respect to art history and the treatment of static visuals.

Bibliography

Asher, Nicholas and Alex Lascarides (1994). Intentions and information in discourse. In: James Pustejovsky (ed.). *32nd. Annual Meeting of the Association for Computational Linguistics*. San Francisco: Morgan Kaufmann, 34–41.

Asher, Nicholas and Alex Lascarides (2003). *Logics of Conversation*. Cambridge: Cambridge University Press.

Bateman, John A. (2007). Towards a *grande paradigmatique* of film: Christian Metz reloaded. *Semiotica* 167, 1, 13–64.

Bateman, John A. (2013). Multimodal analysis of film within the GeM framework. *Revista Ilha do Desterro: a journal of English language, literature in English and cultural studies* 64. Florianópolis: Universidade Federal de Santa Catarina, 49–84.

Bateman, John A. (2014). *Text and Image: A Critical Introduction to the Visual/Verbal Divide*. London and New York: Routledge.

Bateman, John A. (2016). Methodological and theoretical issues for the empirical investigation of multimodality. In: Nina-Maria Klug and Hartmut Stöckl (eds.). *Handbuch Sprache im Multimodalen Kontext* 7. (HSW 1). Berlin: De Gruyter Mouton, 36–74.

Bateman, John A. (2018). Peircean semiotics and multimodality: Towards a new synthesis. *Multimodal Communication* 7, 1.

Bateman, John A. (2019). Transmediality and the end of disembodied semiotics. *International Journal of Semiotics and Visual Rhetoric* 3, 2, 1–23.

Bateman, John A. (2020). The foundational role of discourse semantics beyond language. In: Michele Zappavigna and Shoshana Dreyfus (eds.). *Discourses of Hope and Reconciliation. On J. R. Martin's Contribution to Systemic Functional Linguistics*. London and New York: Bloomsbury, 39–55.

Bateman, John A. (2021). Dimensions of materiality: Towards an external language of description for empirical multimodality research. In: Jana Pflaeging, Janina Wildfeuer and John A. Bateman (eds.). *Empirical Multimodality Research: Methods, Evaluations, Implications*. Berlin: De Gruyter, 35–64.

Bateman, John A. (2022a). Growing theory for practice: empirical multimodality beyond the case study. *Multimodal Communication* 11, 1, 63–74.

Bateman, John A. (2022b). Multimodality, where next? – some meta-methodological considerations. *Multimodality & Society* 2, 1, 41–63.

Bateman, John A. and Janina Wildfeuer (2014). A multimodal discourse theory of visual narrative. *Journal of Pragmatics* 74, 180–218.

Bateman, John A., Janina Wildfeuer and Tuomo Hiippala (2017). *Multimodality – Foundations, Research and Analysis. A Problem-Oriented Introduction*. Berlin: De Gruyter Mouton.

Bucher, Hans-Jürgen (2011). Multimodales Verstehen oder Rezeption als Interaktion. Theoretische und empirische Grundlagen einer systematischen Analyse der Multimodalität. In: Hajo Diekmannshenke, Michael Klemm and Hartmut Stöckl (eds.). *Bildlinguistik. Theorien – Methoden – Fallbeispiele*. Berlin: Schmidt, 123–156.

Burdick, Catherine E. (2010). *Text and Image in Classic Maya Sculpture: A.D. 600–900*. [PhD thesis, Art History]. Chicago, IL: University of Illinois at Chicago.

Chomsky, Noam (1957). *Syntactic Structure*s. The Hague: Mouton.

Cohn, Neil (2013). *The Visual Language of Comics: Introduction to the Structure and Cognition of Sequential Images*. London: Bloomsbury.

Cohn, Neil (2014). Building a better 'comic theory': Shortcomings of theoretical research on comics and how to overcome them. *Studies in Comics* 5, 1, 57–75.

Cohn, Neil (2016). A multimodal parallel architecture: A cognitive framework for multimodal interactions. *Cognition* 146, 304–323.

Cohn, Neil (2018). In defense of a "grammar" in the visual language of comics. *Journal of Pragmatics* 127, 1–19.

Cohn, Neil (2019). Being explicit about the implicit: inference generation techniques in visual narrative. *Language and Cognitio*n 11, 1, 66–97.

Cohn, Neil (2020). Your Brain on comics: A cognitive model of visual narrative comprehension. *Topics in Cognitive Science* 12, 1, 352–386.

De Beaugrande, Robert and Wolfgang U. Dressler (1981). *Introduction to Text Linguistics*. London: Longman.

Eco, Umberto (1976). *A Theory of Semiotics*. Bloomington: Indiana University Press.

Eisner, Will (2008). *Comics and Sequential Art: Principles and Practices from the Legendary Cartoonist*. New York: Norton.

Elleström, Lars (ed.) (2010). *Media Borders, Multimodality and Intermediality*. Basingstoke: Palgrave Macmillan.

Elliot, Kamilla (2003). *Rethinking the Novel/Film Debate*. Cambridge: Cambridge University Press.

Frahm, Ole (2010). *Die Sprache des Comics*. Hamburg: Philo Fine Arts.

Gavaler, Chris (2022). *The Comics Form: The Art of Sequenced Images*. London and New York: Bloomsbury Academic.

Goldberg, Adele E. (1995). *Constructions: A Construction Grammar Approach to Argument Structure*. Chicago: University of Chicago Press.

Greimas, Algirdas J. (1987). Actants, actors, and figures. English translation by Paul J. Perron and Frank H. Collins. In: *On Meaning: Selected Writings in Semiotic Theory. Theory and History of Literature* 38. Minneapolis, MN: University of Minnesota Press, 106–120.

Groensteen, Thierry (2007 [1999]). *The System of Comics*. English translation by Bart Beaty and Nick Nguyen. Jackson, MS: University Press of Mississippi.

Halliday, Michael A. K. (1978). *Language as Social Semiotic*. London: Edward Arnold.

Halliday, Michael A. K. and Ruqaiya Hasan (1976). *Cohesion in English*. London: Longman.

Hick, Darren H. (2012). Aesthetic supervenience revisited. *British Journal of Aesthetics* 52, 3, 301–316.

Hjelmslev, Louis (1961). *Prolegomena to a Theory of Language*. Translated by Francis J. Whitfield. Madison, WI: University of Wisconsin Press.

Horstkotte, Silke (2015). Zooming in and out: panels, frames, sequences, and the building of graphic storyworlds. In: Daniel Stein and Jan-Noël Thon (eds.). *From Comic Strips to Graphic Novels. Contributions to the Theory and History of Graphic Narrative*. Second Edition. Berlin and New York: De Gruyter, 27–48.

Horstkotte, Silke and Nancy Pedri (2011). Focalization in graphic narrative. *Narrative* 19, 3, 330–357.

Horstkotte, Silke and Nancy Pedri (2022). *Experiencing Visual Storyworlds. Focalization in Comics*. Columbus, OH: Ohio State University Press.

Iser, Wolfgang (1978). *The Act of Reading: A Theory of Aesthetic Response.* Baltimore: John Hopkins University Press.

Jackendoff, Ray (1990). *Semantic Structures*. Cambridge, MA: The M.I.T. Press.

Jackendoff, Ray (2002). *Foundations of Language: Brain, Meaning, Evolution*. Oxford: Oxford University Press.

Kammerzell, Frank (2009). Non-textual marking systems, writing and pseudo script from prehistory to modern times. In: Petra Andrássy, Julia Budka and Frank Kammerzell (eds.). *Lingua Aegyptia – Studia monographica* 8, 277–308.

Kamp, Hans (1981). A theory of truth and semantic representation. In: Jeroen A. Groenendijk, Theo M. V. Janssen and Martin J. B. Stokhof (eds.). *Formal Methods in the Study of Language. Part 1', number 136 in 'Mathematical Centre Tracts*. Amsterdam: Mathematisch Centrum Amsterdam, 277–322.

Kamp, Hans and Uwe Reyle (1993). *From Discourse to Logic: Introduction to Modeltheoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory. Studies in Linguistics and Philosophy* 42. London, Boston and Dordrecht: Kluwer Academic Publishers.

Kress, Gunther (2010). *Multimodality: A Social Semiotic Approach to Contemporary Communication*. London: Routledge.

Kukkonen, Karin (2013). *Contemporary Comics Storytelling*. Lincoln, NE and London: University of Nebraska Press.

Kukkonen, Karin (2015). Space, time, and causality in graphic narrative: An embodied approach. In: Daniel Stein and Jan-Noël Thon (eds.). *From Comic Strips to*

graphic novels. Contributions to the Theory and History of Graphic Narrative. 2nd ed. Berlin and New York: De Gruyter, 49–66.

Kutz, Oliver, John A. Bateman, Fabian Neuhaus, Till Mossakowski and Mehul Bhatt (2015). E pluribus unum: Formalisation, use-cases, and computational support for conceptual blending. In: T. R. Besold, M. Schorlemmer and A. Smaill (eds.). *Computational Creativity Research: Towards Creative Machines. Atlantis Thinking Machines* 7. Amsterdam, Paris and Beijing: Atlantis Press, 167–196.

Langacker, Ronald W. (2001). Discourse in cognitive grammar. *Cognitive Linguistics* 12, 2, 143–188.

Lim, Fei V. (2004). Problematising 'semiotic resource'. In: Eija Ventola, Cassily Charles and Martin Kaltenbacher (eds.). *Perspectives on Multimodality*. Amsterdam: John Benjamins, 51–64.

Lotman, Juri M. (1990). *Universe of the Mind: A Semiotic Theory of Culture*. London: I.B. Taurus.

Martin, James R. (1992). *English Text: Systems and Structure*. Amsterdam: John Benjamins.

Martin, James R. (2014). Evolving systemic functional linguistics: beyond the clause. *Functional Linguistics* 1, 3.

Martin, James R. and Christian M. Matthiessen (1991). Systemic typology and topology. In: Frances Christie (ed.). *Literacy in Social Processes: Papers from the Inaugural Australian Systemic Functional Linguistics Conference, Deakin University, January 1990*. Darwin: Centre for Studies of Language in Education, Northern Territory University, 345–383.

Metz, Christian (1974). *Film Language: A Semiotics of the Cinema*. English translation by Michael Taylor. Oxford and Chicago, IL: Oxford University Press and Chicago University Press.

Mikkonen, Kai (2015). Subjectivity and style in graphic narrative. In: Daniel Stein and Jan-Noël Thon (eds.). *From Comic Strips to Graphic Novels: Contributions to the Theory and History of Graphic Narrative*. 2nd ed. Berlin and New York: De Gruyter, 101–126.

Mikkonen, Kai (2017). *The Narratology of Comic Art*. London and New York: Routledge.

Miodrag, Hannah (2013). *Comics and Language: Reimagining Critical Discourse on the Form*. Jackson, MS: University Press of Mississippi.

Mitchell, William J. T. (1992). The pictorial turn. *Artforum* 30, 7, 89–94.

Pedri, Nancy (2015a). Graphic memoir: Neither fact nor fiction. In: Daniel Stein and Jan-Noël Thon (eds.). *From Comic Strips to Graphic Novels: Contributions to the Theory and History of Graphic Narrative*. 2nd ed. Berlin and New York: De Gruyter, 127–153.

Pedri, Nancy (2015b). Thinking about photography in comics. *Image and Narrative* 16, 2, 1–13.

Postema, Barbara (2013). *Narrative Structures in Comics: Making Sense of Fragments*. Rochester, New York: RIT Press.

Rajewsky, Irina O. (2005). Intermediality, intertextuality, and remediation: A literary perspective on intermediality. *Intermédialités* 6, 43–64.

Rajewksy, Irina O. (2010). Border talks: the problematic status of media borders in the current debate about intermediality. In: Lars Elleström (ed.). *Media Borders, Multimodality and Intermediality*. Basingstoke: Palgrave Macmillan, 51–68.

Rippl, Gabrielle and Lukas Etter (2015). Intermediality, Transmediality, and Graphic Narrative. In: Daniel Stein and Jan-Noël Thon (eds.). *From Comic Strips to Graphic Novels: Contributions to the Theory and History of Graphic Narrative*. 2nd ed. Berlin and New York: De Gruyter, 191–218.

Saraceni, Mario (2016). Relatedness: Aspects of textual connectivity in comics. In: Neil Cohn (ed.). *Visual Narrative Reader*. London and New York: Bloomsbury, 115–128.

Smith, Greg M. (2015). Comics in the intersecting histories of windows, the frame, and the panel. In: Daniel Stein and Jan-Noël Thon (eds.). *From Comic Strips to Graphic Novels: Contributions to the Theory and History of Graphic Narrative*. 2nd ed. Berlin and New York: De Gruyter, 219–240.

Smith, Murray (2012). Triangulating aesthetic experience. In: Arthur P. Shimamura and Stephen E. Palmer (eds.). *Aesthetic Science: Connecting Minds, Brains, and Experience*. London and New York: Oxford University Press, 80–106.

Smith, Murray (2022). Triangulation revisited. *Projections: the Journal for Movies and Mind* 16, 1, 11–24.

Steiner, Wendy (2004). Pictorial narrativity. In: Marie-Laure Ryan (ed.). *Narrative Across Media. The Languages of Storytelling*. Lincoln and London: University of Nebraska Press, 145–177.

Sweetser, Eve E. (2012). Introduction: Viewpoint and perspective in language and gesture, from the ground down. In: Barbara Dancygier and Eve E. Sweetser (eds.). *Viewpoint in Language: A Multimodal Perspective*. Cambridge: Cambridge University Press, 1–22.

Tomasello, Michael (2005). *Constructing a Language: A Usage-based Theory of Language Acquisition*. Cambridge, MA: Harvard University Press.

Tseng, Chiao-I (2013). *Cohesion in Film: Tracking Film Elements*. Basingstoke: Palgrave Macmillan.

Van Leeuwen, Theo (1999). *Speech, Music, Sound*. London: MacMillan.

Wildfeuer, Janina and John A. Bateman (2016). Linguistically-oriented comics research in Germany. In: Neil Cohn (ed.). *Visual Narrative Reader*. London and New York: Bloomsbury, 19–66.

Wittenberg, Eve and Jackendoff, Ray (forthcoming). The co-evolution of pragmatics and grammatical complexity. In: Bart Geurts and Richard Moore (eds.). *Evolutionary Pragmatics*. Oxford: Oxford University Press.

Wolf, Werner (2003). Narrative and narrativity: a narratological reconceptualization and its applicability to the Visual Arts. *Word & Image: A journal of verbal/visual enquiry* 19, 3, 180–197.

Sources

Fred (wri/art) (2011). *Philémon. L'Intégrale. Tomes 1 à 5*. Paris: Dargaud.

Moore, Alan (wri) and Dave Gibbons (art) (1986–1987). *Watchmen*. New York: DC Comics.

Spiegelman, Art (wri/art) (1991). *Maus: A Survivor's Tale. II: And here my troubles began*. New York: Pantheon.

Image Sources

Fig. 1. Bateman et al. (2017: 124); Bateman (2016).
Fig. 2. Fred (2011: 152).
Fig. 3. Spiegelman (1991: 50).
Fig. 4. Spiegelman (1991: 50).
Fig. 5. Moore and Gibbons (1986–1987, Chapter I: 2–4).
Fig. 6. Moore and Gibbons (1986–1987, Chapter II: 26–27).

Prof. Dr. John A. Bateman
Universität Bremen
FB10 – Sprach- und Literaturwissenschaften
Bibliothekstr. 1
28334 Bremen
Germany
E-Mail: bateman@uni-bremen.de

Measuring Inter-Subjective Agreement on Units and Attributions in Comics with Annotation Experiments

Lauren Edlin and Joshua Reiss, School of Electronic Engineering and Computer Science, Queen Mary University of London

Summary. The conceptualisation of units of interpretation and analyses remains an inherent issue across comics studies. Despite the many conceptualisations of comics units from numerous theories and disciplines, empirical assessments of their validity as proxies for reader interpretation have yet to receive adequate attention. We argue that unit delineation practically involves classifying groups of visual and textual markings according to type, function or semantic category. Based on this, we present a nascent methodology for collecting and measuring inter-subjective agreement by comics readers on proposed units of comics and their attributes. We create an online tool to facilitate handmade segmentations on digital comic pages and assigning labels or classifications appropriate to the annotation task, resulting in segmentation-attribute pairs. We demonstrate the methodology through two inter-annotator agreement experiments that test a segment-attribute pair of p a n e l s e g m e n t a t i o n and a judgment of b a c k-g r o u n d l o c a t i o n i n f o r m a t i o n . The first experiment shows that assigning a binary classification for panel background judgments requires refinements. The second experiment reconceptualises the task to assess agreement on two scalar methods, namely Likert ratings and a continuous scale. We argue that these experiments support the claim that we can build models of structures in comics with an empirical anchor of reader judgments through this methodology.

Keywords. Comics, visual narrative, corpus annotation, inter-annotator agreement, empirical research methods

Zusammenfassung. Die Konzeptualisierung von Analyseeinheiten bleibt ein grundlegendes Problem der Comicforschung. Obwohl es viele Ansätze zur Differenzierung von Comiceinheiten gibt, wurden die meisten bislang nicht ausreichend empirisch überprüft. In diesem Beitrag argumentieren wir, dass die Abgrenzung von Einheiten aufgrund einer Klassifizierung von Gruppen visueller und textueller Markierungen nach Typ, Funktion oder semantischer Kategorie erfolgen sollte. Hierfür präsentieren wir eine Methodik, die die intersubjektive Übereinstimmung („agreement") von Comic-Leser:innen bei der

Bestimmung von Analyseeinheiten erfasst und prüft. Mithilfe eines Online-Tools ermöglichen wir es, händisch potenzielle Analyseeinheiten auf digitalen Comic-Seiten zu skizzieren und ihnen Attribute und Klassifizierungen in Form einer Annotation zuzuordnen. Wir demonstrieren die Anwendung anhand zweier Experimente, die die Übereinstimmung der Annotationen in der Bestimmung von Segmenten einer Comicseite sowie der Beurteilung von Hintergrundinformationen testen. Das erste Experiment zeigt, dass die Verwendung einer binären Klassifizierung für Hintergrundstandortinformationen unzureichend ist und Verfeinerungen notwendig macht. Das zweite Experiment dient der Bewertung der Übereinstimmung mittels zweier skalarer Methoden – Likert-Skalen und kontinuierlicher Skalen. Diese Experimente, so unsere Argumentation, unsützen die Annahme, dass empirisch verankerte Bestimmungen von Analyseeinheiten durch Comicleser:innen eine Basis für die Erstellung von Modellen für die Struktur von Comics bilden können.

Schlüsselwörter. Comics, visuelle Erzählung, Korpusanmerkung, „Interannotator-Agreement", empirische Forschungsmethoden

1.   Introduction

Comics artists use visual and textual elements to introduce, repeat, emphasise, or de-emphasise information at particular places across a sequence to communicate effectively (Eisner 2008; McCloud 2006). Analysis of comics structures requires precise ways of discussing these elements. A perennial topic in comics studies is therefore conceptualising meaningful and consistent u n i t s   o f   i n t e r p r e t a t i o n (or a n a l y s i s), often through developing o n t o l o g i e s to define relationships between units and constrain possible configurations of visual/textual information (Schalley 2019). Current conceptualisations of comics units are typically established within disciplines such as semiotics, linguistics, and computer vision/artificial intelligence. However, empirical assessment of unit conceptualisation validity as proxies for interpretation by everyday readers has not received adequate attention.

Common practices across implicit and explicit approaches towards defining comics units include grappling with the cognitive gap between perceptions of visual markings and higher-level representations, and appropriately delimiting discrete units in non-discrete images and image sequences. In other words, unit delineation involves classifying groups of visual and textual markings according to type, function or semantic category. In this research, we generalise the process of visual element delimitation and classification to develop a practical method for assessing how proposed units are interpreted by multiple readers. Quantitative measures of inter-subjective interpretation can be used in conjunction with current unit conceptualisations as a type of empirical inflection point to further investigate assertions and descriptions about comics structure.

This article presents a nascent methodology to assess inter-subjective interpretation agreement on proposed comics units within an annotation scheme. In general, an annotation scheme prompts annotators to delimit areas on comics pages and assign each area a classification, label, or rating that reflects the type and/or content of the proposed units. The resulting unit is a segmentation-attribute pair. The annotations are assessed using agreement measures commonly used in computer vision and computational linguistics. We demonstrate this methodology's capacity to refine annotation schemes by developing an annotation task, assessing inter-annotator results, and re-conceptualising and retesting the task. This refinement process follows the MAMA (Model-Annotate-Model-Annotate) cycle (Pustejovsky et al. 2017: 24), which is a term coined to express incremental improvement to annotation schemes for text and language corpora. We develop a prototype of a browser-based annotation tool to facilitate efficient annotation of digital comics pages, in which annotators are prompted to create image segmentations paired with classifications regarding the segmentation's content. We investigate the efficacy of this approach through two inter-annotation experiments to test overall reader agreement. We also assess the specific methodological setup, such as investigating expert versus naïve annotators, recruitment through word-of-mouth or crowd-sourcing, and annotation task conceptualisation.

Building directly on our previous work (Edlin and Reiss 2021) which investigated inter-annotator agreement on segmentation and classification tasks assessing panel, character, and text sections, this article offers an in-depth study on interpretations of a specific conception of background location information amount within panel segmentations. Background information is defined as any non-character and non-textual sections of markings in a panel image that explicates the setting or location depicted in that panel, according to a reader. This concept is not only developed to demonstrate the annotation methodology, but is also an attempt to identify places in a comics narrative where the background location appears to have been 'dropped' – that is, the within-panel image does not provide any indication of the location or setting in that panel, and instead typically depicts a single tone. Dropped backgrounds appear across a wide variety of comics types and artistic styles. Figure 1 shows examples of sequential panels from two comics with different art styles and publication formats. Both sequences depict a change in background information amount, namely from some information to no information – in other words, the background has dropped out. The veracity of this concept is tested by reader judgments to determine whether it can be implemented in future work or requires refinement.

The article proceeds as follows: Section 2 gives a brief survey of unit conceptualisations across comics studies to motivate the approach taken in this research. Section 3 describes the Comics Annotation Tool (CAT), which is a prototype browser-based annotation interface used to collect annotations from individual annotators.

Fig. 1. Examples of background information amount changes between two sequential panels; **Fig. 1a (top).** Markings in the first panel suggest this scene is located in a living room. No such markings appear in the second panel, although a reader could infer the second panel takes place in the same living room; **Fig. 1b (bottom).** The first panel suggests the action is taking place in a bedroom, however the second panel has a neutral tone besides the character and speech bubbles.

The CAT is used in two experiments described in sections 4 and 5. Section 4 explicates the concepts of panel segmentation and background location information amount, and presents an inter-annotator agreement study where

annotators assign a binary classification of some background information, or no background information, to each panel segmentation. Building on these results, section 5 tests several versions of spectrum-based background location amount annotation tasks to refine its previous conception and determine whether readers agree on background information amount using a more fine-grained scale. Reflections on the methodology presented through both experiments are given in section 6. The two experiments demonstrate that the wider background location information conceptualisation needs refinement, although the concept of dropped backgrounds has merit for future work. Through these experiments, the overall methodology is shown to sufficiently capture reader interpretation, although segmentation tasks tend to procure agreement more easily than assignment tasks. Finally, shortcomings of the methodology which point to directions for future work are described.

2.   Background: Conceptualisations of units across comics studies

Comics exhibit textual and visual information through repeated conventions and representations. The complexity of these configurations leads to a variety of analytical approaches at various levels of representation, from sub-representational markings (e.g. image contrast, line groups) to high level page compositions (e.g. panel sequences and page layout). An objective comics ontology based on a comprehensive understanding of cognitive processes from visual perception to higher-level semantic representation is not feasible. Many conceptualisations of comics units have therefore been proposed. Such conceptualisations usually follow an established theory or discipline, and are used to investigate or articulate a particular aspect of comics structure and relations between defined elements according to that theory.

We provide a brief, non-exhaustive survey of prominent comic unit conceptualisations from various disciplines, and focus on how units are defined implicitly or explicitly. The methodology described in this research emerges from generalising unit delineation and attribution across these conceptualisations.

2.1 Semiotic and linguistic approaches

Theories and methods from semiotics and linguistics naturally apply to the study of comics, as understanding the relation of form to meaning is foundational to comics research. The backbone of these approaches often relies on defining one or a set of fundamental comics units necessary for the investigation. Therefore, a range of representational levels of units have been defined and analysed.

On the lower end of the spectrum, structuralist scholars examine markings such as lines, dots, or small line groupings and attempt to identify a paired meaning. The inter-relations and compositions of these sub-representational units, in turn, contribute to the meaning of higher-level units such as characters, scenes and panels, and their subsequent paired meanings (e.g. Gauthier 1976: 113). See Cohn (2012), Meesters (2017) and Miller (2017) for summaries of this research. Similarly, the concept of combinatorial morphology formalises how specific combinations of sub-units create coherent compositions. Visual elements that have representational meaning on their own can be affixed with additional markings that produce a novel meaning. For instance, a depiction of a woman with a light bulb above her head indicates that the woman had a sudden burst of inspiration (Cohn 2018b).

More conventional elements such as speech bubbles, instances of characters, and sound effects, are often a primary unit for analysis. Peircean semiotic traditions, which categorise visual signs within complex taxonomies based on a marking's resemblance to its referent, have been widely applied to comics (e.g. Magnussen 2000; Saraceni 2003), and may focus on analyses of a single element. Such investigations include analyses of panels (Caldwell 2012), onomatopoeia (Guynes 2014), characters displaying gestures and action types, and objects (Szawerna 2013), among others. Practitioners of comics also discuss how to effectively implement elements such as speech bubbles and character design in comics creation (Eisner 2008; McCloud 2006). Lastly, units on this representational level are also conceptualised through discourse representation theory (DRT). DRT (Kamp 1981; Kamp et al. 2011) posits that a receiver of a communicative utterance builds a mental representation reflecting information in the communication (Geurts et al. 2020). The mental representation updates accordingly as new information is introduced – however, exactly what information is being updated needs to be defined. Abusch (2012), for example, identifies instances of an individual character as discrete areas in a picture by colouring in the exact shape of the character on page. This work is further built upon by Maier (2019) and Maier and Bimpikou (2019).

A very common primary unit of analysis is panels, which typically grant sequential structure to comics. Panels are also a fundamental unit of the visual narrative grammar (VNG) (Cohn 2013a, b, 2018a). Each panel in a sequence is assigned a syntactic category based on its particular narrative function. There are valid and non-valid sequences of panel order according to their narrative category, and valid orders are reflected in abstract hierarchical recursive tree-structures as commonly used in linguistic syntactic analyses (Cohn 2013b: 417; Cohn 2015; Cohn and Kutas 2017). McCloud (1993) describes a taxonomy of six panel transitions based on reader judgments regarding amounts of time and action depicted from one panel to another – the fundamental unit here is a panel dyad, with a judgment on the transition type between them (McCloud 1993). Lastly, panels are a fundamental element with compositions that give rise to page structure, facil-

itating investigations into relationships between panel structures, panel content, and overall page structure. Groensteen (2007), for instance, argues that panels are the fundamental signifying unit of comics, and coins 'arthrology' as the study of panel configurations, allowing for holistic analyses of sequential and distantly placed panels. Explicit classifications of panel layout compositions lead to further work on the relationship between lower-level panel content and page layout (e.g. Pederson and Cohn 2016; Bateman et al. 2017).

Finally, many analyses examine units and their relations to one another on lower to higher levels of representation. Bateman and Wildfeuer (2014b; a) provide a comprehensive account of higher-level discourse relations within and between panels by applying concepts from Segmented Discourse Representation Theory (SDRT) (Asher and Lascarides 2003). Markings within panels are identified and interpreted through knowledge and experience (e.g. wavy lines above a pipe means a smoking pipe), and demarcated as non-discrete units with a variable assignment. Wildfeuer (2019) further develops delimiting within panel elements by describing a formal notation that assigns perceptually salient features to existential quantifiers and variables. The notation qualitatively expresses entailment between elements that gives meaning to a panel scene, and is a formal description of reader interpretation. Lastly, Yus (2008) describes stages of inferences which a reader goes through when selecting and reading a comic from the cover image to page layouts and processing within and between panel elements.

2.2 Computational approaches

Computational methods are used across comics studies for a wide variety of purposes, which can be broadly subdivided into automatic content identification and automatic content generation. Automatic identification includes detection of sub-representational and conventional elements as well as predicting the presence of an element based on configurations of other elements and features. Automatic generation describes computation-based creation of parts of comics or whole comics. These investigations require precise specifications of the units under analysis (see Augereau et al. 2018; Laubrock and Dunst 2020 for comprehensive surveys).

On the sub-representational level of automatic content identification, computational methods are used to identify and isolate areas of certain textures (Liu et al. 2017) and screen tones (Ito et al. 2015). Properties of images such as colour contrasts, image brightness, the types and numbers of different shapes are detected and used to find higher-level representations such as characters (Mao et al. 2015), or to associate broader concepts such as artistic style (Dunst and Hartel 2018).

Automatic detection of comics features typically involves recognising low-level visual marker configurations, and using techniques

from document analysis and computer vision to identify and classify objects in an image. Segmentation tasks seek to match the correct labels to boundary-mark regions of comics images, which are often within a rectangular b o u n d i n g b o x containing the sought after visual element. Numerous studies show that text (e.g. Rigaud et al. 2017), speech bubbles and captions (e.g. Dubray and Laubrock 2019), panels (e.g. Pang et al. 2014), and characters (e.g. Nguyen et al. 2017; Qin et al. 2017) can be accurately segmented. Automatic segmentations are compared to a g r o u n d t r u t h , or hand-annotated sets of comics pages, to determine a correct segmentation. Knowledge-based ontology approaches add a higher semantic level to assist lower-level extraction processes by adding additional relation constraints between designated elements to facilitate their correct identification. Guérin et al. (2017) describe a formal ontology that uses the concepts of panel, balloon, balloon tail, text line, and character (Guérin et al. 2017: 22) to classify segmentations, or regions of interest, derived from lower-level extraction processes (Rigaud and Burie 2018; Rigaud et al. 2015).

A u t o m a t i c c o m i c s g e n e r a t i o n , on the other hand, typically describes an ontology of visual and textual elements that a programme selects and organises to create readable comics. The visual and textual elements used to produce comics may be pre-stored – that is, all the drawings and text are already created and not themselves generated – or a series of images is automatically segmented based on some heuristic. The first (to our knowledge) comics generation program, C o m i c C h a t (Kurlander et al. 1996), creates a comic beginning with text from online chat logs. Characters are created by combining pre-drawn heads and bodies, matched with the corresponding text, and placed in a sufficiently sized panel with pre-drawn backgrounds. The placement of character, speech balloons, text elements, and backgrounds on a page are organised according to a set of spatial placement rules. Similar programs inspired by Comic Chat have appeared more recently (e.g. Alves et al. 2008; Soares de Lima et al. 2013; Shamir et al. 2006).

2.3 Proposed approach: Efficient annotation for inter-subjective interpretation measurement and agreement assessment

The overview above shows that unit conceptualisation ranges from low to high levels of representation across theoretical and methodological practices. Units are often delineated on different levels within a framework to investigate links from 'parts to wholes' in visual content. In other words, each approach grapples with the cognitive gap between perceptions of visual markings and higher-level categories, either by describing how to delimit discrete units from non-discrete images, or by explaining how defined subcomponents contribute to larger visual compositions. The level of unit granularity ranges from high-level image descriptions of panels or

page layout as foundations of analysis all the way down to grouping areas of pixels in a comics page image under distinct labels.

What the approaches have in common is that they rely on the judgement of researchers to characterise units, sub-components, and their attributes. What has not received adequate attention is the assessment of inter-subjective agreement by everyday comics readers about the interpretation of the proposed units. Quantifying subjective interpretation is not only useful for understanding whether a unit has been well-conceptualised, but also serves as an empirical grounding for developing links between higher and lower units. Reader interpretation reveals ambiguity and vagueness in particular where readers may have consistently differing interpretations: For instance, if 'character' is a fundamental unit, can we be sure that readers actually discern the same set of markings as representing the same character across a story? Disagreement between readers may reveal intended ambiguity that can be incorporated into a theory, or show that the concept of 'character' is not well-formed.

In light of this, we explore a methodology that facilitates efficient comics annotation from multiple readers to quantify inter-annotator judgments. By abstracting the process of unit conceptualisation from across various methods and approaches, an operational definition of units is achieved by classifying groups of visual and textual markings according to type, function, or semantic category. The annotation process implemented here is therefore developed to have annotators outlining, or segmenting, areas on a comics page and assigning each segment an attribute resulting in a segmentation-attribute pair (segmentations can, of course, have numerous attributes, as tested in one or over several studies). The motivation for the definition of each segment and its attribution is described as a task in an annotation scheme. We borrow several practices and measures from computational linguistics and computer vision to quantity and measure agreement, and follow a MAMA (Model-Annotate-Model-Annotate) cycle (Pustejovsky et al. 2017: 24), where an initial model of units and attributes is created from theoretical assumptions, evaluated, and updated based on the results of inter-annotator experiments. This methodology is demonstrated below through two experiments which test a proposed segmentation-attribute pair unit.

3.   The Comics Annotation Tool (CAT)

We aim to assess annotator agreement on the basic amount and type of information across comics pages. The general method we use involves recruiting a number of annotators and providing them with an annotation scheme, which instructs them to demarcate areas of comics pages and assign labels to these segments. We develop the Comics Annotation Tool (CAT) to accommodate these tasks.[1] The CAT is a browser-based comics
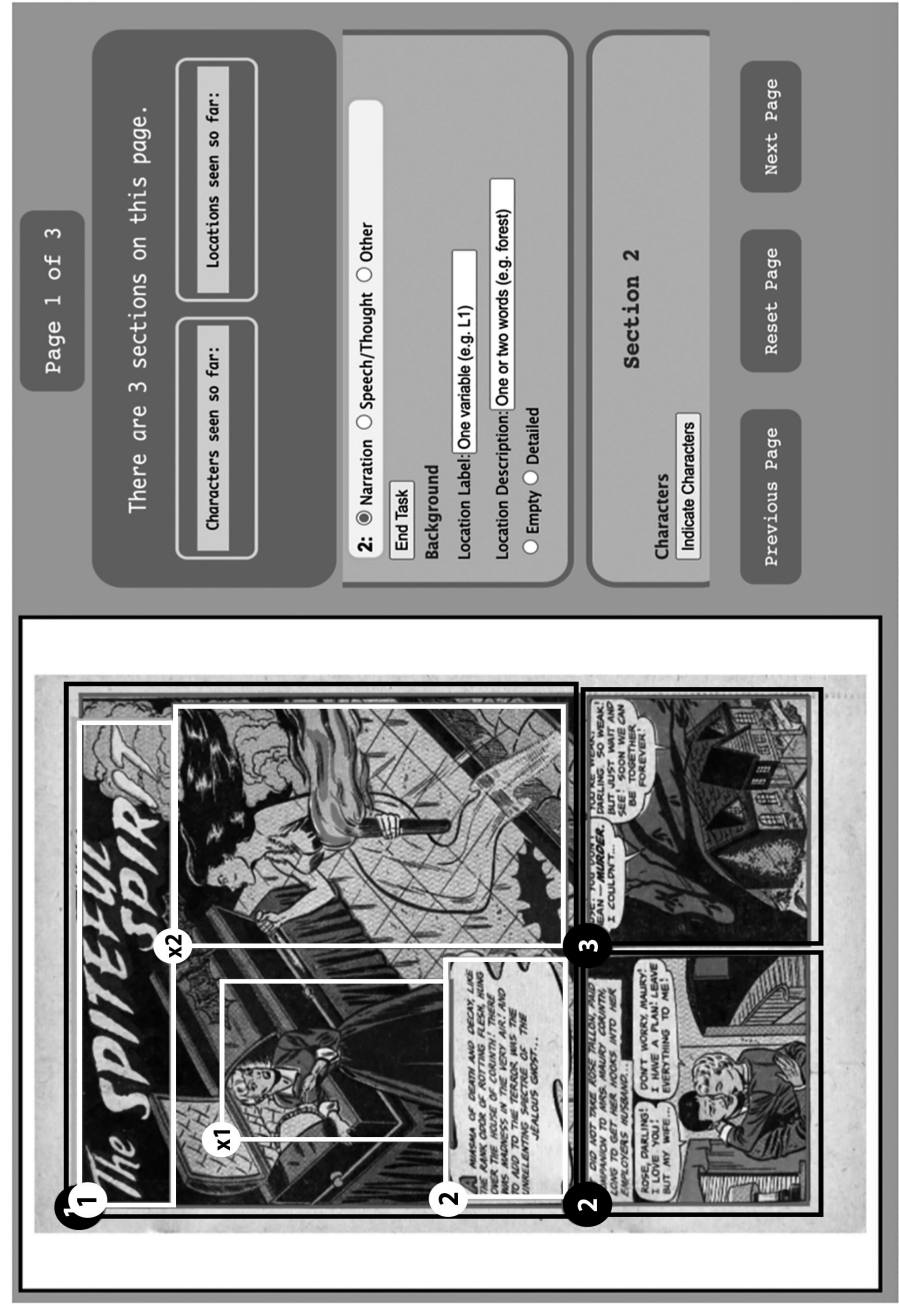
Fig. 2. The CAT's main annotation interface. This version of the CAT is set up for character, text section and panel segmentation on the right, and their associated assignments on the left. Checkboxes for the background information amount task are shown at the bottom of the first light blue section.

mark-up tool that facilitates remote annotation of digital comics pages. Individual annotators can access it via a URL, and are prompted to perform a series of pre-configured annotation tasks in a specific order by instructions and responsive features provided directly in the CAT.

Figure 2 depicts the main CAT interface setup for an experiment testing several segment classification tasks, which are described in previous work (Edlin and Reiss 2021). The comics page on the left is annotated with bounding boxes that are coloured according to the segmentation task.[2] On the right are the reference and labelling prompts, matched to their associated segments on the left by number and colour. Annotators can navigate between pages in the story using buttons at the bottom of the right-hand section of the interface. Once an annotator completes all prompted annotation tasks for all pages in a given story, all segmentations (e.g. pixel positions of bounding boxes on the page) and their associated labels are collected in JSON format and stored in an external database.

When a segmentation task is required, the annotator is prompted to outline areas on the digital comics page image by clicking and dragging rectangular bounding boxes over the desired area. Bounding boxes are used for several reasons: the annotator only has to click and drag on the comic image once, allowing for more efficient and scalable annotation; only two pixel coordinates need to be recorded to assess the size or reproduce the bounding box; and many segmentation tasks for automatic detection of comics elements use bounding boxes, which may facilitate integrating reader-interpreted segmentations described here into corpora that only contain a ground truth, or one interpretation, of the units. Each newly created segmentation generates an input that asks for the annotators' judgment regarding each segmentation's classification, label or reference. The formulation of the input depends on how the labelling task is conceptualised. For example, a binary classification task may provide a checkbox, while a reference labelling task presents a text input.

4. Experiment 1: Inter-Annotator agreement on panel segmentation and binary classification of background information

4.1 Methodology

4.1.1 Annotation scheme and CAT setup

This first experiment investigates panel segmentations and associated background location information, and assesses whether these concepts reach sufficient agreement by aligning readers' interpretations. Panels are conceptualised as coherent and distinct sections based on image structures. Background location

information is a judgment about whether visual evidence regarding the location or setting of the narrative is given in a panel.

The agreements we focus on in this paper were carried out in tandem with other annotations tasks (as described and evaluated in Edlin and Reiss 2021). This overall experiment investigated inter-annotator agreement for several proposed annotation tasks regarding paradigmatic comic unit concepts, including panel segmentation and associated location reference, text section segmentation and classification, and character segmentation and reference. A brief overview of the annotation scheme from the overall experiment is provided in Table 1, which specifies the type of label assigned to each type of segmentation, and a summary of the instructions given to annotators.

Tab. 1. An overview of the annotation scheme assessed in the experiment described in Edlin and Reiss (2021).

ANNOTATION	TYPE	SUMMARY OF INSTRUCTION	CAT INPUT
Panel	Segmentation	Outline clear and distinct image sections, even if there is no clear panel outline.	Bounding box
	Reference (Location)	Assign a label that will stay the same for that specific location throughout the story.	Reference Label (e.g. L1, L2)
	Background Information Classification	Indicate whether information about the setting or location is present in a panel background.	Checkbox {Empty, Detailed}
Text Section	Segmentation	Outline any sections of text within a panel.	Bounding Box
	Classification	Indicate the type of text section.	Checkbox {Narration, Speech/Thought, Other}
Character	Segmentation	Outline active participants in the comic narrative within a panel.	Bounding Box
	Reference	Give each character a label that will stay the same for that specific character throughout the story.	Reference Label (e.g. X1, X2)

The concept of background information amount is more precisely defined according to Table 1. Technically, the concept is understood as a judgment regarding the remaining visual markings outside text and character segmentations within a panel segmentation. After creating a panel segmentation with additional text and character segmentations within the panel, annotators were prompted to judge whether there is any information about the wider location or setting in that particular panel without considering markings in the previously segmented text and character areas. If an annotator perceived any such information, they were to select the detailed category. The empty category should be selected if no such information is perceived. Any objects that a character is interacting with are not considered part of a panel background according to the full annotation scheme.

Fig. 3. A series of panels demonstrating the area meant to be interpreted as the background; **Fig. 3a (top).** The original, unaltered panel; **Fig. 3b (middle).** The same panel with added bounding boxes segmenting text sections and characters; **Fig. 3c (bottom).** The same panel showing only the background areas.

Figure 3 provides an illustrative example of markings intended to be interpreted for background information. Figure 3a depicts a typical panel of a similar style to the comics used in this experiment. Figure 3b shows the same panel with added text section and character segmentations as they would appear in the CAT. Despite the previously described benefits of using bounding box segmentations, they do not precisely outline the complex shapes that often constitute areas depicting text section and characters. Each bounding box includes markings that are technically not meant to be judged for the associated label assignment task. The inverse of this causes some markings that should be included as background segmented within a bounding box. Annotators were therefore instructed to regard bounding boxes as only rough approximations for distinctions between groups of markings intended for interpretation. Figure 3c shows the same panel with a more reasonable expectation of what should be interpreted as background area. The text sections and characters are covered in black, therefore showing only the remaining markings intended for judgment regarding background information amount.

The binary categories of empty and detailed are specifically constructed as an attempt to distinguish panels where all background information is 'dropped' – that is, panels where the location or setting can only be inferred. These types of panels seem to occur across comics from a wide variety of artistic styles, genres, and cultural backgrounds. Refer back to figure 1 in the intro-

duction for two examples. Furthermore, this concept of background location information is constructed to be applicable across many artistic styles; since the annotation task focuses on reader's judgments regarding representations of setting, the actual style of markings – whether an artist uses a richer versus a sparser style, for example – is not a factor in the judgment. However, while the use of dropped backgrounds looks to be prevalent across comics, its use is likely to have different meanings for various authors, genres, or cultural contexts. Identifying cases of these occurrences through reader agreement may therefore be useful in further work to explicate these meanings.

Finally, this unit judgment is used to demonstrate this annotation methodology. As a complex high-level compositional representation, it may be composed of many sub-parts, such as single or groups of objects, other aspects of setting, or more sub-representational aspects such as tones and textures. Since there is incredible potential for image configurations, forcing annotators into a binary choice may reveal instructive disagreements on how to proceed in future work and refine the proposed conceptual unit.

We assessed this annotation scheme on comics stories from *Alarming Tales* comics magazine, published by Harvey for six issues between 1957–1958. We chose these stories because they all have a similar art style that is typical of Silver Age comics, are all of the fantasy sci-fi genre, and are created by several different writers and artists. We limit the scope to one publication to get a precise assessment of agreement on a small set of comparable comics, as disagreements will be difficult to parse on a wider variety of comics at this stage. Finally, these comics have a style that appears to exhibit a range of background information, including potentially clear instances of the 'dropped' background concept. The digital comics were downloaded from the *Comic Book Plus* (2006) Internet archive of comics. Four comics stories with five pages each (for a total of 20 pages) were selected for this experiment.

4.1.2 Participants

A total of ten participants (six female, four male) produced the annotations. All participants are postgraduate students or friends and partners thereof, and were recruited from Queen Mary University of London. All participants speak and read English as their native language or to a fluent level, as all the comics are written in English. Participants were compensated £10/hour and could choose the number of stories they wished to annotate, therefore not all stories were annotated by all annotators. Table 2 lists annotators per story by their ID number and the total annotator pairs for inter-annotator agreement assessment. Annotations made by the first author (annotator 0 in Table 2) are included to assess the effectiveness of naïve versus expert annotators. All other annotators were only given the annotation scheme and instructions on how to use the CAT remotely, and did not receive training or further insight into the experiment.

Readers interpret visual information, including comics images and sequential structures, differently due to cultural background and biases towards particular meanings in visual information. In addition, readers may be more adept at interpretation through more exposure to comics and other visual media. Annotators were therefore given the Visual Language Fluency Index (VLFI) (Cohn 2014) questionnaire which was used to compute a quantitative metric of visual fluency per annotator. All participants scored in the average fluency range, except Annotators 8 and 9 who scored in the low fluency range. The mean VLFI score across all participants is 13.49, indicating average fluency overall.

Tab. 2. Annotator Number and Total Annotator Pairs per Story.
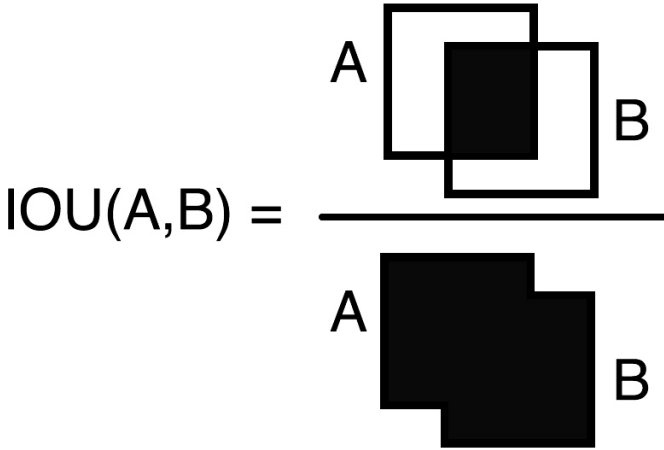
STORY NO.	ANNOTATORS (BY NO.)	TOTAL ANNOTATOR PAIRS
1	0, 1, 2, 3, 6, 7, 8	21
2	0, 1, 2, 4, 8	10
3	0, 1, 2, 5, 8	10
4	0, 1, 2, 5, 9	10

4.1.3   Inter-Annotator agreement measures

Segmentation agreement was measured using Intersection over Union (IOU, or Jaccard Index). IOU is a quantitative metric of the similarity between two sets, and is defined as the size of the intersection of two sets divided by the size of the union of the same sets: $IOU(A,B) = |A{\cap}B|/|A{\cup}B|$. It is a widely used evaluation metric in computer vision (Rezatofighi et al. 2019; Rosebrock 2016; Szeliski 2020). Object recognition in particular evaluates the amount of overlap between an algorithmically generated bounding box against a ground truth bounding box, with the latter surrounding a depicted object attempting to be detected. Figure 4 depicts a visualisation of the IOU metric.

In this experiment, a set is understood as the pixels within a bounding box on a comics page. An IOU score between a pair of annotators' segmentations measures the amount of overlap between both annotator's respective bounding boxes; the more overlap between the segmentations, the higher the IOU score will be. This measurement therefore indicates whether annotators agree on the judgements regarding the use of a comics page's 'space'. For instance, annotators show agreement that the top right of a page shows a panel by segmenting the same areas. Scores are within a range of [0, 1], and what is considered a good score is up to some

interpretation. In automatic object detection, a score of 0.5 and above is typically considered a correct detection against a ground truth (Everingham et al. 2015). Previous work (Edlin and Reiss 2021) supports that bounding box segmentations for text agreement should reach a threshold of 0.6, and 0.5 and above for character segmentation.

Fig. 4. Visualisation of intersection over union (IOU) between two bounding boxes.

Annotators are instructed to create panel segmentations in the order they read them, and segments are numbered to reflect that order. However, disagreement will typically occur due to a difference in the number of created segmentations between annotators. A mapping algorithm is developed to assess the IOU scores between all possible segmentation pairs between two annotators, per page. The best overall IOU score for a permutation of segmentation pairs gives a 'mapping' of corresponding panels between annotators. The measure of overall panel segmentation agreement is the total mean IOU score of matched panel pairs per annotator, per story. The full mapping algorithm and overall IOU scores that include non-mapped panels are available (Edlin and Reiss 2021). Since the inclusion of non-mapped segments produces a very similar score, only mapped segment results are reported here.

For the background information task, we use Krippendorff's α (KA) (Krippendorff 2011). KA is a quantitative metric of inter-annotator agreement that is widely used in corpus linguistics and content analysis (Artstein and Poesio 2008). The metric rates the extent to which annotators assign the same category or value to the same segmentation. The score ranges from -1 (complete disagreement/negative agreement) to 1 (perfect agreement), with 0 being chance agreement – that is, annotators appear to be randomly assigning values. Although there are no clear thresholds for sufficient agreement, 0.68 is typically considered adequate while 0.8 indicates excellent agreement (Artstein and Poesio 2008: 591). Scores meeting these

thresholds indicate a well-conceptualised classification pertaining to content on comics pages, as annotators perform similarly under the same instructions. Lower scores suggest that a task is difficult to understand either due to bad instructions, an incoherent concept, or that annotators are unreliable and choosing random categories.

A KA score is calculated for all annotators against all annotators, as well as between each pair of annotators, per story. Pairwise scores allow for a precise assessment of agreement between each annotator. This gives further context to the all-against-all score, and may indicate potential unreliable annotators. Finally, KA scores were calculated only for mapped segmentations. KA scores were calculated using the *Fast Krippendorf* python package (2017), which is based on the implementation in Grill (2017).

4.2 Results

The results in Table 3 show the means and standard deviations of annotator pair mean IOU scores per story, as well as the mean and standard deviations for annotator pair and all-against-all KAs. The panel segmentations exhibit very high agreement for all stories, except for Story 4, which shows a lower agreement. All scores exceedingly pass the traditional threshold of 0.5, indicating significant overlap between mapped segmentations.

Tab. 3. Panel Segmentation IOU Scores and Background Information KA Scores per Story.

	PANEL SEGMENTATIONS		BACKGROUND INFORMATION		
	Pairwise IOUs		Pairwise KAs		All-against-all KAs
	Mean	*Std.*	*Mean*	*Std.*	
Story 1	0.931	0.027	0.3568	0.241	0.3812
Story 2	0.955	0.016	0.526	0.2715	0.5369
Story 3	0.933	0.023	0.6639	0.1347	0.6733
Story 4	0.784	0.094	0.6545	0.1548	0.599

Mapping disagreements show differences in parsing panel structures. There were only a few non-mapped panels in Stories 1–3. These can be attributed to title segmentation – annotators diverged on whether to include large title text within adjacent image segmentations, or place them in their own segmentation. Story 4, however, has many more mapping disagreements,

primarily due to different decisions on including text section blocks with neighbouring images. While Stories 1–3 have text section boxes to the left and right of distinct image sections, Story 4 also has text boxes above image sections. The latter composition appears to have a wider range of interpretation in relation to adjacent images. Subsequently, the relationship between areas of text and image in the interpretation of panel should be examined in future work; See our previous work (Edlin and Reiss 2021) for a more detailed discussion.
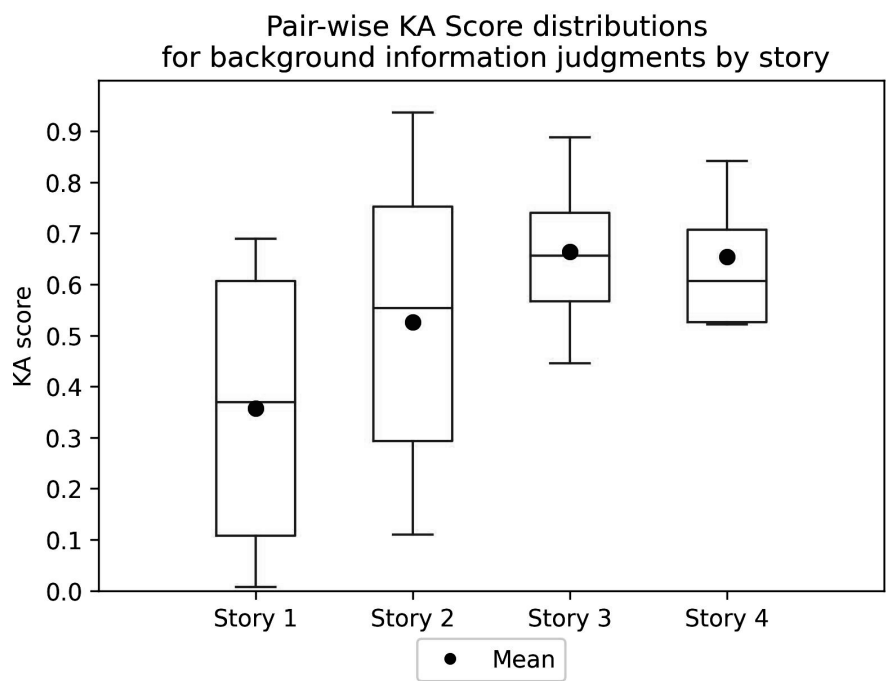
Fig. 5. Boxplots of the distribution of pair-wise KA scores between all annotators for the background location information task, per story.

The background information task generally achieved low agreement across stories, with only the all-against-all KA for Story 3 nearly reaching the 0.68 threshold for adequate agreement. Figure 5 presents the distributions of each individual annotator pair's mean IOU score per story. The pairwise KA distributions show a large range of scores, meaning that some pairs of annotators agreed much more than others. Story 3 exhibits the most overall agreement as it has the highest mean pairwise KA score and the lowest standard deviation between annotators. Story 1 has the least overall agreement, although Story 2 has the highest standard deviation.

4.2.1   Per-panel analyses of background information

Within the 119 mapped panels annotated across all stories, 66 panels (55.5%) exhibit unanimous agreement between annotators, while 53 panels (44.5%) present disagreement from at least one annotator. Story 1 has the highest number of disagreed panels at 74.2%, followed by Story 2 at 54.8%. Stories 3 and 4 exhibit similar percentages of disagreed panels with Story 3 at 22.2%, and Story 4 at 23.3%.

All stories contain unanimously agreed upon empty and detailed panel segmentation assignments. An example of an image that all annotators classified as detailed from Story 4 is shown in figure 6a. These panels often appear to show images from a 'zoomed out' viewpoint. Story 4, followed closely by Story 3, has the most panels with unanimous detailed agreement. Unanimously labelled empty panel segmentations, on the other hand, appear to often depict close-up images of a single character or multiple characters in conversation, accompanied by a solid tone in the remaining space. Figure 6b shows an image from Story 2 classified as empty by all annotators and exemplifies a typical empty panel segmentation. Story 2 has the most panels with unanimous empty agreement and appears to feature many panel segmentations depicting two characters in conversation.

Fig. 6. Examples of background information classifications with unanimous agreement; **Fig. 6a (left).** A panel segmentation classified as detailed by all annotators (Story 4, Page 4, Panel 8); **Fig. 6b (right).** A panel segmentation classified as empty by all annotators (Story 2, Page 5, Panel 3).

A qualitative assessment of disagreed upon panels reveals several potential causes. Images with discernible and prominent objects in the foreground that also show a single colour or gradient area in the remaining space frequently exhibit disagreement. Figure 7 provides an example. In this image, the scientific instrument is taking up most of the non-character and non-textual space. Since the characters are interacting with the instrument, it should

Fig. 7. An example of image with disagreement possibly due to the image's foreground configuration, which belies the category 'empty' (Story 1, Page 2, Panel 2).

technically be discounted from the background. The remainder of the image is a black and blue tone that can be interpreted to be a neutral tone with no location or setting information. This image can therefore be classified as empty according to the annotation scheme. However, most of the image area is taken up by an object, making the category of 'empty' an unintuitive descriptor. Annotators 0 and 7 assigned the segmentation as empty, while Annotators 1, 6, and 8 classified it as detailed. Story 1 seems to exhibit many of these panel types, and is the story with the most disagreed upon panels. Stories 3 and 4 appear to have very few panels with similar prominent foregrounds, and both have the most unanimously agreed detailed judgements.

Second, disagreement commonly occurred for images that appear to show a relatively small indication of the setting, often through only one or two markings. Figure 8 provides an example of such an image. The image background

contains a primarily green tone embellished with black shapes. One may infer that these shapes represent a nearby wooded area, meaning that location or setting is present. However, Annotators 1 and 8 classified this section as empty, while Annotators 0, 6, and 7 assigned this section as detailed.

Lastly, disagreements occur when single colour, two-toned, or gradient areas are interpreted to have a meaning relevant to or inferring a wider setting, such as depictions of shadows or sky. Figure 9 depicts two characters looking at two giant intertwined plants, with the remain-

Fig. 8. An example of a disagreement based on several background markings (Story 1, Page 4, Panel 6).

Fig. 9. Example of background disagreement based on interpretations of lighting, sky, or shadows (Story 3, Page 5, Panel 3).

ing non-textual areas filled in with a light blue colour. Annotators 0 and 5 classified this image as detailed, while Annotators 1 and 8 classified it as empty. The blue areas may have been interpreted by Annotators 1 and 8 as a neutral tone to fill the space and therefore assigned empty. Annotators 0 and 5, on the other hand, may have interpreted these areas as the sky, which gives some information on the setting such as being outside in an open area, leading to a classification as detailed.

Annotator 1 and 8 may have also interpreted that area as sky, but did not consider the sky to be indicating a more specific or relevant location. In addition, these images tend to have prominent objects presented distinctly in the foreground. The black area in figure 7 above could also be interpreted as a shadow, providing another potential reason for disagreement in this image.

4.2.2   Annotator reliability on background information

Low KA scores are primarily attributed to an incoherent concept for annotation, but can also indicate annotator unreliability. While many annotators tended to agree amongst one another, several annotators consistently disagreed with others across the board. Annotators 3 and 6 in Story 1 consistently produced low agreement, and both were prone to assigning the detailed category against all other annotators empty assignments. Nevertheless, these annotators had an average score on the VLFI metric, which suggests proficient visual language literacy. With these annotators excluded, the all-against-all KA score for Story 1 is 0.5836. It cannot be clearly determined, however, whether these annotators interpreted the instructions in a way conducive to more detailed assignments, or through unreliability. Heatmaps depicting the KA agreement between each annotator pair for each story are available.[3] Finally, there did not appear to be consistent disagreement between expert (Annotator 0) versus all other naïve annotators.

4.3 Discussion

The panel segmentation tasks show very high agreement, while the low to adequate agreement scores for background location information judgments

show that forcing a binary choice between two categories is not a robust conceptualisation of background image space. Overall, segmentation tasks proved to yield higher inter-annotator ratings compared to assignment tasks according to previous work (Edlin and Reiss 2021). The panel segmentation results in particular may skew high because most panels in these stories have a rectangular shape that fits well in a bounding box.

A variety of entangled structural and semantic factors contribute to background location information disagreements. Generally, structural disagreements are caused by differences in demarcation of actual background areas, while semantic disagreements occur when the same area is interpreted differently. While some annotators consistently disagreed with all others, disparate interpretations of the task are the main reason for disagreement. The use of the term 'background' is likely to have added confusion, as the term may not intuitively align with the image areas intended to be so attributed. Despite these disagreements, the binary categories appear more suitable for some stories than others. It seems that more agreement is accomplished when there is a substantial amount of unanimously agreed detailed panels. This is the case for Stories 3 and 4, both of which had higher scores across the board compared with Stories 1 and 2. None of the stories had a relatively high number of agreed upon empty panels, although it can be speculated that some stories – for instance, comic strips that typically show two characters in conversation – would be suited to a binary classification task.

Nevertheless, the presence of unanimously agreed upon empty classifications across all stories supports that there are image configurations that can intuitively be labelled 'empty'. This suggests that the panel 'drop out' concept is not without merit. We find that these panels tend to have similar visual configurations, and often show one or several characters in conversation with a background tone that does not lend itself to semantic interpretation of setting (e.g. a shadow or sky). It may be beneficial to try and isolate these particular types of images in a refined annotation scheme.

Overall, the low scores for the binary classification and the type of disagreements suggest that the annotation tasks can be improved by taking into account information 'amount', or a range of interpretation between no information and some information present. While the survey of potential disagreement sources given above suggests either more structural-based and semantic-based interpretations of background information, these potential causes for disagreement remain intertwined. However, we note that visual configurations that most often produced disagreement have areas of a single colour with additional objects or markings, such as prominent foreground figures or small markings in a space with an otherwise neutral tone. Making additional categories or gradients available to annotators may capture more fine-grained interpretations, without having to specifically address structural and semantic causes of disagreement separately. Gradient scores may also give the description of 'empty' a clearer definition. It is unclear, however, what type of scale would most accurately capture interpretations of background information amount.

5.   Experiment 2: Crowd-sourced judgements between three scales of background information amount

Building on the results from Experiment 1, we re-conceptualise background location information to be interpreted on a spectrum from no information to full information. We conduct inter-annotator agreement experiments on three different scale types – c o n t i n u o u s, o r d i n a l, and b i n a r y – to determine which best measures more fine-grained perceptions of background information. We use a between-subjects experiment design and crowd-source a unique set of participants to annotate each scale for each story. Crowd-sourcing comics annotations appears a promising route for gathering large numbers of annotation from a number of annotators (Tufis and Ganascia 2019), therefore we try this approach here.

5.1 Methodology

5.1.1   Annotation scheme and CAT setup

Three versions of the background information amount annotation task are tested: one version prompts annotators to indicate information amount on a c o n t i n u o u s spectrum between 1 (no information) and 5 (full information), another version is a 5-point o r d i n a l scale between categories 1 (no information) and 5 (full information), and the final version presents the b i n a r y   c a t e g o r i c a l choice of 0 (no information) and 1 (information).[4]
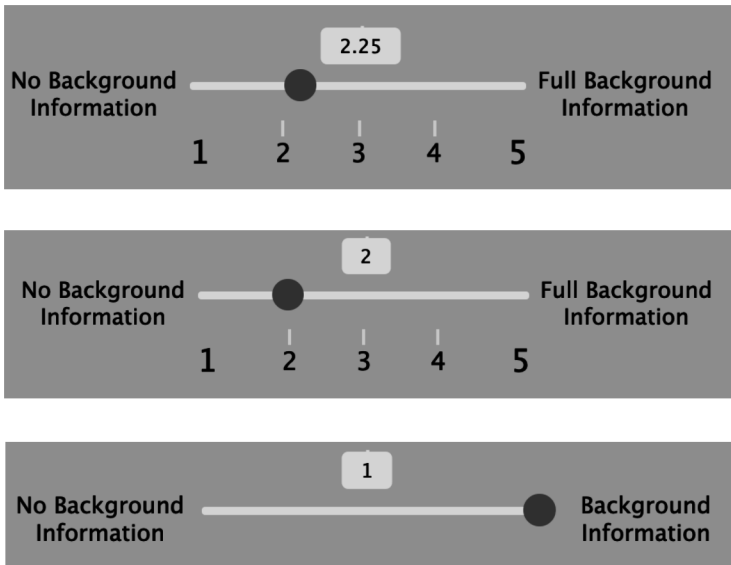
Fig. 10. The continuous (top), ordinal (middle) and binary (bottom) scales as presented on the CAT.

Three configurations of the CAT were developed to reflect each annotation scale type. Figure 10 provides an example of how each scale is presented on the right side of the main CAT interface. Annotators can drag the purple circle left and right along the grey line to specify a numerical information amount. The continuous scale on the top allows for indications between integers 1 to 5 up to two decimal places. The ordinal scale in the middle allows the selection of whole integers between 1 and 5. The binary scale allows for the choice between 0 for no information, and 1 for some information. Note that the binary choice task is presented on a scale rather than as a checkbox as in Experiment 1.

Unlike Experiment 1, the comics pages were shown with pre-segmented bounding box panel segmentations to guide annotators to the corresponding section on the right side of the interface. These page segmentations reflect the agreed upon segmentations from Experiment 1, using the lead author's annotations when in doubt.

Stories 1 and 2 from the first experiment are selected for annotation. We use the same stories to allow for a comparison between recruited and crowd-sourced annotations.

5.1.2 Participants

Participants were recruited on the online crowd-sourcing platform *Prolific* (2022). While there are a number of alternative participant recruitment platforms, *Prolific* was chosen for several reasons: i) it was developed specifically for academic research, ii) it could be easily linked to the CAT, iii) it has clear rights and obligations including minimum pay of £5.00/$6.50 per hour, and iv) participants are typically more naïve to experimental research tasks than on other common crowdsourcing platforms (e.g. Mechanical Turk) (Peer et al. 2017).

Tab. 4. Descriptive statistics of participants for Experiment 2.

SCALE TYPE	STORY NO.	NO. COMPLETED (RETURNED /REJECTED)	MEAN AGE	GENDER*	VLFI (MEAN/STD)
Continuous	1	10 (2/0)	26.1	5M/5F	4.9/3.86
	2	10 (6/0)	29.9	1M/8F/1NB	9.8/6.82
Ordinal	1	10 (1/0)	33.1	4M/6F	10.9/6.44
	2	9 (1/1)	30.4	1M/8F	7.75/3.27
Binary	1	10 (7/0)	33.8	3M/6F/1U	5.9/4.17
	2	10 (3/0)	32.2	5M/5F	6.83/5.35

A different set of participants was recruited for each type of scale and each story. Table 4 provides descriptive statistics of these participants, including mean VLFI score per annotator set. R e t u r n e d participants started the annotation task but did not complete it, and their data was not collected. R e j e c t e d participants did not pass the attention check and were also not included in the analysis. Each mean VLFI score per set indicates low flu-

ency, and is below the mean VFLI score for the participants in Experiment 1. Participants were pre-screened through *Prolific* to be between the ages of 21–75, have UK or US nationality, exhibit fluency in English or have English as a first language, and to have achieved an undergrad degree. These attributes were selected to be similar to the demographics of the participants recruited in Experiment 1.

5.1.3   Inter-annotator agreement measures

The new scales produce continuous and ordinal data in addition to the categorical data from the binary scale. KA is used to test for all-against-all annotator agreement as a unifying measure. In addition, we assess the strength of correlation between annotator pairs using Pearson's correlation and Spearman's rank correlation. Spearman's rank is a common method for measuring correlation between Likert scale items, and shows whether the ratings of one annotator correspond to the same rating given by another annotator. Both Pearson's *R* and Spearman's rank correlation coefficients range from -1 to +1, with values close to -1 indicating a negative correlation, values close to +1 indicating a positive correlation, and values near 0 representing no correlation. All calculations were done using the inbuilt functions in the SciPy python library (Virtanen et al. 2020).

5.2 Results

All-against-all KA scores are reported in Table 5. None of the scores reach a sufficient threshold of agreement. However, the binary scale performed better on Story 2 – with results similar to those in Experiment 1 – while the ordinal and continuous scales performed better on Story 1.

Tab. 5. All-against-all KA scores per scale, per story.

	BINARY	ORDINAL	CONTINUOUS
Story 1	0.222	0.5636	0.5033
Story 2	0.5234	0.360	0.3302

Pair-wise agreement was calculated in terms of Pearson's *R* and Spearman's rank for all pairs of annotators for both stories in each condition (binary, ordinal and continuous scales). Overall, the mean of the pairwise correlation coefficients between annotators for each scale type for both stories indicate a moderate mean positive correlation, with the exception of the Story 1 binary scale which exhibits a low positive correlation. Figure 11 displays the distributions of all pairwise correlation coefficients per scale type, per story. For Story 1 in figure 11a, the mean Pearson's correlation coefficients for the continuous ($M = 0.54$, $SD = 0.16$) and ordinal ($M = 0.63$, SD

= 0.15) scales, and their respective mean Spearman's coefficients (continuous $M = 0.61$, $SD = 0.13$; ordinal $M = 0.62$, $SD = 0.16$), are all between 0.5 and 0.65. This supports a moderately strong positive correlation. However, the average Pearson's correlation for the binary scale was weak to low-moderate ($M = 0.27$, $SD = 0.22$).
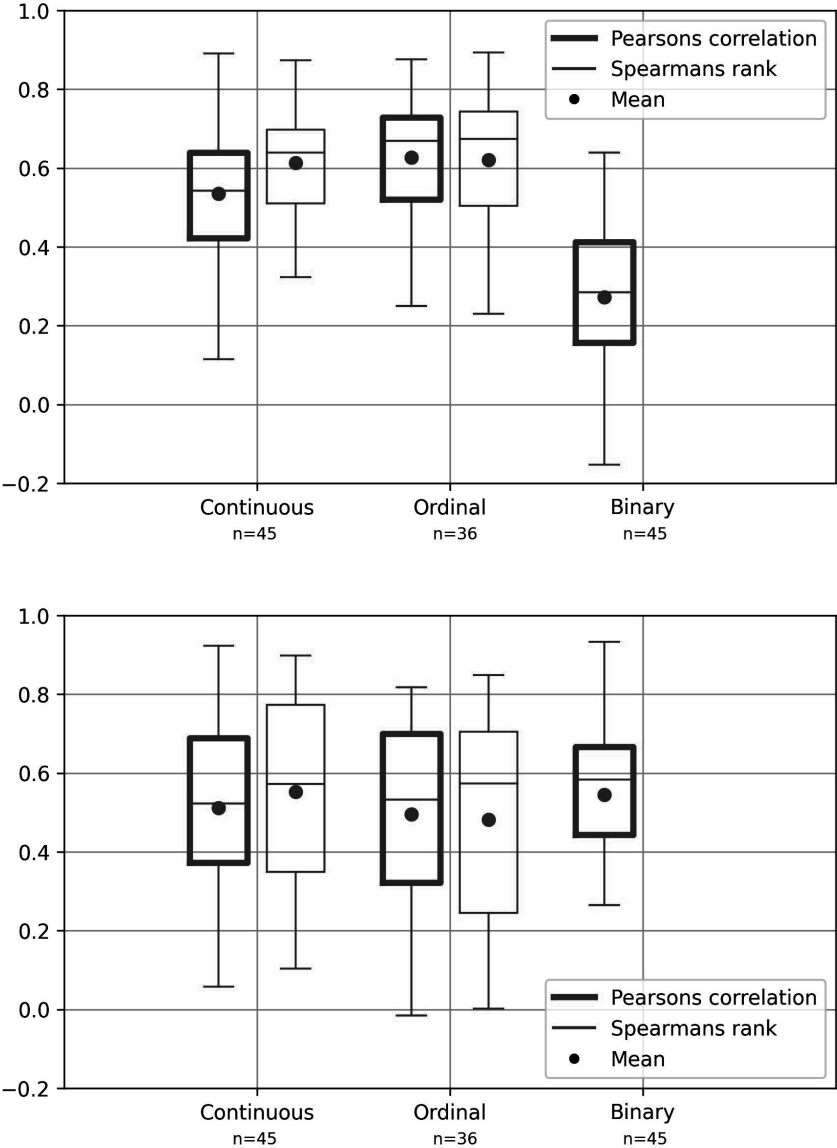
Fig. 11. Distributions of all pairwise correlations coefficients per scale type per story; **Fig. 11a (top).** Story 1; **Fig. 11b (bottom).** Story 2.

For Story 2 in figure 11b, the mean Pearson's correlation coefficients for the continuous ($M = 0.51$, $SD = 0.22$), ordinal ($M = 0.5$, $SD = 0.24$), and binary ($M = 0.55$, $SD = 0.18$) scales all indicate a moderate positive correlation. This is also the case for the mean Spearman's correlation coefficients (continuous $M = 0.55$, $SD = 0.24$; ordinal $M = 0.48$, $SD = 0.26$).

In terms of raw means, the continuous and ordinal scales exhibit higher annotator agreement through higher mean pairwise correlation scores than the binary scale for Story 1, while there is little difference between scales for Story 2. To assess whether any of the distributions are statistically significantly different, independent t-tests were performed.[5] An alpha level of 0.01 is used for all statistical tests unless otherwise stated. There are significant differences between the Pearson's correlation score distributions for Story 1 between all three scales: continuous ($M = 0.54$, $SD = 0.16$) and ordinal ($M = 0.63$, $SD = 0.15$) with $t(88) = -2.71$ $p < 0.01$, continuous and binary ($M = 0.27$, $SD = 0.22$) with $t(88) = 6.35$ $p < 0.01$ and ordinal and binary with $t(88) = 8.77$ $p < 0.01$. However, there is no significant difference between the continuous and ordinal Spearman's correlation score distributions for Story 1. There are also no significant differences between the Pearson's or Spearman's distributions for Story 2 between any of the conditions.

5.2.1   Per-panel analyses

To further investigate patterns of higher and lower agreement for panels between and within each condition, we standardise the raw scores per participant by using Z-score transformation and compute the mean and standard deviations of these Z-scores per panel for the continuous and ordinal scales. We compute the percentage agreement for the binary scale.

Within the continuous and ordinal scales, the most disagreed upon panels exhibit similar standard deviations – the five most disagreed upon panels per story for the continuous scale have standard deviations between 0.88–1.1 for Story 1, and between 0.87–1.22 for Story 2. This was similar for the ordinal scale, with Story 1 between 0.83–1.0 and between 0.84–0.96 for Story 2. These panels tend to feature images with objects, scenery and/or characters in the foreground with little visual detail in the background. For the binary scale, Story 1 has three panels that show 50% agreement (the lowest possible score, an even split between annotators) and seven more panels with 60% agreement, while Story 2 only has one panel with 50% agreement and four panels with 60% agreement. The images that exhibit the most disagreement for the binary scale are similar to those that do so for the continuous and ordinal scales.

The most agreed upon panels also produce similar standard deviations for both continuous and ordinal scales. The five most agreed upon panels for the continuous scale have standard deviations between 0.17–0.24 for Story 1, and 0.27–0.44. For the ordinal scale, Story 1 has standard deviations between 0.18–0.24 and 0.17–0.35 for Story 2. The binary scale has

five out of thirty-two unanimously agreed upon panels in Story 1, and eleven of thirty-one panels for Story 2. The panels that exhibit high agreement across all scales are overwhelmingly images of characters in conversation against a solid colour background. Exceptions to this occur for several panels; Panel 31 in Story 1 appears to elicit disagreement due to a prominent foreground aspect in the image. Panel 10 in Story 2 shows only a small number of markings in the background which cause disagreements in interpretation. Both these types of disagreements are also found, and further described, in Experiment 1.
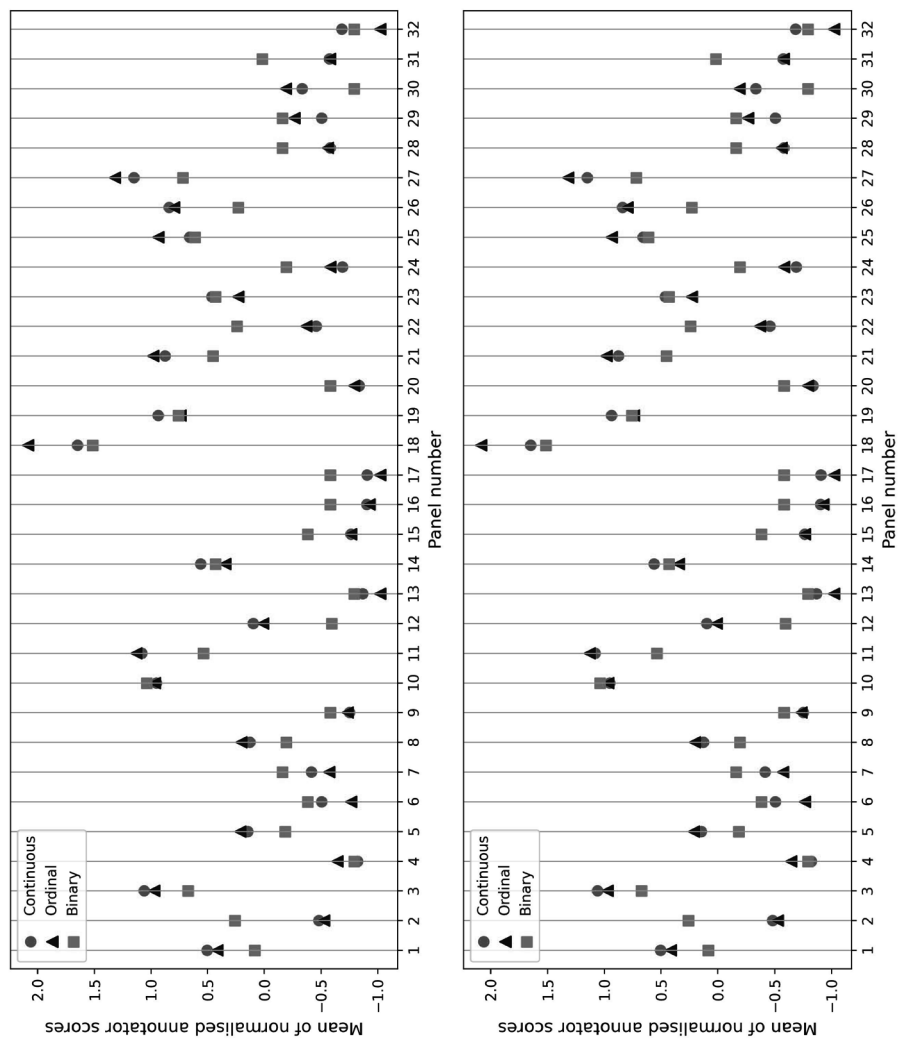
Fig. 12. Mean of annotators values (Z-score normalised) per panel;
Fig. 12a (left). Story 1; **Fig. 12b (right).** Story 2.

The overlap in the most agreed/disagreed panels appears to support generally consistent perceptions of information amounts. We plot the Z-score normalised mean of all annotators scores per scale per panel, which is displayed in figure 12, to further investigate agreement between conditions. The per-panel means for each scale of Story 1 in figure 12a show that the binary scale mean score often deviates from the relatively close ordinal and continuous mean scores, while in Story 2 the binary scale is more closely grouped to the other two scales, as shown in figure 12b.

It also appears that agreement is strongest for panels perceived to be towards the no information end of the scales, and becomes weaker as more background information is perceived. We test this idea statistically by correlating the Z-score normalised mean plotted against the standard deviations per scale, per story, as per figure 13. Pearson's correlations for each scale distribution in figure 13 for Story 1 show that there is a strong positive correlation between mean of annotator score and its standard deviation for the continuous ($r(30) = 0.86$, $p < 0.001$), and binary ($r(30) = 0.75$, $p < 0.001$) scales, and a moderately strong positive correlation for the ordinal scale ($r(30) = 0.55$, $p = 0.001$) – this means that the higher the assigned score, the higher the disagreement between annotators – and conversely, the lower the assigned score, more agreement is observed between annotators. This relationship is only evident for the continuous scale in Story 2 which shows a low-moderate positive correlation ($r(29) = 0.4$, $p = 0.025$, with an alpha level of 0.5), while the ordinal ($r(29 = 0.29$, $p = 0.12$)) and binary ($r(29) = 0.25$, $p = 0.18$) scales are not correlated significantly.
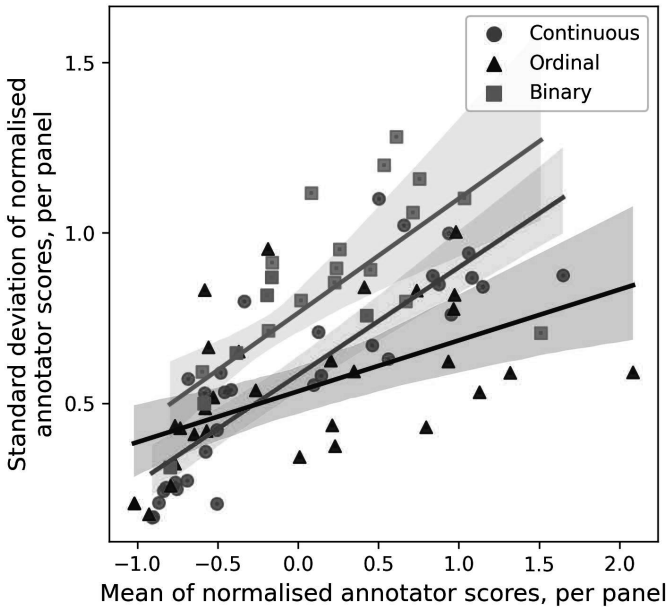
Fig. 13. Z-score normalised mean vs. standard deviation per panel in Story 1.

5.2.2   Annotator reliability

Each condition contained a few annotators who consistently disagreed with most others. However, most annotators were consistent and highly reliable. Therefore, a large range of interpretations of the task by annotators seems to be the primary cause for disagreement. Visual language fluency, as measured by the mean VLFI scores shown in Table 4, does not predict the mean agreement differences in Story 1 – while the participants in the binary condition had a lower VLFI (5.9) than the ordinal condition (10.9), they had a higher VLFI than the continuous condition (4.9), despite having significantly lower mean agreement; for example, see Annotator 3 in the continuous scale condition in Story 2.[6]

5.3 Discussion

Owing to its superior correlations in Story 1, the ordinal scale is tentatively understood as the scale most suitable for achieving better inter-annotator agreement on background information amount when compared to the binary classification. However, the KA results show the background information amount concept does not reach a threshold of agreement for implementation in further comics analysis, even with a finer-grained scale. Similar results for the binary scales between experiments reinforces that reader judgments are consistent between word-of-mouth and crowd-sourced recruitment.

However, the moderate positive mean correlations between annotator pairs for the continuous and ordinal scales convey a consistent relation between annotator perceptions of visual background information amount in a comics panel. The ordinal scale appears to be relatively the most robust, and exhibits the highest inter-annotator agreement for Story 1, but there is no significant difference between the ordinal, continuous and binary scales for Story 2. The binary scale results show a discrepancy in agreement between Story 1 and Story 2, with Story 1 causing more disagreement, consistent with the results from Experiment 1. The KA results for ordinal and continuous scales, however, also show an opposite difference in agreement, with Story 1 exhibiting higher agreement than Story 2. We can conclude that while a particular scale may produce sufficient agreement for stories of a particular style or narrative, no scale is obviously generalisable across comics even within the same publication.

Finally, the notion of the empty or 'dropped' background is again shown to have some conceptual legitimacy, as annotator agreement was highest towards the empty side of the spectrum. This was particularly evident for Story 1, where the ordinal scale showed a less severe increase in disagreement as the mean level of detail increased. This suggests that using a scale of background information amount can be more reliable for empty panel identification but not for gradations of some information – perhaps

future work can explore whether the ordinal scale can be used to 'collapse' back into a binary classification, where the choice of 1 indicates empty, and a choice of 2–5 indicates non-empty.

6.   Discussion

Both experiments, as well as previous work on other segmentation and classification tasks, overall demonstrate that this methodology produces reliable quantified inter-subjective interpretations. The effectiveness of the method is supported in several ways. First, annotators were shown to be more-or-less reliable with only some annotators substantially disagreeing with others. This suggests that disagreements are primarily due to the annotation judgment conceptualisation itself. Second, both experiments produced comparable results for Story 1 and Story 2 using the binary scale. This consistency in annotator judgments across two implementations of the experiment indicates that the practical annotation method itself is sound. It additionally supports the claim that there is no clear benefit between word-of-mouth over crowdsourced recruiting. Panel segmentation with background information amount judgment also seems to be an appropriate task for recruiting everyday readers – a very complex concept requiring theoretical knowledge or expertise is likely to need hand-picked annotators. Segmentation tasks overall tend to attain higher agreement than categorisation or labelling tasks, although agreement thresholds were met for such tasks in previous work. Finally, evidence of minor improvement from the binary to the ordinal scale for the background information location judgement between Experiments 1 and 2 shows that task refinement per the MAMA (Model-Annotate-Model-Annotate) cycle can indeed develop a more robust annotation scheme.

On reflection, the conception of the proposed unit itself determines whether a segmentation-attribute pair is a feasible construction. Sub-representational and lower-level perceptual elements do not appear to be appropriate for this method – recall that many of the disagreements from the experiments were due to different interpretations of background areas with non-representational image components such as neutral tones. However, these disagreements point to places where addressing or incorporating sub-representational elements may be beneficial. Since the overall image area for making a judgment is agreed upon through high-level panel segmentation agreement, perhaps adding a measure of sub-representational marking amounts using computer vision, such as image contrast or number of lines, can be added to panel segments as additional features. Exploring these attributes together may supplement or bolster inter-annotator results – for instance, checking whether correlations of more or less background information in Experiment 2 correspond with certain sub-representational structural features. On the other hand, segmentation can be useful for identifying parts of wholes in higher-level rep-

resentations. Either way, this refinement methodology provides an initial empirical anchor to the validity of the concept.

7. Future work

There are several shortcomings and limitations with this work that motivate further development of this methodology. First, only a small-sized corpus containing comics exhibiting only one artistic style and genre is used. This limits the findings to a narrow range of comics that were all created at the same point in time, stem from the same cultural background, and show similar artistic styles. Although the findings from the experiments are not generalisable, they do reveal an effectiveness of the presented methodology. Useful disagreements were found even within a small corpus of similar comics, shown here as well as in previous work. Future work may expand on the types and amounts of comics annotated, but it is beneficial to find agreements and parse disagreements in similar stories before implementing the units across comics of different styles, genres, and cultural origins.

Second, the current CAT set-up of implementing bounding box delimiters for annotation are likely to be inappropriate for many segmentations. Bounding boxes were used in this initial implementation to facilitate efficient annotation, and for future comparisons between reader-made segmentations and segmentations found in ground truth corpora for computational analysis. However, many comics use much more diverse panel structures. Many areas of page use, such as sound effects, character outlines, and even sub-representational markings, do not appear to fit well within bounding boxes. A good direction for future work is providing more flexible segmentation tools to the prototype CAT, such as allowing prompting annotators to 'colour in' areas of the comics page as described in Abusch (2012). In fact, a good direction for future work is developing the CAT to be flexible in allowing researchers to define and implement novel annotation tasks using a set of in-built segmentation and classification tools.

Turning to the background information concept itself, further segmentation or clarification of what area within a panel is meant to be judged is beneficial. Having annotators make a segmentation distinction between traditional understandings of foreground and background, for instance, would help to explain the task. Furthermore, using new terms instead of 'empty', 'detailed' and 'background' may reduce confusion. Nevertheless, since the 'no information' end of the scales in Experiment 2 produced relatively higher agreement, the concept of 'dropped' background has merit to pursue in future work. Eventually applying a robust concept of dropped background to further studies, such as determining their frequency across comics, their meaning, and other relations between use of background space and other features would be informative.

There are also many opportunities for future work using verified segment-attribute pairs using this method. Other examples of how this metho-

dology can be applied include: assessing perceptions along a spectrum of iconicity to representation according Peircean semiotics (how much does a particular visual element resemble what it is signifying, according to annotators?), delimiting areas across a story that make up character discourse chains (Tseng et al. 2018) and other studies of discourse cohesion, and quantitatively modelling amounts of information used by readers to generate meaning from entailment that were qualitatively modelled by Wildfeuer (2019), among others.

Finally, this methodology is conducive to creating comics corpora with multiple reader perceptions per delineated unit. Further analyses of annotator disagreement may indicate cases of deliberate ambiguity or vagueness from the comics' author, and offer a quantified measure of ambiguity. In a practical sense, this methodology produces JSON formatted information about a unit's size and associated attributes according to a number of annotators. This data is therefore useful for further computational analyses of the distribution and constraints of information across comics. A direction for future work is building models which predict selected elements from others. Corpora with empirically verified ground truths could also be useful for automatic content extraction, especially in cases with ambiguity in the comics narrative.

8.   Conclusion

The purpose of these two experiments was to provide an example of a general methodology for testing inter-reader interpretations on aspects of comics. The process begins with developing a preliminary concept for a segmentation-classification pair. The concept can be a low, sub-part, or a high-level categorisation, and is typically based on a theory or intuition about comics structure. The concept is translated into an annotation task to test whether everyday readers make the same judgments on areas of comics pages. While the results from previous work showed high agreement for some salient aspects of comics, with segmentation often exhibiting higher agreement than attributions, the background information task discussed here required refinement. The refinement and re-testing follows the spirit of a MAMA (Model-Annotate-Model-Annotate) cycle (Pustejovsky et al. 2017). Achieving high agreement between annotators supports that a proposed unit is well conceptualised, while disagreements are instructive for re-conceptualising stronger units, and may also reveal intended ambiguity or vagueness within the comics narrative. We hope that an efficient annotation methodology allows for developing robust and empirically verified units for future research across comics studies.

Notes

1   The current version of the CAT is available to use at *GitHub* (https://comicsanno-tationtool.github.io) and the source code can be downloaded there (https://github.com/ComicsAnnotationTool/comicsannotationtool.github.io) [retrieved April 16, 2024].
2   The panel segmentations are outlined in red, the character segmentations in purple, and text section segmentations in green. Due to printing constraints the colours are not able to be shown in Figure 2.
3   See *GitHub* (https://github.com/le300/CAT_Annotation_Experiment_1 [retrieved April 17, 2024]) for heatmaps depicting the KA agreement between each annotator pair for each story.
4   See *GitHub* (https://github.com/le300/CAT_Annotation_ Experiment_2 [retrieved April 17, 2024]) for the full annotation scheme, which offers detailed instructions and examples.
5   All assumptions were met to perform independent t-tests between each distribution (data groups are independent, normally distributed, and exhibit similar variance), except for the ordinal scale data for Story 2. This distribution has fewer data points ($N = 36$) than the others ($N = 45$) and does not meet the homogeneity of variance assumption using Levene's test between ordinal and binary data $F(1,79) = 5.74$, $p = 0.019$, and is not normally distributed according to a Shapiro-Wilk test ($W = 0.94$, $p = 0.045$). Non-parametric Welch's and Mann-Whitney U tests were performed between the distributions for Story 2, with all results showing no significant differences between scale types.
6   Heatmaps depicting the agreement between each annotator for each story are available at *GitHub* (https://github.com/le300/CAT_Annotation_Experiment_2 [retrieved April 17, 2024]).

Bibliography

Abusch, Dorit (2012). Applying discourse semantics and pragmatics to co-reference in picture sequences. *Proceedings of Sinn und Bedeutung*, 17, 9–25.

Alves, Tiago, Ana Simoes, Rui Figueiredo, Marco Vala, Ana Paiva and Ruth Aylett (2008). So tell me what happened: Turning agent-based interactive drama into comics. *Proceedings of the 7th International Joint Conference on Autonomous Agents and Multiagent Systems* 3, 1269–1272.

Artstein, Ron and Massimo Poesio (2008). Inter-coder agreement for computational linguistics. In: Hwee Tou Ng (ed.). *Computational Linguistics* 34, 4, 555–596.

Asher, Nicholas and Alex Lascarides (2003). *Logics of Conversation*. Cambridge: Cambridge University Press.

Augereau, Olivier, Motoi Iwata and Koichi Kise (2018). A survey of comics research in computer science. *Journal of Imaging* 4, 7, 87, 1–19.

Bateman, John A., Francisco OD Veloso, Janina Wildfeuer, Felix Hiu Laam Cheung and Nancy Songdan Guo (2017). An open multilevel classification scheme for the visual

layout of comics and graphic novels: Motivation and design. *Digital Scholarship in the Humanities* 32, 3, 476–510.

Bateman, John A. and Janina Wildfeuer (2014a). A multimodal discourse theory of visual narrative. *Journal of Pragmatics* 74, 180–208.

Bateman, John A. and Janina Wildfeuer (2014b). Defining units of analysis for the systematic analysis of comics: A discourse-based approach. *Studies in Comics* 5, 2, 373–403.

Caldwell, Joshua (2012). Comic panel layout: A peircean analysis. *Studies in Comics* 2, 2, 317–338.

Cohn, Neil (2012). Comics, linguistics, and visual language: The past and future of a field. In: Frank Bramlett (ed.). *Linguistics and the Study of Comics*. London: Palgrave Macmillan, 92–118.

Cohn, Neil (2013a). *The Visual Language of Comics: Introduction to the Structure and Cognition of Sequential Images*. London: Bloomsbury Academic.

Cohn, Neil (2013b). Visual narrative structure. *Cognitive Science* 37, 3, 413–452.

Cohn, Neil (2014). The visual language fluency index: A measure of "comic reading expertise". *Visual Language Lab*. URL: http://www.visuallanguagelab.com/resources [retrieved October 12, 2020].

Cohn, Neil (2015). Narrative conjunction's junction function: The interface of narrative grammar and semantics in sequential images. *Journal of Pragmatics* 88, 105–132.

Cohn, Neil (2018a). In defense of a "grammar" in the visual language of comics. *Journal of Pragmatics* 127, 1–19.

Cohn, Neil (2018b). Combinatorial morphology in visual languages. In: Geert Booij (ed.). *The Construction of Words*. Cham: Springer, 175–199.

Cohn, Neil and Marta Kutas (2017). What's your neural function, visual narrative conjunction? Grammar, meaning, and fluency in sequential image processing. *Cognitive Research: Principles and Implications* 2, 1, 1–13.

Comic book plus: No.1 free & legal (2022). URL: http://www.comicbookplus.com [retrieved September 30, 2022].

Dubray, David and Jochen Laubrock (2019). Deep CNN-based speech balloon detection and segmentation for comic books. *2019 International Conference on Document Analysis and Recognition (ICDAR)*. IEEE, 1237–1243.

Dunst, Alexander and Rita Hartel (2018). The quantitative analysis of comics: Towards a visual stylometry of graphic narrative. In: Alexander Dunst, Jochen Laubrock, Janina Wildfeuer (eds.). *Empirical Comics Research: Digital, Multimodal, and Cognitive Methods*. New York: Routledge, 43–61.

Edlin, Lauren and Joshua Reiss (2021). *An empirically-based spatial segmentation and coreference annotation scheme for comics*. New York, NY: Association for Computing Machinery, 8, 1–8.

Eisner, Will (2008). *Comics and Sequential Art: Principles and Practices from the Legendary Cartoonist*. New York: WW Norton & Company.

Everingham, Mark, S. M. Ali Eslami, Luc Van Gool, Christopher K. I. Williams, John Winn and Andrew Zisserman (2015). The PASCAL visual object classes challenge: A retrospective. *International Journal of Computer Vision* 111, 1, 98–136.

Fast Krippendorf (2017). URL: https://github.com/pln-fing-udelar/fast-krippendorff.git. San Francisco, CA: GitHub [retrieved November 13, 2022].

Gauthier, Guy (1976). Les peanuts: un graphisme idiomatique. *Communications* 24, 1, 108–139.

Geurts, Bart, David I. Beaver and Emar Maier (2020). Discourse representation theory. In: Edward N. Zalta (ed.). *The Stanford Encyclopedia of Philosophy*. Spring 2020 ed. Stanford: Metaphysics research Lab, Stanford University.

Grill, Thomas (2017). *Python Implementation of Krippendorff's Alpha – Inter-Rater Reliability*. URL: https://github.com/grrrr/krippendorff-alpha.git. San Francisco, CA: GitHub [retrieved November 13, 2022].

Groensteen, Thierry (2007). *The System of Comics*. Mississippi: University Press of Mississippi.

Guérin, Clement, Christophe Rigaud, Karell Bertet and Arnaud Revel (2017). An ontology-based framework for the automated analysis and interpretation of comic books' images. *Information Sciences* 378, 109–130.

Guynes, Sean A. (2014). Four-color sound: A peircean semiotics of comic book onomatopoeia. *Public Journal of Semiotics* 6, 1, 58–72.

Ito, Kota, Yusuke Matsui, Toshihiko Yamasaki and Kiyoharu Aizawa (2015). Separation of manga line drawings and screentones. *Eurographics (Short Papers)*, 73–76.

Kamp, Hans (1981). A theory of truth and semantic representation. In: Paul Portner and Barbara Partee (eds.). *Formal Semantics: The Essential Readings*. Oxford: Blackwell Publishers Ltd, 189–222.

Kamp, Hans, Josef Van Genabith and Uwe Reyle (2011). Discourse representation theory. In: Dov M. Gabbay and Franz Guenthner. *Handbook of Philosophical Logic* 15. Dordrecht: Springer, 125–394.

Krippendorff, Klaus (2011). *Computing Krippendorf's Alpha-Reliability*. URL: http://www.asc.upenn.edu/Krippendorff/2007. University of Pennsylvania: Departmental papers [retrieved October 8, 2022].

Kurlander, David, Tim Skelly and David Salesin (1996). Comic chat. *Proceedings of the 23rd Annual Conference on Computer Graphics and Interactive Techniques*. New York: Association for Computing Machinery, 225–236.

Laubrock, Jochen and Alexander Dunst (2020). Computational approaches to comics analysis. *Topics in Cognitive Science* 12, 1, 274–310.

Liu, Xueting, Chengze Li and Tien-Tsin Wong (2017). Boundary-aware texture region segmentation from manga. *Computational Visual Media* 3, 1, 61–71.

Maier, Emar (2019). Picturing words: The semantics of speech balloons. *Proceedings of the 22nd Amsterdam Colloquium*. Amsterdam: ILLC, 584–592.

Maier, Emar and Sofia Bimpikou (2019). Shifting perspectives in pictorial narratives. In: Uli Sauerland and Stephanie Solt (eds.). *Proceeding of Sinn und Bedeutung* 23, 2, 91–106.

Magnussen, Anne (2000). The semiotics of C. S. Peirce as a theoretical framework for the understanding of comics. *Comics & Culture: Analytical and Theoretical Approaches to Comics* 193–208.

Mao, Xiangyu, Xueting Liu, Tien-Tsin Wong and Xuemiao Xu (2015). Region-based structure line detection for cartoons. *Computational Visual Media* 1, 1, 69–78.

McCloud, Scott (1993). *Understanding Comics: The Invisible Art*. Northampton, MA: Kitchen Sink Press.

McCloud, Scott (2006). *Making Comics: Storytelling Secrets of Comics, Manga and Graphic Novels*. New York: Harper New York.

Meesters, Gert (2017). Semiotics and linguistics. In: Matthew Smith and Randy Duncan (eds.). *The Secret Origins of Comics Studies*. New York and Oxon: Routledge, 100–104.

Miller, Ann (2017). Formalist theory: Academics. *The Secret Origins of Comics Studies*. New York: Routledge, 62–76.

Nguyen, Nhu-Van, Christophe Rigaud and Jean-Christophe Burie (2017). Comic characters detection using deep learning. *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)* 3, 41–46.

Pang, Xufang, Ying Cao, Rynson W. H. Lau and Antoni B. Chan (2014). A robust panel extraction method for manga. *Proceedings of the 22nd ACM international conference on Multimedia*, 1125–1128.

Pederson, Kaitlin and Neil Cohn (2016). The changing pages of comics: Page layouts across eight decades of American superhero comics. *Studies in Comics* 7, 1, 7–28.

Peer, Eyal, Laura Brandimarte, Sonam Samat and Alessandro Acquisti (2017). Beyond the turk: Alternative platforms for crowdsourcing behavioral research. *Journal of Experimental Social Psychology* 70, 153–163.

Prolific: A Higher Standard of Online Research (2022). URL: http://www.prolific.co [retrieved September 30, 2022].

Pustejovsky, James, Harry Bunt and Annie Zaenen (2017). Designing annotation schemes: From theory to model. In: Nancy Ide and James Pustejovsky (eds.). *Handbook of Linguistic Annotation*. Dordrecht: Springer, 21–72.

Qin, Xiaoran, Yafeng Zhou, Zheqi He, Yongtao Wang and Zhi Tang (2017). A faster r-cnn based method for comic characters face detection. *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)* 1, 1074–1080.

Rezatofighi, Hamid, Nathan Tsoi, JunYoung Gwak, Amir Sadeghian, Ian Reid and Silvio Savarese (2019). Generalized intersection over union: A metric and a loss for bounding box regression. *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 658–666.

Rigaud, Christophe and Jean-Christophe Burie (2018). Computer vision applied to comic book images. In: Alexander Dunst, Jochen Laubrock, Janina Wildfeuer (eds.). *Empirical Comics Research: Digital, Multimodal, and Cognitive Methods*. New York: Routledge, 104–124.

Rigaud, Christophe, Jean-Christophe Burie and Jean-Marc Ogier (2017). Segmentation-free speech text recognition for comic books. *2017 14th IAPR International Conference on Document Analysis and Recognition (ICDAR)* 3, 29–34.

Rigaud, Christophe, Clement Guérin, Dimosthenis Karatzas, Jean-Christophe Burie and Jean-Marc Ogier (2015). Knowledge-driven understanding of images in comic books. *International Journal on Document Analysis and Recognition (IJDAR)* 18, 3, 199–221.

Rosebrock, Adrian (2016). *Intersection over Union (IoU) for Object Detection*. URL: https://pyimagesearch.com/2016/11/07/intersection-over-union-iou-for-object-detection/ [retrieved December 2, 2020].

Saraceni, Mario (2003). *The Language of Comics*. London: Routledge.

Schalley, Andrea C (2019). Ontologies and ontological methods in linguistics. *Language and Linguistics Compass* 13, 11, 1–19.

Shamir, Ariel, Michael Rubinstein and Tomer Levinboim (2006). Generating comics from 3d interactive computer graphics. *IEEE Computer Graphics and Applications* 26, 3, 53–61.

Soares De Lima, Edirlei, Bruno Feijo, Antonio L. Furtado, Simone Diniz Junqueira Barbosa, Cesar T. Pozzer and Angelo E.M. Ciarlini (2013). Non-branching interactive comics. In: Dennis Reidsma, Haruhiro Katayose and Anton Nijholt (eds.). *International Conference on Advances in Computer Entertainment Technology*. Boekelo: Springer, 230–245.

Szawerna, Michał (2013). A peircean characterization of the semiotic properties of selected visual signs found in comics. *Ars Aeterna* 5, 1, 51–72.

Szeliski, Richard (2020). *Computer Vision: Algorithms and Applications*. 2nd ed. URL: http://szeliski.org/Book/ [retrieved December 13, 2020].

Tseng, Chiao-I, Jochen Laubrock and Jana Pflaeging (2018). Character developments in comics and graphic novels: A systematic analytical scheme. In: Alexander Dunst, Jochen Laubrock and Janina Wildfeuer (eds.). *Empirical Comics Research: Digital, Multimodal, and Cognitive Methods*. New York: Routledge, 154–175.

Tufis, Mihnea and Jean-Gabriel Ganascia (2019). Crowdsourcing comics annotations. In: Alexander Dunst, Jochen Laubrock and Janina Wildfeuer (eds.). *Empirical Comics Research: Digital, Multimodal, and Cognitive Methods*. New York: Routledge, 85–103.

Virtanen, Pauli, Ralf Gommers, Travis E. Oliphant, Matt Haberland, Tyler Reddy, David Cournapeau, Evgeni Burovski, Pearu Peterson, Warren Weckesser, Jonathan Bright, Stéfan J. van der Walt, Matthew Brett, Joshua Wilson, K. Jarrod Millman, Nikolay Mayorov, Andrew R. J. Nelson, Eric Jones, Robert Kern, Eric Larson, C.J. Carey, İlhan Polat, Yu Feng, Eric W. Moore, Jake VanderPlas, Denis Laxalde, Josef Perktold, Robert Cimrman, Ian Henriksen, E. A. Quintero, Charles R. Harris, Anne M. Archibald, Antônio H. Ribeiro, Fabian Pedregosa, Paul van Mulbregt and SciPy 1.0 Contributors (2020). SciPy 1.0: Fundamental algorithms for scientific computing in Python. *Nature Methods* 17, 3, 261–272.

Wildfeuer, Janina (2019). The inferential semantics of comics: Panels and their meanings. *Poetics Today* 40, 2, 215–234.

Yus, Francisco (2008). Inferring from comics: A multi-stage account. *Quaderns de Filologia: Estudis de Comunicacio* 3, 223–249.

Sources

Kida, Fred (art) and unknown (wri) (1957–1958). *My Robot Plants*. New York: Harvey Comics.

Kirby, Jack (wri/art) (1957–1958). *The Cadmus Seed!* New York: Harvey Comics.

Kirby, Jack (wri/art) (1957–1958). *The Fourth Dimension is a Many Splattered Thing!* New York: Harvey Comics.

Oleck, Jack (wri) and Doug Wildey (art) (1957–1958). *The Man Who Never Lived*. New York: Harvey Comics.

Watterson, Bill (wri/art) (1985–1995). *Calvin and Hobbes*. Andrews McMeel Universal syndicate.

Quinlan Sr., Charles (art) and unknown (wri) (1941–1942). *Cat-Man Comics: …Introducing The Kitten*. Helnit Publishing Co. Inc.

Image Sources

Fig. 1a. Watterson (1987). URL: https://www.gocomics.com/calvinandhobbes/1987/06/05 [retrieved May 10, 2024].
Fig. 1b. Kirby (1975: 10).
Fig. 3. Quinlan (1941: 3).
Fig. 6a. Oleck and Wildey (1957, Story 4: 4).
Fig. 6b. Kirby (1957, Story 2: 5).
Fig. 7. Kirby (1957, Story 1: 2).
Fig. 8. Kirby (1957, Story 1: 4).
Fig. 9. Kida (1957, Story 3: 5).

Lauren Edlin, PhD Student
Queen Mary University of London
Electronic Engineering and Computer Science
Peter Landin Building
Mile End Rd, Bethnal Green
London E1 4FZ
United Kingdom
E-Mail: l.edlin@qmul.ac.uk

Prof. Dr. Joshua Reiss
Queen Mary University of London
Electronic Engineering and Computer Science
Peter Landin Building
Mile End Rd, Bethnal Green
London E1 4FZ
United Kingdom
E-Mail: joshua.reiss@qmul.ac.uk

Visualising an Oral Epic: Lobačev's Comic Book Tsar Dušan's Wedding*

Miloš Tasić, University of Niš and Dušan Stamenković, Södertörn University

Summary. This study analyses visual and verbal material found in Đorđe Lobačev's comic book *Tsar Dušan's Wedding* (orig. *Ženidba Cara Dušana* [1989]) using a set of tools coming from the domains of intersemiotic translation, intermediality, and adaptation. The comic book is based on the Serbian oral epic ballad *Dušan's Wedding* (orig. *Ženidba Dušanova* [1815/1975]), which focuses on the hero wedding theme. The study will try to present verbal and graphic devices used to transfer the epic narrative to the medium of comics and compare them to the ones used in the original story. It will also address the aspect of the epic ballad structure and its transposition into the realm of comics. The analysis will take into account the historical context of the comic book in question and compare the findings with some of the results of contemporary comics studies dealing with certain graphic devices employed in comics. These findings reveal that Lobačev's comic book contains certain graphic flourishes related to describing motion, speed and spatiality, but lacks the ones used to depict affect, which is in line with basic traits of oral epic poetry.

Keywords. Comics, oral epic, wedding theme, intersemiotic translation, medial transposition, adaptation

Zusammenfassung. Dieser Artikel untersucht das visuelle und sprachliche Material in Đorđe Lobačevs Comic *Kaiser Dušans Hochzeit* (Orig. *Ženidba Cara Dušana* [1989]). Die Analyse folgt Methoden aus der Forschung zur intersemiotischen Übersetzung, zur Intermedialität und zur Adaptation. Das Comicheft basiert auf der serbischen epischen Volksballade *Dušans Hochzeit* (Orig. *Ženidba Dušanova* [1815/1975]), die die Hochzeit des Helden thematisiert. Der Artikel stellt sprachliche und grafische Mittel dar, die die epische Erzählung ins Medium des Comics übertragen, und vergleicht sie mit denjenigen des Originals. Die Struktur der Volksballade wird ebenso diskutiert wie ihre Übertragung auf und in den Comic. Die Analyse zieht auch den historischen Kontext des betreffenden Comichefts in Betracht und vergleicht die Ergebnisse mit einigen gegenwärtigen Ansätzen der Forschung zu grafischen Verfahren im Comic. Die Ergebnisse zeigen, dass Lobačevs Comic bestimmte grafische Verzierungen enthält, die mit der

Beschreibung von Bewegung, Geschwindigkeit und Räumlichkeit zusammenhängen. Es fehlen jedoch diejenigen, die zur Darstellung von Affekt verwendet werden, was im Einklang mit den grundlegenden Merkmalen der mündlichen epischen Dichtung steht.

Schlüsselwörter. Comics, mündliche Epik, Hochzeit, intersemiotische Übersetzung, mediale Transposition, Adaptation

1.   Introduction: Study rationale, focus and outline

Using findings from the realms of intermediality, intersemiotic translation, and adaptation, this study analyses visual and verbal material found in Đorđe Lobačev's[1] comic book *Tsar Dušan's Wedding* (orig. *Ženidba Cara Dušana* [1989]), which is based on the Serbian oral epic of a similar name – *Dušan's Wedding* (orig. *Ženidba Dušanova*[2] [1815/1975]) – and focuses on the hero wedding theme. It belongs to the Pre-Kosovo cycle (orig. *Pretkosovski ciklus*), which includes epics about events that predate the Battle of Kosovo (1389). The study will try to present verbal and graphic devices used to transfer the epic narrative to the medium of comics and compare them to the devices used in the original story.

We will look at this artefact as an example of medial transposition, or transformation of one media product into another medium, understood as a concept of a narrative that exists in different media regardless of the novelty of the adaptation. The analysis will also take into account the historical context of the comic book in question (pre-WWII) and compare the findings with some of the research results of contemporary comics studies. The subject matter of Lobačev's work represents a visual rendition of a popular Serbian oral epic, which is in its own way a highly specific form of creative expression, as it follows a specific narrative formula, among other things. With this in mind, we will also look at the ways in which this formula is transferred from the realm of the epic ballad to the comic book. We will keep track of transmedial processes in portraying each of these aspects, with a special focus on how the coherence between the elements of the traditional oral epic formula was maintained in the process of transforming the examined epic into a comic book, which uses a different semiotic toolkit.

We will start by outlining the most notable differences between the two works, which include, among others, the modernisation of language and its adaptation to the assumed preferences of the target audience, the omission and addition of certain narrative elements, as well as changes in the tone of storytelling, specifically the lack of explicit violence in the comic book. We will also examine in detail how page layout is employed by the artist to propel the narrative and how it differs in rhythm to the strict decasyllabic structure of the source material, which was traditionally sung with the accompaniment of the g u s l e , a single-stringed instrument. Furthermore, particular attention will be paid to the use of colour in the comic book

(added in the 1989 special edition), primarily in the instances where it is utilised for emphasis, and the ways in which readers can recognise such artistic intentions. Preliminary results reveal that Lobačev's comic contains some graphic devices related to describing motion, speed and spatiality, but lacks the ones used to depict affect, which is in line with the basic traits of oral epic poetry. Exploring all of these aspects of the examined adaptation of a monomodal source material – as we will show – to a multimodal media product will allow us to draw parallels between the likely perception and interpretation of both versions.

2.   Theoretical framework

The present analysis is grounded in several theoretical constructs. First, we will take a look at Jakobson's (1959: 233) intersemiotic translation, as one of the three ways of interpreting a verbal sign, where words are translated using a nonverbal sign system, in our case images. Then, we will delve deeper into the relation between the different media of the original text and the comic book by examining the concept of intermediality, both in its wider sense, as any type of relation between two or more media, and its narrower sense, which focuses on "concrete medial configurations and their specific intermedial qualities" (Rajewsky 2005: 51). Finally, as this is, indeed, a case of adaptation of one media product into another (or we could even argue one media product into another into yet another, as we will see later on), the paper will turn to the adaptation process itself, with a particular emphasis on adapting other texts into the comics format. This eclectic approach will reflect the complexity that comes with the study of comics in the first place – many studies stress the overall heterogeneity of this research field and note that many theories and methods have been brought in and adapted for various purposes in comics studies (see Bramlett et al. 2017; Smith and Duncan 2017; Packard et al. 2019). We hope to go beyond the confines of a single and a purely linguistic perspective (for a criticism of narrow linguistics-based approaches, see Bateman and Wildfeuer 2014).

When discussing the interpretation of verbal signs, Jakobson (1959: 233) differentiates between three kinds of translation. The first is the process of interpreting verbal signs of one language by means of other verbal signs of that same language, called intralingual translation, a rewording usually through synonyms or circumlocution. If verbal signs are rendered through verbal signs of a different language, we are dealing with translation proper, or what Jakobson dubs "interlingual translation". If, however, verbal signs are interpreted by a system of nonverbal signs, for example, a visual language as in the present case, such a process is labelled "intersemiotic translation" or "transmutation". Interestingly, Eco (2003) believes that the term "translation" in Jakobson's third type is, in fact, metaphorical, thus he opts for "intersemiotic conversion" as a term that more accurately describes the process at hand. Furthermore,

recent research has drawn attention to the fact that translation between different semiotic systems does not need to be limited to linguistic systems, but may include transmutations between all kinds of media including cinema, theatre, visual arts and others (Dusi 2015: 182). This broadening of the term "intersemiotic translation", which now also encompasses not exclusively linguistic semiotic resources (e.g. Kourdis and Yoka 2014; O'Halloran et al. 2016), has led many to connect intersemiotic translation to other concepts from translation studies, such as adaptation, multimodality, audiovisual translation, etc.

For instance, an adaptation of one media product into another can be considered as an act of translation where one sign system undergoes intersemiotic transposition into a different sign system (Tsui 2012), or where different signs and conventions are used to transcode and transmutate (Hutcheon 2013). There are examples of intersemiotic translation actually yielding new forms perhaps better suited for conveying the messages found in original works: Perteghella (2019) illustrates this by using poetry podcasts that engage the listener in the orality and aurality of poetry, and filmic poetry or video poems that can bring forth the visual, iconic elements of poetry. Other examples include the application of digital technologies in the analysis of poems that increase our appreciation of poetry (Alghadeer 2014). Pârlog (2019) discusses the relation between intersemiotic translation and multimodality, stating that the latter is based on different types of signs and symbols and the ways in which they are intertwined, and concluding that intersemiotic translation can make knowledge more accessible to those that find it difficult to grasp it in its original form. Other applications of intersemiotic translation include audiovisual translation (Taylor 2020) and creating book illustrations (Pereira 2008) and book covers (Sonzogni 2011). O'Halloran et al. (2016) focus on the difficulties of analysing and modelling intersemiotic translation, emphasising the need for introducing various computational methods, such as multimodal annotation software, visualisation techniques or mathematical modelling, in order to tackle the issues that might arise in such procedures.

As, in this case, we are dealing with an adaptation that traverses different media, it is paramount that we take a closer look into the ways in which these media (and others as well) relate to each other. The analysed material, in fact, represents an extension of a narrative from the medium of oral poetry, via its written form recorded, most probably, centuries after the creation of the original work, to the medium of comics, which serves as our main point of interest in this study. With that in mind, what follows is a brief overview of the concept of intermediality, which can serve as background for the present analysis. Intermediality is a rather wide and variously defined term, yet we can start from Jensen's (2016: 972) understanding of it as "the interconnectedness of modern media of communication". Put this way, intermediality covers a number of different relations between media, and Jensen goes on to group them into three distinctive categories: (i) 'discursive intermediality', which implies simultaneous communication through several dis-

courses and modalities, (ii) 'material intermediality', where the focus is on different material vehicles of representation, and (iii) 'institutional intermediality', which describes the interplay between media as institutions.

Further, Schröter (2011) proposes four types of intermediality based on the discourse in which it can be found. Thus one may find (i) 'synthetic intermediality', in which different media are fused together, (ii) 'formal' or 'transmedial intermediality', where formal structures are present in different media, (iii) 'transformational intermediality', meaning one medium is represented through another, and (iv) 'ontological intermediality', where what comes first is the intermedial relation between two or more media rather than any one of those media themselves. Some authors even go so far as to describe the medium of comics as intermedial in its essence (Rajewsky 2005; Rippl and Etter 2013; Stein 2015), since it is a combination of two distinct media (text and image). However, if we turn to Kress and van Leeuwen (2001: 21–22), who define modes as semiotic resources employed in the realisation of discourses and interactions on the one hand, and media as material resources used in the creation of semiotic products and events on the other, we can conclude that the medium of comics is better explained as multimodal and not intermedial. There are at least two semiotic modes present in comics, the verbal (written text) and the pictorial (drawings), and Kukkonen (2011: 35) even argues that sequence can be taken as the third mode. Thus, meaning in comics is made and communicated via its multimodality, while the theory of intermediality can be applied to comics when a connection has been established with a different medium.

This connection can be seen more clearly in Rajewsky's (2005: 51–53) definition of intermediality in the narrow sense, where this concept is used in order to analyse texts and other media products. If we are to direct our attention to concrete medial configurations, as Rajewsky suggests, we can divide intermediality in this narrow sense into three subcategories. The first is 'm e d i a l t r a n s p o s i t i o n', where a product is transformed from one medium into another, as is the case in our analysis. The second is 'media combination', in which different media are integrated into a single product, such as in film, theatre, opera and others. Finally, the third subcategory is what Rajewsky calls 'intermedial references', e.g. references in film to painting, or in painting to photography, and so on. We will be focusing on the first subcategory, examining the analysed comic book both as an instance of intersemiotic translation and medial transposition.

The changes that are introduced by Đorđe Lobačev, which occur with the narrative being transformed from the "telling mode" of the ballad to the "showing mode" of the comic book (Hutcheon 2013: 22), are all parts of a d a p t a t i o n p r o p e r, a process which implies a number of informing and deforming constraints dictated by the intrinsic configuration of the comics medium (Gaudreault and Marion 2004: 58). One should also bear in mind that the original text had gone through further processes of adaptation prior to it being tackled by Lobačev, since the ballad itself had first existed as an oral epic passed down from generation to generation for centu-

ries before finally being recorded in the written form by Vuk Karadžić in 1815 (according to Lobačev 1989). It was this second text that served as the source material for the comic book. Gaudreault and Marion (2004: 61) suggest that any process of adaptation has to take into account the "incarnations" inherent in the encounter between a story and a medium that are related to the materiality of the media. They discuss the conceptual categories of m e d i a t i v i t y and narrativity, which are of great importance for the process of adaptation, with mediativity being particularly interesting for our approach as well. According to these authors, mediativity is "a medium's intrinsic capacity to represent – and to communicate that representation" (Gaudreault and Marion 2004: 66). Regardless of the fact that this idea comes from Gaudreault and Marion's treatment of stage performances, we believe that each medium has its own intrinsic capacity to represent something. Thus, mediativity is determined by the technical possibilities of the medium, or its internal semiotic configurations, which, in our case, would mean a combination of image and text. These specificities, in turn, result in the inevitable deletion or addition of material in the adaptation (Lefèvre 2007: 3–4), which will also be examined closely in the present study. In what follows, we will see how these theoretical underpinnings inform our analysis of the adaptation of an oral epic ballad into a comic book.

3.   Methodology and materials

The present study contains the analysis of Đorđe Lobačev's comic book *Tsar Dušan's Wedding* (orig. *Ženidba cara Dušana*), first published in 1938, with a particular insight into the process of adaptation of the original epic ballad that served as the source material for the comic book. The ballad itself, titled *Dušan's Wedding* (orig. *Ženidba Dušanova*), belongs to the Pre-Kosovo cycle (orig. *Pretkosovski ciklus*), which includes epics about events that predate the Battle of Kosovo (1389). The exact date or period of its creation is not known, but based on its themes, similar to other oral epics related to the persons and events prior to the Battle of Kosovo, it is most probably one of the oldest extant Serbian epics. It was finally written down in 1815 by Vuk Karadžić, who heard it from Tešan Podrugović, one of the most prolific Serbian oral poets and storytellers of the late 18th and early 19th century (see Perović and Vučković 2017). The recorded version originates from the region of Herzegovina and is, in fact, only one of several wedding ballads recorded by Vuk Karadžić as performed by Tešan Podrugović (Mitić 2017). These oral epics were traditionally (but not necessarily) sung with the accompaniment of the g u s l e (see Suvajdžić 2010), a single-stringed instrument. Tešan Podrugović was an exception to this rule, because he did not use the gusle in telling these stories but recited them with no musical background (consequently keeping the original medium monomodal). As is usually the case with this type of folk poetry, the epic also employs the strict decasyllabic structure that maintains the rhythm throughout the ballad, which

does not include any rhyming patterns. Thus, at least as far as the form is concerned, the analysed narrative has gone through two important adaptations, first by being transformed from its oral variant into the written text, changing the medium yet retaining the monomodal character[3], and then followed by the adaptation of the written form into the multimodal medium of comics, which also involved an introduction of another mode, the pictorial one. This second adaptation is the subject of our study.

The comic book was first drawn in black and white and published in 1938, but later redrawn in 1976 from the author's memory after the first version was lost during WWII. Colour was added for the 1989 special edition.[4] This final edition is the one examined in this paper. We opted for the coloured edition of the comic book simply because colour represents one of the specific semiotic resources used by comics authors in their works, and as such adds yet another layer of meaning to the examined text, as we will see in the next section. The comic book is 24 pages long and represents a more concise version of the original narrative, which is why it served as the starting point for comparing the two texts.

Our analysis, which is primarily of a qualitative and descriptive type, was focused on three major lines of inquiry. The first dealt with comparing the s t r u c t u r e of the original text and its adaptation by examining how their plotlines followed the hero wedding formula found in Serbian epic poetry (Petković 2019). This was done in an attempt to examine how true the comic remained to the oral ballad regarding this specific issue during the process of intersemiotic translation. In doing so, we also tried to assess the general level of fidelity (Kukkonen 2013: 80–85) of Lobačev's work to the source material. The second line of inquiry was further related to this, as it examined all the s t y l i s t i c and n a r r a t i v e d i f f e r e n c e s found in the two texts. In line with one of the challenges of adaptation listed by Lefèvre (2007: 3–4), that is, the problems that may occur during the deletion (omission) or addition of the material in the medial transposition of one media product to another, we paid special attention to the language and the narrative tone of the comic book compared to the ballad, and singled out the most important deviations in storytelling from the original text. Finally, the third point of interest covered various c o m i c s - r e l a t e d c o n c e r n s such as page layout, use of colour, representation of speed and motion, and absence of graphic devices to depict affect. Here we drew on the work by Cohn (2013) and Forceville (2011), as well as some of our previous studies (Stamenković and Tasić 2014; Tasić and Stamenković 2017, 2022), to analyse the use of upfixes and pictorial runes in representing speed, motion and emotion in the comic book. We also consulted Groensteen (2013) in examining whether the rhythm of the comic book expressed in its page layout emulated the strict meter of the oral epic. All of these tools were integrated into an elaborate methodological approach with the aim of scrutinising this specific case of adaptation (or intersemiotic translation) in as many aspects as possible.

We proceeded with the analysis in the following manner. Having read both the ballad and the comic book, we first identified certain points of inter-

est in the comic book and then referred back to the ballad for comparison. This strategy was more efficient in our view, since the ballad contained a number of secondary plot threads not found in the comic, which was not the case vice versa. Our methodological approach comprised the following steps (steps 1 through 3 are related to the first two major lines of inquiry, while steps 4 through 7 are related to the third one, i.e. comics-specific resources):

1. Track and match the key structural elements of the two plotlines, with a particular focus on the space allocated to the different parts of the hero wedding formula;
2. Search for and compare any existing stylistic differences (language, tone, etc.);
3. Pinpoint the main narrative differences between the two texts by closely looking at both;
4. Study the page layout of the comic, particularly with its rhythmic structure in mind;
5. Examine the use of colour and its ability to emphasize certain elements;
6. Identify and analyse any graphic devices used to express speed and motion;
7. Identify and analyse any graphic devices used to express emotion.

Lastly, due to the rather short length of the comic book and the mainly qualitative and descriptive nature of the present study, no quantitative aspects were taken into consideration. All of the major lines of inquiry mentioned above will be addressed in the following section.

4.   Ballad and comic book structure and plot outline

The first point of comparison between the ballad and the comic book has to be their structure. Namely, just like many other forms of folk literature (see Thompson 1955–1958), the Serbian epic ballads were developed in accordance with one of several available 'formulas'. The inventory of formulas of the Serbian epic ballads includes (i) protection of the weak and fight for justice, (ii) liberation, (iii) hero competitions, (iv) wedding, (v) family relations, (vi) social status, (vii) death of a hero, many of which have numerous variations (Petković 2019). *Dušan's Wedding* obviously belongs to the wedding formula and its variation "wedding with obstacles". Petković (2019: 84–85) observes that it is one of four ballads in which the main protagonist encounters obstacles both before and after taking the bride and one of only two ballads in which there are multiple obstacles before taking the bride. As such, the formula of *Dušan's Wedding* includes the following elements: the proposal with negotiations and conditions set by the in-laws (includes the disguised fraud attempt and the detection of the fraud), the gathering of the wedding guests and the journey towards the bride's home

(which initiates the wedding campaign), obstacles set by the in-laws before taking the bride, obstacles on the way back home (after taking the bride), and the arrival at the bridegroom's home. *Dušan's Wedding* also includes an uninvited hero as a saviour (which is an element borrowed from the liberation formula).

The comic book largely follows the formula employed by the ballad (and by the more general set of formulas from the oral epic tradition, as discussed above). It contains an opening page[5] which introduces six major characters: Tsar Dušan (the bridegroom), Miloš Vojinović (the main hero and protagonist), Roksanda (the bride), Balačko vojvoda (Miloš's most difficult obstacle), King Mihailo (Roksanda's father) and his mage counsellor (Mihailo and he are the main antagonists of the story). The remainder of the comic book renders the epic in a mostly consistent manner. We will track the key structural elements within the comic's panels. The proposal containing the negotiations process and conditions set by the in-laws opens the comic book: it includes all panels on the second page and the first three panels of the third page: here, Tsar Dušan sends out his envoy Todor to propose to Roksanda by negotiating the terms with Mihailo (called Mijailo in the original epic), King of Leđan (an imaginary city from South Slavic folklore). Roksanda's father sets the following condition: Tsar Dušan's nephews, the Vojinović brothers, are not to attend the wedding campaign and ceremony, as Mihailo claims they are trouble-makers (which is the disguised fraud attempt). The gathering of the wedding guests takes up a single panel (panel 5 on page 9[3], Fig. 1), in which we see them from the perspective of the Vojinović brothers.
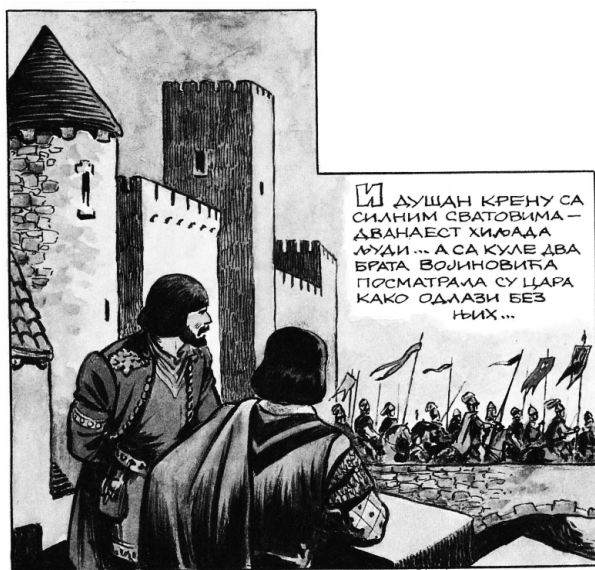
Fig. 1. Wedding guests gather and set out towards Leđan. Lobačev (1989: 9).

Fig. 2. Wedding campaign obstacles before taking the bride. Lobačev (1989: 15, 18, 22).

The two elder Vojinović brothers (Vukašin and Petrašin) detect this as a fraud set by the villain Mihailo and, aided by their mother, summon the youngest brother, Miloš Vojinović, who, in the structure of the ballad, acts as an uninvited hero, or a hero in disguise. He then joins the wedding guests in the campaign, i.e. during their journey towards Leđan, and throughout most of the comic he appears in shepherd's robes and is addressed as Bugarče ('young Bulgarian') or Čobanče ('young shepherd'). The comic book continues to follow the formula, so the next part involves three obstacles set by the in-laws: first a duel against a hero designated by King Mihailo, second jumping over three horses carrying three flaming swords, and third recognising the bride among several women (Fig. 2).

The portrayal of these obstacles takes up nearly 11 pages (pages 13[7]–23[17]), dominating the comic book and thus foregrounding the action elements. Miloš Vojinović (still in disguise) volunteers to tackle all three obstacles and is successful. Once the bride is allowed to join the wedding guests on their way back home, two further obstacles occur, again arranged by the malicious King Mihailo. They are again packed with action and take up another 6 pages (24[18]–29[23]). The wedding procession is attacked by Balačko vojvoda, an extremely strong enemy, who is also defeated. This is followed by the attack of 600 members of armoured cavalry – here Miloš Vojinović is aided by his friends and they fight off the assault (Fig. 3). Finally, Miloš Vojinović reveals his identity and the campaign ends when the procession brings the bride to Tsar Dušan's castle. This is how the formula reaches a happy ending. In the closing panel

we see Miloš Vojinović travelling back to the mountains to continue his usual lifestyle. The conclusion of the comic book all happens on one page only (page 30[24]).

Fig. 3. Wedding campaign obstacles after taking the bride. Lobačev (1989: 25, 28).

The overview of the comic book's structure allows us to see some of the differences in the allocation of space to the different parts of the epic formula. The most notable difference is that the action-packed parts describing the obstacles and the manner in which they are surmounted take up around 45% of the ballad (313 out of 690 verses), while they cover as much as 70% of the comic book and more (17 out of 24 pages). Therefore, we can once again confirm, that the notion of action (most prominent in the scenes depicting fights), is much more prevalent in the comic book. This, of course, takes its toll on the rest of the structure – we have shown that the gathering of the wedding guests, traditionally seen as a separate element in the ballad formula, is presented in only one frame, whereas the whole conclusion of the epic occupies only one page of the comic book.

5. Major stylistic and narrative differences between the ballad and the comic book

5.1 Changes in style

One of the most important and conspicuous changes when it comes to stylistic differences is the modernisation of language and its adjustment to the contemporary audience. The decasyllabic verse of the original epic is also missing from the comic book, with the text now in prose, again to better suit the audience for which the comic book is intended. This change is perhaps most notably illustrated by the fact that the eponymous Tsar Dušan, exclusively addressed as such in the comic book, is never mentioned by that name in the ballad, where he is only known as Stjepan, a moniker derived from his full name Stefan Uroš IV Dušan. Stjepan is the version of his first

name most likely used in Herzegovina, the birthplace of the poet Tešan Podrugović, who recited the epic to Vuk Karadžić. Even back in the 1930s, the name Stjepan would not resonate with the target audience, primarily comprising children and adolescents. The 1976 version and its coloured 1989 edition were even further removed from the original text, and the initial change was retained in them as well.

Another important adaptation to the target audience concerns the tone of the narration and the absence of explicit violence found throughout the epic ballad. Probably due to its graphic form being intended for younger readers (the coloured 1989 version was published by a renowned Yugoslav children's publishing company *Children's newspapers* [orig. *Dečje novine*]), Lobačev opted to leave out any scenes of explicit violence, which are explained in detail in the ballad. If he had tried to maintain an extremely high level of fidelity in his adaptation, the graphic nature of the ballad would have made the comic book unsuitable for Lobačev's target audience. Therefore, certain key moments in the plot that include violence in the ballad are represented in the comic book in such a way that actual violence takes place off-panel, while a number of other, arguably gratuitous violent scenes included in the ballad are completely absent from the comic. Even though the epic itself is taught to primary schoolchildren in its unabridged form, it is clear that a faithful visual representation of all the details of its storyline would simply be deemed too graphic for children that age.

5.2 Changes in narration

Three significant differences occur in the storytelling. The first is the introduction of a mage-like character who serves as counsellor to King Mihailo, the antagonist. This is accompanied by the omission of the Queen, whose place alongside the King is now occupied by the mage, who remains unnamed throughout the comic book, and whose appearance (as can be seen in Fig. 6, bottom right) is reminiscent of certain popular representations of the wizard Merlin, complete with long grey hair and beard, a dark blue robe and a pointed hat with crescent moons and stars of different shapes and sizes. Even though the ballad itself contains several fantastic elements, no such wizard can be found in the Serbian oral epic tradition, and his addition to the comic book is most probably inspired by the popularity of similar characters in comics and cartoons of the time, particularly the ones found in Hal Foster's *Prince Valiant*, which was launched only a year or so before Lobačev's comic book. Even though Lobačev did not mention any such influence in the introduction to the 1989 edition (he speaks only of being influenced by American detective comics and Disney), we believe that this connection between the two comics can further be corroborated by the fact that the story's protagonist, Miloš Vojinović, has an almost identical haircut as Prince Valiant (Fig. 4).

Fig. 4. Prince Valiant (left) and Miloš Vojinović (right). Foster (2013: 52); Lobačev (1989: 10).

The other important difference lies in the role that the robes worn by Miloš Vojinović play in the source material and the lack of such relevance in the comic book. In the ballad, Miloš's shepherd clothing and his wedding guest attire underneath serve a more prominent purpose than a mere disguise. They represent a major narrative tool that introduces the duality of the character who is both a shepherd and a nobleman, which, in fact, allows him to overcome all the obstacles put before him. The point in the story at which the hero removes his shepherd clothing to reveal the elegant garb beneath also differs, taking place in the ballad in the narratively charged moment of the recognition of the bride, as opposed to the very last page of the comic book, when Miloš reveals his true identity to Tsar Dušan. Đorđe Lobačev must have again changed this detail with an eye for his intended audience, who would perhaps need a more detailed explanation of the metaphorical and symbolic meaning of the different types of robes worn by the protagonist, which would in turn require more space than the rather short length of this comic.

Finally, the fantastic three-headed character of Balačko vojvoda, the strongest of Miloš's enemies, is slightly altered. In the epic, he is a more articulate character, who is familiar with the hero and acts accordingly, whereas he does not seem to recognise the protagonist at all in the comic book. Before he appears in the narrative in person, he is portrayed in the ballad as having three heads, spewing cerulean flames from one and icy winds from the other, with no mention of the purpose of the third head. In the comic book, probably with the aim of preparing for a more graphically

powerful impression on readers, he is first described to the King by the mage as having a specific 'power' (blizzard, fire and wind) in each one of his three heads, which are then colour-coded for added effect on the next page when Balačko appears for the first time (see Fig. 6, left). This change affects the narrative less than the other two, though it may still be attributed to the process of adaptation.

6.   Semiotic resources not available in the original text

In this part of our analysis, we will examine those aspects of the comics medium that cannot be found in the original text and that are closely related to the multimodal nature of comics. We will look in detail how page layout is employed by Đorđe Lobačev to drive the narrative forward and how the choice of panel arrangements differs in rhythm from the strict decasyllabic structure of the source material. Even though a substantial majority of pages follow the dominant page layout, we were not able to detect any rhythmic regularities that could perhaps indicate a significant level of fidelity to the structure of the epic ballad. Another specific affordance of the medium is the use of colour which conveys the creator's specific artistic intentions. Besides all this, the final point of interest in this part of analysis lies in the manner in which the author applies certain graphic devices describing motion, speed and spatiality, such as speed and motion lines, but forgoes the use of others that would emphasize affect, such as upfixes (Cohn 2013) or pictorial runes (Forceville 2011; Tasić and Stamenković 2017). This decision can be seen as mirroring a basic trait of oral epic poetry, which tends to eschew overt displays of emotion.

6.1 Page layout

As already mentioned above, the page layout of the comic book is uniform for the most part, excluding three pages that contain, in their order of appearance in the comic, a sort of visual *dramatis personae* at the very beginning (page 7[1]), a single-panel page showing Miloš Vojinović overcoming one of the obstacles (page 18[12], middle panel in Fig. 2), and a two-panel page depicting the final battle between Miloš's and Balačko's cavalries (page 28[22]). The remaining 21 pages follow a more or less stable pattern (e.g. page 10[4], Fig. 5), albeit with a varying number of panels (from five to eight), which leads us to believe that there is no clearly intended rhythm that would be evocative of the decasyllabic structure of the source material. In Groensteen's (2013: 135–138) terms, the beat of the multiframe in the comic does not match the strict rhythm of the poem, and Lobačev's constant slight changes in the number and positioning of panels seem to renounce the fixed metric form of regular layout that would be taken as perhaps truer to the original.

Fig. 5. An example of the typical page layout. Lobačev (1989: 10).

As can be seen from the above Figure, pages often contain entirely textual panels (not counted here as panels proper) that further the narrative and serve as a bridge between two panels either divided by a longer period of time or happening at two separate locations. In addition to the varying number of panels per page, the panels, though angular, do not follow a strict grid (Fig. 5). This contributes to the action-packed narrative by providing a sense of constant haste and motion. Everything seems even more lively and almost dynamic in an inherently static medium. Nevertheless, all of this distances the comic book even further from the epic ballad in terms of the immanent pace of storytelling, where changes in both rhythm and tempo, reinforced by the distinct page layout, appear much more prominent in the comic book, adding a new layer to the text.

6.2 Use of colour

The special 1989 coloured edition of the comic book introduces another dimension to the meaning-making process of the adaptation. The comic book was originally drawn in black and white and published in the Serbian newspaper *Politika* in 1938, while colour was added some 50 years after that. Bearing in mind that *Tsar Dušan's Wedding* was not originally conceived as a colour comic, our focus here will primarily be on how Lobačev's artwork employs colours as means of emphasis of specific physical manifestations, on the one hand, and of certain personality traits, on the other.
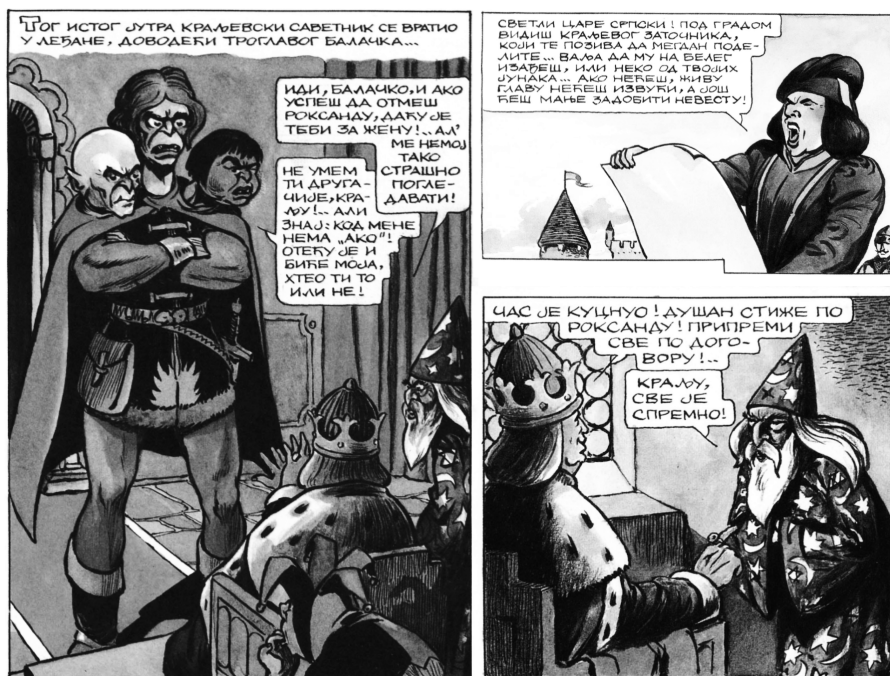
Fig. 6. Use of colour for emphasis. Lobačev (1989: 12, 14, 24).[6]

The left panel in Fig. 6 (page 24[18], panel 1) represents the first appearance of the character of Balačko vojvoda, the three-headed nemesis of the ballad's hero, Miloš Vojinović. We have already described that each of the heads possesses a particular 'power', and Lobačev here uses colour to accentuate the exact nature of those powers. The leftmost head, from the reader's perspective, is coloured icy blue to indicate that it spews blizzard, as will be seen later on in the comic (see Fig. 3, left). The middle head spews cerulean flame, and the rightmost is supposed to blow strong winds. The choice of colours for these latter two heads might be a bit confusing at first, since the reddest one would probably be more readily associated with fire, but the fact that the flames coming out of the middle head are coloured in

sky blue and orange, fitting the description as cerulean in the source text, justifies the use of orange for the fire head, while the red colour of the wind head might simply imply the consequences of the extra effort of producing gusts of wind. What is also important is that none of the colours used for the heads match the colour of Balačko's arms, which only further supports the interpretation of the application of colours as calculated. The top right panel in Fig. 6 (page 14[8], panel 1) may offer another explanation for the use of red in the wind head, as we can see a town crier making a public pronouncement from the top of his lungs, his face red from all the strain of shouting and emitting a great amount of air in the process. Again, colour is used to emphasise a physical manifestation of sorts, and to facilitate the reader's understanding of the image. The realism of the scene is expressed not only through the wide open mouth and closed eyes, unchanged from the original black-and-white version, but also through Lobačev's colouring technique. Finally, the bottom right panel in Fig. 6 (page 12[6], panel 5), illustrates another type of emphasis, which is more metaphorical in character. Here, certain personality traits are hinted at by colours conventionally used to indicate malevolence or ill intent. The mage can be seen both in the bottom right panel in Fig. 6, where his countenance is predominantly grey, and in the left panel, where it takes on a greenish hue. Both of these colours can be associated with envy and moral decay, and the juxtaposition of the mage's visage against all other characters' faces, barring Balačko vojvoda, of course, clearly underscores his wicked intentions as the character that comes up with all the different obstacles put before the hero in the comics adaptation, though completely absent in the original.

6.3 Speed and motion

One of the most challenging tasks for any comics artist is to try and succeed in depicting motion in what is intrinsically a static medium. If the work at hand is additionally intended to be part of the action genre, where speed and movement are some of the basic characteristics, this task becomes even more important for its eventual felicity. The artist here draws on two techniques to represent motion and speed of movement.

First (Fig. 7, left), he uses speed and motion lines to indicate action in panel 3 on page 16[10]. Such lines (Forceville 2011; Tasić and Stamenković 2017) are graphic devices that come in different shapes, sizes, and positions, yet all with the intention of implying motion paths (trajectory lines) and the velocity at which a movement occurs. They are usually drawn as extensions of the person or the object that is supposed to be moving in the depicted scene, or adjacent to them, as is the case in motion lines marking the position previously occupied. The example from Fig. 6 illustrates this by employing two lines as an extension of Miloš Vojinović's arm as he swings his mace in an attempt to hit the runaway knight. The two lines show the motion path of the hero's hand, but also accentuate the speed at which this

movement is happening. Furthermore, the curved lines that contour the clouds of dust rising behind the galloping adversaries are not merely used to delineate dust but can at once be read as motion lines that show the direction in which the dust is moving. Another technique is exemplified in the right panel in Fig. 7 (page 13[7], panel 5), where Lobačev's realistic style comes to the fore. Namely, the banners hanging from the heralds' trumpets appear to be suspended in the midst of fluttering, while the varying extent to which they are displaced from their resting position further evokes the possible real-life unfolding of such a scene. These graphic devices are present throughout the comic book and their subtle use by the artist enhances the experience of the action-oriented narrative.

Fig. 7. Speed and motion. Lobačev (1989: 13, 16).

6.4 Graphic devices depicting affect

There are different ways in which comics artists can attempt to vividly depict affect in their work. Apart from the obvious choice, which would be to draw characters as realistically as possible, artists have a number of comics-specific tools at their disposal, including upfixes (Cohn 2013) and pictorial runes (Forceville 2011; Tasić and Stamenković 2017). These may include squiggly or straight lines, spirals, twirls, droplets of liquid, etc., drawn either in a halo-like fashion around a character's head or as singular graphic decorations usually appearing above a character and indicating various emotions such as anxiety, anger, surprise, fear, etc. Unlike the previously described speed and motion lines, which are all-present, the absence of these affective graphic devices in this comic book is striking. Most probably due to his realistic style the artist relies fully on facial expressions and expressive anatomy in trying to convey emotive meaning to his audience. Even so, such explicit depictions of emotions are rare, and the occasional smile or frown are nearly the only communications of the characters' emotional states.

Apart from the fit to the artist's style, we believe that the subdued expression of emotional content can also be seen as fitting the basic traits of Serbian oral epic poetry, which focuses more on events than on emotions. In the process of transmutation or intermedial transposition from the epic ballad to the comic book, this appears to be one of the aspects in which a very high level of fidelity was maintained.

7.   Conclusions

The present study attempted to trace the process of adapting an oral epic ballad into a comic book. As we have seen, the original oral epic was first transferred into the written form, which later served as the source material for the comic book created by Đorđe Lobačev, one of the founding fathers of the Serbian school of comics. We have tried to examine this process by drawing on a number of theoretical concepts to link the original text and its comics adaptation. The comic is a product of intersemiotic translation or transmutation, as it uses its own pictorial sign system to interpret the verbal signs of the epic ballad. It can also be seen as an instance of medial transposition, since the process of adaptation actually transforms one media product into another, establishing an intermedial connection between the two texts.

More closely, we have observed the structural similarities and differences between the ballad and the comic, concluding that the latter follows the hero wedding formula from the former, albeit with a greater emphasis on the action aspect of the plot. A similar comparative analysis of other stylistic and narrative elements has revealed some major differences, which primarily stem from the artist's desire to adjust the narrative to the taste of the comic book's intended audience. These variations include, above all, the modernisation of the language and the less violent representation of certain events from the ballad, in an attempt to bring the original text to modern, mainly younger, readers. The observed changes in some of the characters in the comic were also probably made for the same reasons. As for the specific semiotic resources available only in the comics adaptation and not in the ballad, we have seen how the artist uses specific page layouts and colouring techniques to emphasise particular storytelling aspects and create a more dramatic impact on his readers. The absence of some other graphic devices, such as upfixes or pictorial runes, which are often used in comics to depict affect, is equally striking. We suggest that this is in line with the source material, which is known for its rather sparse depiction of emotions.

Directions for further research could be provided by Đorđe Lobačev's diverse body of work, either by choosing similar source material and comparing the process of adaptation to the present analysis, or by looking into other types and genres of original texts adapted by Lobačev into comic

books. Furthermore, there are other Serbian comics artists (e.g. Petar Meseldžija, Mijat Mijatović, Nikola Mitrović – Kokan) who have done similar adaptations and whose procedures can be compared to Lobačev's. Lastly, using a more formal discourse analysis approach (e.g. based on Kamp and Reyle 1993; Asher and Lascarides 2003) to investigate comparable stretches of discourse from the oral epic ballad and the comic book could reveal additional facets of similarities and differences between the two portrayals of same events.

Notes

1   Đorđe Lobačev (Yuriy Lobachev; 1909–2002) was a famed Soviet Russian and Serbian-Yugoslav author, known as "the father of Serbian comics". He began publishing comics in 1935, and was one of the first comics artists in pre-WWII Yugoslavia. His series *Bloody Inheritance* (orig. *Krvavo nasledstvo*) was, in fact, the second ever published Serbian comic. More importantly, he was the first author who consciously tried to steer away from foreign influences in Serbian comics at the time, and the first to choose a local theme as a backdrop for his work, which became his most invaluable contribution to the school of Serbian comics (Grujičić 2002).

2   The epic was first translated into English by Geoffrey N. W. Locke as *Tsar Dushan's Wedding* in the anthology *The Serbian epic ballads* (1997). That is why we have decided to retain the original title of *tsar*, which designates East and South Slavic monarchs, instead of using its Western European counterpart *emperor*.

3   With regard to our claim that the oral epic is monomodal and does not contain any additional separate modes, such as rhythmic structure, for example, we consider rhythm and other similar aspects here to be part of the verbal mode, the same way colour is part of the pictorial mode in our analysis.

4   Apart from the fact that the author redrew his original work at the request of Žika Bogdanović, one of the most important Serbian theoreticians of comics and visual narration in general, and the then director of the "Yugoslavia" publishing company (Lobačev 1989: 4), to the best of our knowledge, there are no mentions of any specific differences between the two versions of the comic book. The colour was later added to make the comic more appealing to the new generation of readers.

5   The page numbering used in the paper follows the numbering from the source material, which is a collection of Lobačev's works published under the title *Tracking the folk imagination* (orig. *Tragom narodne mašte*) that contains the examined comic. Where pertinent to the analysis, the actual pages of the comic book itself are given in square brackets.

6   Due to printing constraints, the original coloured images are only available in black
       and white.

Bibliography

Alghadeer, Hessa A. (2014). Digital landscapes: Rethinking poetry interpretation in mul-
       timodal texts. *Journal of Arts and Humanities* 3, 2, 87–96. URL: https://www.thearts-
       journal.org/index.php/site/article/view/354 [retrieved February 13, 2023].

Asher, Nicholas and Alex Lascarides (2003). *Logics of Conversation*. Cambridge: Cam-
       bridge University Press.

Bateman, John A. and Janina Wildfeuer (2014). A multimodal theory of visual narrative.
       Journal of Pragmatics 74, 180–208.

Bramlett, Frank, Roy T. Cook and Aaron Meskin (eds.) (2017). *The Routledge Compan-
       ion to Comics*. New York: Routledge.

Cohn, Neil (2013). *The Visual Language of Comics: Introduction to the Structure and
       Cognition of Sequential Images*. London: Bloomsbury.

Dusi, Nicola (2015). Intersemiotic translation: Theories, problems, analysis. *Semiotica*
       206, 181–205.

Eco, Umberto (2003). *Mouse or Rat? Translation as Negotiation*. London: Orion Pub-
       lishing Group.

Forceville, Charles (2011). Pictorial runes in *Tintin and the Picaros. Journal of Pragmat-
       ics* 43, 3, 875–890.

Gaudreault, André and Philippe Marion (2004). Transécriture and narrative mediatics:
       The stakes of intermediality. English translation by Robert Stam. In: Robert Stam
       and Alessandra Raengo (eds.). *A Companion to Literature and Film*. Malden: Black-
       well, 58–70.

Groensteen, Thierry (2013). *Comics and Narration*. English translation by Ann Miller.
       Jackson, MS: University Press of Mississippi.

Grujičić, Nebojša (2002). Avantura je šampanjac života [Adventure is the champagne
       of life]. *Vreme*. 1 August 2002. URL: https://www.vreme.com/kultura/avantura-je-sam-
       panjac-zivota/ [retrieved November 27, 2022].

Hutcheon, Linda (2013). *A Theory of Adaptation*. 2nd ed. Oxford: Routledge.

Jakobson, Roman (1959). On linguistic aspects of translation. In: Reuben Brower (ed.).
       On Translation. Cambridge, MA: Harvard University Press, 232–239.

Jensen, Klaus B. (2016). Intermediality. In: Klaus B. Jensen and Robert T. Craig (eds.).
       The International Encyclopedia of Communication Theory and Philosophy. Vol. 2.
       Hoboken, NJ: Wiley-Blackwell, 972–982.

Kamp, Hans and Uwe Reyle (1993). *From Discourse to Logic: Introduction to Modelthe-
       oretic Semantics of Natural Language, Formal Logic and Discourse Representation
       Theory*. Amsterdam: Kluwer Academic Publishers.

Kourdis, Evangelos and Charikleia Yoka (2014). Intericonicity as intersemiotic transla-
       tion in a globalized culture. In: Yongxiang Wang and Haihong Ji (eds.). *Our World: A
       Kaleidoscopic Semiotic Network, Proceedings of the 11th World Congress of the
       IASS/AIS, 5–9 October 2012*. Nanjing: Hohai University Press, 162–176.

Kress, Gunther and Theo van Leeuwen (2001). *Multimodal Discourse: The Modes and Media of Contemporary Communication*. London: Arnold.

Kukkonen, Karin (2011). Comics as a test case for transmedial narratology. *SubStance* 40, 1, 34–52.

Kukkonen, Karin (2013). *Studying Comics and Graphic Novels*. Chichester: Wiley-Blackwell.

Lefèvre, Pascal (2007). Incompatible visual ontologies? The problematic adaptation of drawn images. In: Ian Gordon, Mark Jancovich and Matthew P. McAllister (eds.). *Film and Comic Books*. Jackson, MS: University Press of Mississippi, 1–12.

Mitić, Kristina (2017). Ženidba Ajduka Trskoloma [Haiduk Trskolom's wedding]. In Drago Perović and Radomir Vučković (eds.). *200 godina susreta Tešana i Vuka* [*200th Anniversary of Tešan and Vuk's Meeting*]. Gacko: SPKD Prosvjeta, 87–106.

O'Halloran, Kay L., Sabine Tan and Peter Wignell (2016). Intersemiotic translation as resemiotisation: A multimodal perspective. *Signata* 7, 199–229.

Packard, Stephan, Andreas Rauscher, Véronique Sina, Jan-Noël Thon, Lukas R.A. Wilde and Janina Wildfeuer (2019). *Comicanalyse: Eine Einführung*. Berlin: J. B. Metzler.

Pârlog, Aba-Carina (2019). *Intersemiotic Translation: Literary and Linguistic Multimodality*. Cham: Palgrave Macmillan.

Pereira, Nilce M. (2008). Book illustration as (intersemiotic) translation: Pictures translating words. *Meta* 53, 1, 104–119.

Perović, Drago and Radomir Vučković (eds.) (2017). *200 godina susreta Tešana i Vuka* [*200th Anniversary of Tešan and Vuk's Meeting*]. Gacko: SPKD Prosvjeta.

Perteghella, Manuela (2019). The case of the poem in motion: Translation, movement and the poetic landscape. In: Madeleine Campbell and Ricarda Vidal (eds.). *Translating Across Sensory and Linguistic Borders: Intersemiotic Journeys Between Media*. Cham: Palgrave Macmillan, 63–85.

Petković, Danijela (2019). *Junak i siže epske pesme* [*The Hero and the Structure of Epic Ballads*]. Beograd: Institut za književnost i umetnost.

Rajewsky, Irina O. (2005). Intermediality, intertextuality and remediation: A literary perspective on intermediality. *Intermédialités* 6, 43–64.

Rippl, Gabriele and Lukas Etter (2013). Intermediality, transmediality, and graphic narrative. In: Daniel Stein and Jan-Noël Thon (eds.). *From Comic Strips to Graphic Novels: Contributions to the Theory and History of Graphic Narrative*. Berlin: De Gruyter, 191–218.

Schröter, Jens (2011). Discourses and models of intermediality. *CLCWeb: Comparative Literature and Culture* 13, 3.

The Serbian Epic Ballads: An Anthology (English translation by Geoffrey N. W. Locke). (1997). Belgrade: Nolit.

Smith, Matthew J. and Randy Duncan (eds.) (2017). *The Secret Origins of Comics Studies*. New York: Routledge.

Sonzogni, Marco (2011). *Re-covered Rose: A Case Study in Book Cover Design as Intersemiotic Translation*. Amsterdam: John Benjamins.

Stamenković, Dušan and Miloš Tasić (2014). The contribution of cognitive linguistics to comics studies. *Balkan Journal of Philosophy* 6, 155–162.

Stein, Daniel (2015). Comics and graphic novels. In: Gabriele Rippl (ed.). *Handbook of Intermediality: Literature – Image – Sound – Music*. Berlin: De Gruyter, 420–438.

Suvajdžić, Boško (2010). *Pevač i tradicija* [*The Oral Poet and the Tradition*]. Beograd: Zavod za udžbenike.

Tasić, Miloš and Dušan Stamenković (2017). Exploring pictorial runes in Luca Enoch's comic book series "Gea". *Facta Universitatis: Series Linguistics and Literature* 15, 2, 123–141.

Tasić, Miloš and Dušan Stamenković (2022). From statics to dynamics: Intersemiotic conversion of metaphor and its consequences. In: Shyam Wuppuluri and A. C. Grayling (eds.). *Metaphors and Analogies in Sciences and Humanities: Words and Worlds*. Cham: Springer, 523–551.

Taylor, Christopher (2020). Multimodality and intersemiotic translation. In: Łukasz Bogucki and Mikołaj Deckert (eds.). *The Palgrave Handbook of Audiovisual Translation and Media Accessibility*. Cham: Palgrave Macmillan, 83–99.

Thompson, Stith (1955–1958). *Motif-Index of Folk-Literature: A Classification of Narrative Elements in Folktales, Ballads, Myths, Fables, Mediaeval Romances, Exempla, Fabliaux, Jest-Books, and Local Legends*. Vol. I–VI. Bloomington, IN: Indiana University Press.

Tsui, Cynthia S. (2012). The authenticity in 'adaptation': A theoretical perspective from translation studies. In: Laurence Raw (ed.). *Translation, Adaptation and Transformation*. London: Bloomsbury, 54–60.

Sources

Foster, Harold R. (wri/art) (2013). *Princ Valijant*. Tom I, 1937–1938. [Prince Valiant. Vol. I, 1937–1938]. Belgrade: Čarobna knjiga.

Lobačev, Đorđe (wri/art) (1989). *Tragom narodne mašte*. Tom I. [Tracking the folk imagination. Vol. I]. Gornji Milanovac: Dečje novine.

Ženidba Dušanova [Dušan's Wedding] (1815/1975). In: Vuk St. Karadžić (ed.). *Srpske narodne pjesme II* [Serbian Folk Ballads II]. Beograd: Nolit, 85–99.

Image Sources

Fig. 1. Lobačev (1989: 9).
Fig. 2. Lobačev (1989: 15, 18, 22).
Fig. 3. Lobačev (1989: 25, 28).
Fig. 4. Foster (2013: 52); Lobačev (1989: 10).
Fig. 5. Lobačev (1989: 10).
Fig. 6. Lobačev (1989: 12, 14, 24).
Fig. 7. Lobačev (1989: 13, 16).

Assoc.-Prof. Dr. Miloš Tasić
University of Niš
Faculty of Mechanical Engineering
Kabinet 227
Aleksandra Medvedeva 14
18000 Niš
Serbia
E-Mail: milos.tasic@masfak.ni.ac.rs

Assoc.-Prof. Dr. Dušan Stamenković
Södertörn University
School of Culture and Education
Primus PA 217
Alfred Nobels allé 7
141 52 Huddinge
Sweden
E-Mail: dusan.stamenkovic@sh.se

Unreliable Iconicity, or: Accounting for the Cartoonish Pictures of Comics in Multimodal Reasoning

Lukas R.A. Wilde, Norwegian University of Science and Technology (NTNU)

Abstract. In this article, I explore how pictoriality in comics cannot be conceived as a single semiotic mode, but rather as two distinct thresholds of pictorial comprehension and interpretation: a mostly pre-attentional cognitive reconfiguration of two-dimensional lines on paper into three-dimensional bodies in space, as well as a more conscious interpretational mapping (or disentanglement) of perceivable features with storyworld entities, guided by often conflicting multimodal forces of specific textual cues, generic traditions, and paratextual markers. I analyze two comics by German artist Sascha Hommer that are typical for a medium-specific unreliability of iconicity in which we can never be sure how the inhabitants of both Hommer's fantastic as well as of his autobiographical storyworld may be perceived by other characters. These questions, however, remain crucial for evaluating the thematic point of both works, especially as readers have to revise their earlier assumptions throughout their multimodal reasoning. My analysis of Hommer's works will indicate how the two thresholds described prove indispensable for any account of the cartoonish pictures of comics and their media-specific unreliability.

Keywords. Cartoon, cartoonization, comics, iconicity, pictoriality, semiotics, transmedia narratology, unreliability, Sascha Hommer

Zusammenfassung. Der Beitrag argumentiert, dass Bildlichkeit im Comic nicht als eine distinkte Zeichenmodalität, sondern als zwei unterschiedliche Schwellen des piktorialen Verstehens und Interpretierens adressiert werden muss: eine zumeist prä-attentionale kognitive Rekonfiguration von zweidimensionalen Linien auf Papier zu dreidimensionalen Körpern im Raum sowie eine bewusste(re-)interpretatorische Zuordnung (oder Entflechtung) von wahrnehmbaren Merkmalen zu Storyworld-Entitäten. Diese Zuordnung wird entlang oft widersprüchlicher multimodaler Signale geleitet, die sowohl spezifische textuelle Hinweise als auch generische Traditionen sowie paratextuelle Markierungen umfassen. Der Beitrag analysiert zwei Comics des deutschen Künstlers Sascha Hommer, die typisch für eine solch medienspezifische ikonische Unzuverlässigkeit sind, da wir nie sicher sein können, wie die Bewohner:innen sowohl von Hommers phantastischen als auch seiner autobiografischen Welten von anderen Figuren wahrgenommen

werden können. Diese Fragen bleiben dennoch entscheidend für die Erschließung der thematischen Pointe beider Werke, zumal die Leser:innen im Laufe der multimodalen Lektüre frühere Annahmen revidieren müssen. Die Analyse von Hommers Werken zeigt so, wie sich die beiden beschriebenen Schwellen als unverzichtbar für die angemessene Einschätzung cartoonisierter Bilder im Comic sowie ihrer medienspezifischen Unzuverlässigkeiten erweisen.

Schlüsselwörter. Cartoon, Cartoonisierung, Comics, Ikonizität, Piktorialität, Semiotik, Transmediale Narratologie, Unzuverlässigkeit, Sascha Hommer

Even though multimodality is widely accepted as a suitable or even necessary theoretical framework to analyze and theorize comics, the respective accounts of p i c t o r i a l i t y as a d i s t i n c t m o d e differ widely (cf. Machin 2014). Put in a broader context, this is hardly surprising, since "[t]he use made of *iconicity* in multimodal studies varies considerably", as John Bateman has pointed out (2018: 18). For comic studies, this becomes especially conspicuous if we look at cartoonish representations of fictional and non-fictional characters, their bodies and faces. In this article I would like to show that there are three different ways of dealing with the 'mode' of pictoriality within comprehension and interpretation – sometimes on a global textual level of a storyworld as a whole, sometimes on a specific, local level of individual characters in contrast to backgrounds:

1. The abstract, simplified cartoon lines have to be e n r i c h e d towards a more 'natural' perceivability (adding something to the material means of representation).
2. The cartoon lines have to be r e p l a c e d by a different perceivability within the imagination by ignoring some of their perceivable features, perhaps because they are taken as metaphorical or allegorical (subtracting something from the material means of representation).
3. The cartoon lines should be interpreted as 'l i t e r a l' as possible within the mimetic (diegetic) domain: characters and objects would then look just as they are represented, merely within three dimensions instead of two (aligning the imagination as closely as possible to the material means of representation).

Since all three options are, in theory, always available, cartoonish pictures of comics – and a cartoonish mode of pictoriality more generally – are not inherently vague and underdetermined because of option 1), but because it remains often intentionally unreliable whether they must, should, or could be interpreted along options 1), 2), or 3). While this article is thus not primarily concerned with comics' multimodality, it will focus on the inherent tensions to the 'mode' of pictoriality within multimodal meaning-making (or storyworld-construction). With recourse to Charles Sanders Peirce's dis-

tinction between iconic and hypoiconic reasoning, pictoriality entails two distinct thresholds of pictorial comprehension and interpretation: a mostly pre-attentional cognitive reconfiguration of two-dimensional lines on paper into three-dimensional bodies in space on the one hand, and a more conscious interpretational mapping of perceivable features to storyworld entities on the other. The latter will be guided by conflicting multimodal information provided by specific textual cues, paratextual markers, and generic traditions.

Fig. 1. McCloud (1993: 36).

Discussions of cartoonish pictures in comics are initially derived from artist Scott McCloud, who reinterpreted the term "cartoon" for a specific pictorial style that he described as "amplification through simplification" (1993: 30). Andrei Molotiu recently addressed cartooning as a key term for comic studies:

> [A] graphic simplification of figurative shapes for purposes of communication, humor, and so on in comic strip and comic book rendering (as well as, of course, in gag cartoons, animation, and other fields of visual media) (Molotiu 2020: 153).

Chris Gavaler (2022: 6–7) even discussed this style as a possible criterion for defining comics in general. McCloud anyways locates comic book drawings on a scale ranging from photography to a completely simplified smiley face lacking any individual features. The idea is then that there is a ref-

erential reality 'behind' any comic book picture that, basically, 'looks like our's' – which is then abstracted, distorted, or stylized by the artist (cf. McCloud 1993: 36; Fig. 1).

Fig. 2. Delisle (2008: 34).

This assumption certainly seems plausible if we look at autobiographical works such as Riad Sattouf's *The Arab of the Future* (2014–2018, see Sattouf 2015) or Guy Delisle's travel comics (such as *Pyongyang*, 2005, or *Burma Chronicles*, 2008). Sattouf recounts the story of his upbringing in Middle Eastern countries (especially Syria), while Delisle documents journeys into places such as Myanmar, Israel, or North Korea. Both artists have been praised for their subjective, yet faithful and sincere representations of places and cultures foreign to most international readers. That these works – and many others – are accepted as authentic may be surprising to someone not acquainted with the media form, as all the characters are strongly cartoonized. Heavily relying on caricature, bodies and faces are reduced to mere outlines, bulbous noses, and pop-eyes. There is little general doubt about these comics' faithfulness to their artists' actual experiences, however, hence readers can be sure that the cartoonization is entirely on the side of the representation, not the represented characters and worlds themselves. When Delisle's self-representation wonders "Aw Geez! If I looked a bit more Burmese, they would've let me through" (Delisle 2008: 34; Fig. 2), he most certainly does not mean the black outline contours without colors or internal features that we can see on the page. These are instead intended to represent 'regular', three-dimensional human beings within the storyworld and to Delisle's avatar himself, just as their personalities are complex and full of contradictions. Or, as Gavaler puts it:

> When Alison Bechdel draws her and her family members' mouths as single dots in her 2006 *Fun Home*, viewers likely do not imagine that the actual individuals' mouths are so impossibly proportioned (Gavaler 2022: 47).

On the level of the intersubjective communicative construct (cf. Thon 2016: 54–56), such characters are clearly p e o p l e made of flesh and blood. Their specific visible appearances, however – to other characters within their diegetic environment – remain largely undefined. At best, we can make r e l a t i o n a l claims, such as has-a-larger-nose-than or is-bigger-than. We usually have no other, 'unmediated' access to the corresponding world(s). Depictions of cartoon protagonists are thus always inherently vague, leaving ample room for the individual imagination. As Roy T. Cook (2015) puts it: "[T]he physical appearance of drawn characters in general is indirect, partial, inferential, and imperfect" (2015: 25).

1.   The ultimate conundrum of comic studies

The assumption that comic book characters and storyworlds are always abstractions from the regular "visual ontology" (cf. Lefèvre 2007) of our world, however, becomes more difficult to uphold for entirely fictional or fantastic works which make no claims to any sort of p e r c e p t i o n a l   r e a l i s m . Indeed, one of the most prominent features of comics' mediality might be the fact that their pictures do n o t need to be taken as abstractions, but that they open up m e d i a - s p e c i f i c spaces for the imagination (cf. Wolk 2007: 141). Gavaler revealed a general lack of attention in comic narratology, and especially in picture theories of comics, towards problems of fiction:

> To identify physically plausible exaggerations, viewers need to reference the subject's actual face, which is impossible if the subject is fictional (2022: 48).

All questions of abstraction, stylization, or underdetermination can only be discussed in relation to a given "baseline reality" (2022: 57) which, in comics, may easily deviate from non-fictional worlds:

> Though cartoon objects are impossible in our reality, their transparently drawn qualities could accurately depict a cartoon reality (Gavaler 2022: 46).

What is the 'actual' perceptibility of Donald Duck as a fictional entity anyways? Does he look like his representations in full body suits in Disneyland theme parks, or do those rather try to represent him as faithfully as possible with three-dimensional materiality? It is, in fact, possible to consider all representations of Donald as 'metaphorical' or 'non-literal'– just like Art Spiegelman's drawing of his father Vlad as a cartoon mouse in *MAUS* (1980–1991). We cannot be certain whether the members of the Duck family are (special, human-like) ducks – or whether they are regular humans, just rep-

resented as ducks, as Disney artist Don Rosa upholds emphatically (cf. Rosa 2014: 8). As Thomas Lamarre has put it for all anthropomorphic animals of comics, manga, and cartoons, "[a]re these humanized animals or animalized humans?" (Lamarre 2008: 82; see also the contributions in Herman 2017). The problem of unreliable iconicity in comics, however, does not stop at such problems of "bestial ambivalence" (Wells 2009: 72), that is, the "strategic blurring of boundary between animal and human" (Alaniz 2020: 329). Comics scholar Martin Schüwer (2008: 23, 510) addressed what he considered the u l t i m a t e   c o n u n d r u m of comic studies – for human protagonists as well: Should one attribute the caricature-style of Charles M. Schultz' Peanuts solely to their representations and imagine that Charlie Brown and Snoopy 'actually' look quite differently in the context of the narrated world? The assertion that they are 'only drawn that way', but 'actually' look like photographs of 'real' people, does not seem to do justice to the drawing styles and their media forms. There is a systematic alternative to this assumption, of course: Fantastic worlds of comics, manga, or animation can not only break with physical laws (characters possessing superpowers or magic), but could also be taken to exhibit a special "visual ontology" that l o o k s entirely differently from ours. To Gavaler, all stylistic elements in comics could then possess a peculiarly "semi-representational" (Gavaler 2022: 48) quality that can oscillate freely between discourse and diegesis and may be attributed only case-by-case to one side or the other. A detachment of cartoon worlds from all demands of everyday reality, not only but especially with regard to perception, is particularly prominent in the Japanese manga and anime discourse as anthropologist Shunsuke Nozawa summarized:

> Character design strives to give characters a sui generis reality, one that is irreducible to our kind of reality (Nozawa 2013: n.pag.; cf. also Berndt 2013)

We must always 'correct' s o m e t h i n g to what we see with recourse to world knowledge, however. Black and white pictures, for instance, will usually be interpreted as a colorful world. We can deduce this from the fact that many collected manga volumes, for instance, include a few color pages in their openings, only to 'switch' to monochrome representations later. Usually, cover illustrations are also in color, as are many fan art interpretations.

Taken together, we arrive at our three interpretational options of adding, subtracting, and aligning, as distinguished earlier. Transmedia narratology (cf. Thon 2016: 39–46) has provided powerful vocabulary to describe these different options mostly alongside option 1) (adding), a "principle of minimal departure" (adding something to the representation that is only "implied", cf. Ryan 1991: 48–54) and option 2) (subtracting) a "principle of charity" (ignoring aspects of the representation that run contrary to the intersubjective construction of the storyworld, cf. Walton 1993: 174–187; Gavaler 2022: 89–91). Option 3) remains largely unexplored within narratological accounts (cf., however, Wilde forthcoming), although there are strong claims towards it from phenomenological comic theories:

> The fact that small resin sculptures of comic book characters are remarkably often produced in exactly this sense [... that they look like three-dimensional approximations of their two-dimensional drawings] is an indication that the images of comics are usually not used in such a way that the visible world is stylistically interpreted through them, but that the style of the image in a panel serves to present objects that possess this style themselves. The style of the images is interpreted by the viewer not as an interpretation of a visible reality, but as the design of a visible object itself [...]. The style of a comic book character is a property of the presented character (Balzer and Wiesing 2010: 62, my translation).

Intermedial transcriptions from comic books into animated films seem to support this idea. *The Peanuts Movie* film (2015) is a particularly good example for that: Even though its pictorial style, computer rendered 3D graphics, is materially and semiotically quite different from Schultz' drawings – the film contains colors instead of black and white pictures, the visible outlines have given way to simulated, shaded 3D bodies – they retain not only all the proportions and internal relations of bodies and faces, but also implement drawn lines within facial representations that approximate or remediate the aesthetics of the original cartoons (Fig. 3). Whereas the discontinuous comic pictures have been discussed as "an art of suggestion, not of mimesis" (cf. Lefèvre 2011: 29; Fresnault-Deruelle 1977: 31), continuous animation makes it easily possible to approximate 'direct glances' into an unmediated diegetic space that merely looks notably different from 'our' perception.

Fig. 3. *The Peanuts Movie* (USA 2015, directed by Steve Martino), 24:30.

In the real life/animation hybrid film *Who Framed Roger Rabbit* (1988), an animated cartoon world of its own physical laws is g e o g r a p h i c a l l y j u x t a p o s e d to the 'realistic world' (filmed with actors in front of a cam-

era). The distinction is thus turned into a part of the overall storyworld itself. When anvils fall on heads 'over there' – in the realm of cartoon characters – their inhabitants see stars and birds which are also visible as part of the diegetic environment. Cartoonization is thus expanded from a v i s u a l mode into a uniquely n a r r a t i v e affordance which 'literalizes' such visual metaphors (cf. Rauscher 2018). It would do cartoon aesthetics little justice to generalize a 'natural' world that is 'cleansed' from the conventionality of the representation (cf. also Limoges 2011). While all these cases vary significantly, we can still approach them from one of our three initial interpretational options: adding, subtracting, or aligning.

In this article, I will discuss how theories of multimodality are equipped to describe and reconstruct these options. While I have brought forth the same argument from the perspectives of Japanese studies and manga semiotics (Wilde 2020a), from phenomenology and cognitive semiotics (Wilde 2020c), as well as from transmedia narratology (forthcoming), I would now like to develop it from both Lars Elleström's (2019) and Charles Forceville's (2020) respective notions of multimodality (section 2). While both are certainly not the only, or maybe not even the most appropriate conceptions of multimodality in comics, their 'missing links' point in interesting directions. I am going to turn to the two (rather different) conceptions of multimodality by these authors, particularly because they developed their accounts specifically for narrative, fictional media – in Forceville's case for comics proper (2020: 185–216), in Elleström's from a transmedial perspective. In both publications, however, a Peircean notion of i c o n i c i t y is mostly taken as self-explanatory. In the subsequent section I am going to discuss serious shortcomings of such conceptions, especially with regard to two close readings of comics by German artist Sascha Hommer, one (*Insekt*, 2007, section 3) fictional and another (*In China*, 2016, section 4) non-fictional. I am then going to contextualize my findings with regard to a larger corpus of more prominent works and their discussion within comic studies (section 5). In the subsequent section, I am going to indicate how we can describe these artistic strategies – and the interpretational gaps they generate (section 6). While multimodality is well equipped to do so, as my conclusions shall show (section 7), we have to go beyond the discussions of Elleström's transmedia narration and Forceville's chapters on comics and turn to a more fine-grained notion of Peircean iconicity that has been proposed for multimodality before, but not with a focus on comics and their aesthetics of cartoonization.

2.   Multimodality and iconicity

In *Transmedial Narration*, Lars Elleström's recent (2019) multimodal approach to transmedia storytelling, the author accuses social semiotics of a "rather coarse notion of mode" (2019: 57). This seems especially true for problems of iconicity, as Bateman has pointed out as well:

> [I]n some approaches, such as social semiotics, there is little more than a passing mention of 'iconicity' using the traditional terms of 'resemblance' (2018: 18; cf. especially Kress and van Leeuwen 2006: 8).

Elleström instead proposes a "more fine-grained concept […] circumscribed as four kinds of multimodality" (2019: 58). Based on his rather idiosyncratic distinction between "modalities" and "modes" from 2010 (that, to my knowledge, not many scholars have accepted), the author distinguishes between four different kinds of multimodalities (each with a range of further internal "modes"), namely multimateriality, multispatiotemporality, multisensoriality, and multisemioticity. For our purpose, only the latter is relevant here, the "multisemioticity" between iconic and symbolic sign, in so far as they are most relevant for Elleström's understanding of representation:

> To say that a media product represents something is to say that it triggers a certain type of interpretation (Elleström 2019: 23).

This guided interpretation is then addressed, in unusual, yet precise terms, as "cognitive import" (Elleström 2019: 22) creating "virtual spheres" (2019: 24). The latter can be understood as mental models about a structure of represented events temporally interrelated in a meaningful way (cf. Elleström 2019: 39). Elleström's account of pictoriality in *Transmedial Narration* (cf. 2019: 49–58), as well as in his more detailed model of multimodality from 2010, however, uses a rather basic Peircean understanding of iconicity ("iconicity is based on similarity", Elleström 2010: 22). This can hardly account for different degrees of cartoonization and stylistic abstraction in comics discussed earlier. Instead, Elleström, too, subscribes to a variety of Mary-Laure Ryan's (1991: 48–54) p r i n c i p l e   o f   m i n i m a l   d e p a r t u r e , the assumption

> that one construes the intracommunicational domain as being the closest possible to the extracommunicational domain and allows for deviations only when they cannot be avoided (Elleström 2019: 27).

In other words, "collateral experience" (2019: 40) considerably shapes the virtual spheres. This determines our interpretational stances on the earlier three potential options (adding, subtracting, or aligning) firmly towards 1), adding something along the "reality" principle of minimal departure, or towards 2), subtracting something when it contradicts reality uncomfortably. Although Ryan (and others) include many thoughts on deviations from reality as a point of departure toward "generic landscapes" (cf. Ryan 1991: 52–57), this is usually not discussed in p e r c e p t i o n a l terms outside of comics' studies (Ryan dedicates a passing thought to perceptional deviation in 2014: 42f., but without developing it any further). To be fair, it is hard to argue against Thon's (2016: 90f.) claim that readers attribute the frequently changing drawing styles of a series with rotating artists (such as

Neil Gaiman's *The Sandman*, 1989–1996) merely to the medium of representation, not to the represented storyworlds and the characters in it. Morpheus, *The Sandman's* supernatural protagonist, does possess shapeshifting powers and is perceived differently by individual characters, but entirely human protagonists like Rose Walker also look strikingly disparate in interpretations of Mike Dringenberg (#10) or Marc Hempel (#65). In terms of storyworld properties, these differences are then certainly 'ignored' or 'subtracted' – not taken into account (see Gavaler 2022: 101–109 for a more extensive discussion).

In a similar way, yet arguing from an entirely different direction, Charles Forceville's recent *Visual and Multimodal Communication* (2020) proposes an authoritative model of multimodal meaning-making on the grounds of Cognitive Relevance Theory (cf. also Forceville 2014). He arrives at similar conclusions for cartoonish pictures of comics, as we shall see. Forceville builds his model on Dan Sperber's and Deirdre Wilson's (1995) relevance theory, assuming that communication is a process where an intelligent, human agent retrieves an alleged set of assumptions "made manifest" (Sperber and Wilson 1995: 58; Forceville 2020: 35) by an interlocuter. Pictoriality is taken as a distinct communicative mode which is not processed along stable codes, but according to a notion of iconicity, defined with Daniel Chandler (2017: 41; original emphasis) once again as a "perceived r e s e m b l a n c e or imitation." We interpret pictorial signs "because they very closely resemble objects, people, and events in everyday life" (Forceville 2020: 77). While it is not immediately apparent how lines in comic books can "resemble" entirely fictional entities, we certainly interpret all the character representations under discussion as anthropomorphic configurations, because humans and cartoons do share this very configuration or schema. The question, however, is whether – or, more precisely, t o w h a t e x t e n t – this schema, that we recognize in simple line drawings, necessitates a similar or a rather different a p p e a r a n c e from the material signs as part of the intended meaning: for how we imagine that the storyworld looks like, and for how we imagine what it looks like to the characters in it.

Forceville's model is less focused on imagination to begin with, but rather on propositional forms of knowledge made accessible by multimodal signs. Comics are addressed as v i s u a l l y c o n t a i n i n g (or at least triggering) explicatures like "Tintin and Snowy walk [in a certain way] in the direction of a hut in the wood/ jungle" (Forceville 2020: 192). Note that we have inserted a v e r b a l t r a n s c r i p t i o n here that 'cleanses' the represented entity from stylistic and perceptional aspects – although linguistic means are also merely representations of these explicatures and propositions. Both views are complementary or even deeply interrelated, as Elleström has addressed elsewhere, too: "We think both in an abstract way and in a concrete (visual and spatial) way" (2010: 22). We can also talk about all storyworlds, storyworld situations, and storyworld entities propositionally, since "[t]he narrated world is, strictly speaking, a world of singular facts" (Wulff 2007: 46, my translation; cf. also Wilde 2019a). The drawing style

(cartoonization), however, can then only be addressed in terms of 'reference assignment', 'ambiguity', and 'enrichment'. Reference assignment and ambiguity are initially only introduced in non-fictional terms (Forceville 2020: 74–78), but they must also be taken into account for characters that exist o n l y in comics: "consulting their mental lexicon, the addressees decide that these characters are Tintin and Snowy" (Forceville 2020: 190). Since we have no other, unmediated access to Tintin and Snowy, is this "consultation" based on iconicity/resemblance itself, or is it a matter of iconography – which is mostly understood as guided by convention? In other words: is the recognition of a human figure based on the same i c o n i c competence as recognizing Tintin? Regardless of how we might answer this question, it merely displaces the crucial issue of how the character Tintin is perceived within his world. Forceville's thoughts on "enrichment" make it clear that he d o e s conceive it in terms of 'omissions' that have to be 'filled in' by readers:

> in most comics, cartoons, and animation films, an artist deliberately l e a v e s o u t many details […]. Stick figures in some comics lack body parts […], and some manga artists o m i t characters' noses (Forceville 2020: 85, my emphasis).

Abstract cartoon drawings are then a form of "loose use of visuals" (Forceville 2020: 86), just as we sometimes 'omit' parts of a message in 'loose talk'. Again, we are back to our interpretational option 1), we have to 'add' something within our imagination – or to derive potential explicatures. Option 2) – subtracting something from the material means of representation – can be addressed as well (cf. Forceville 2016), but only as an exceptional deviation from the regular, less metaphorical form of pictorial comprehension. This is not only theoretically insufficient, I'd like to argue, it also makes us miss some of the most interesting thematic interpretations of comics which are entirely based on the hypothetical, yet often u n r e l i a b l e perceivability of characters within their world.

3.   Close reading #1: *Insekt* (2006)

I would like to analyze two very different works by German artist Sascha Hommer (*Insekt*, 2007, and *In China*, 2016). Born in 1979, Sascha Hommer is surely one of the most important German independent comic artists. Sebastian Bartosch and Andreas Stuhlmann describe his drawing style as follows:

> Referring back to the stereotypical drawing style of newspaper strips, he also borrows heavily from the industrial graphic design of the time as well as from classic Japanese comics such as Osamu Tezuka's *Astro Boy* (1952–1968) or the work of Hideko Mizuno (Bartosch and Stuhlmann 2013: 62–64).

Hommer's books are typical for a medium-specific narrative unreliability (cf. Packard 2018: 133) – precisely at the intersection of our three options for interpreting iconicity in comics. I borrow the term from the extensive literature on unreliable narrators (cf. Shen 2013) that addresses literary examples where it is not clear to what extent a narrating instance is communicating 'truthful' information about the narrated world. Apparent facts can thus turn out to be unreliable when they have to be revised later on due to contradicting new information, so that a reader has to assume they have been intentionally misled in their storyworld construction. In all the following cases it is not any aspect of the verbal narration that must be mistrusted and revised, however, but (aspects of the) pictoriality and the assumed iconicity between cartoonish drawings and their diegetic meaning. Whether we want to attribute the pictoriality of comics to a medium-specific 'visual narrator' or to the actual (or hypothetical) author is a disputed question (cf. Gavaler 2022: 184–191) that does not need to be resolved here, for it does not change the fact that parts of t h e n a r r a t i o n – whoever we might attribute it to – seem purposefully misleading and thus unreliable in their referential function. My two examples are also intended to show how these distinctions cannot be drawn between fiction and non-fiction, but remain an essential part of comics' pictoriality (or even mediality) across this divide. As Ole Frahm has aptly remarked (2011: 12), Hommer's comics open many spaces for a plurality of interpretations between multimodal texts and images. Such interpretational tensions and 'stitches' can also be observed w i t h i n the individual drawings and between them – but only with recourse to the overall diegetic world which rests on multimodal information. Readers can never be sure how the inhabitants of both Hommer's fantastic as well as of his autobiographical storyworld can be perceived by other characters. More importantly, these questions are even crucial for evaluating the thematic point of both works.

Hommer's 128-page book *Insekt*, published by Reprodukt in 2006, is set in a fantastic world in which some large, unnamed metropolis ('the city') seems to be shrouded constantly in a kind of fog or black haze. The comic is rendered in sharply inked black and white contrasts, many white areas are additionally darkened with halftone film (Fig. 4). This could certainly represent a perception clouded by fog – or obscured by poor lighting conditions. As readers, however, we can clearly see the sharp ink lines underneath as well as those representing the characters. On the mimetic (diegetic) level, the comic asserts something quite different, however: the city initially appears to be populated by human people, even though they are clearly cartoonized in stark exaggeration. In the case of children, head-to-body proportions roughly correspond to the extremely popular Funko Pop vinyl Figures, the heads being roughly the same size as the rest of their bodies (Fig. 4). Round eyes, in turn, take up a significant part of the head. This aesthetics is obviously quite common in comics; in manga, it is referred to as 'chibi' (or 'super deformed') (cf. Wilde 2020a). It is important to note, once again, that 'chibi-fication' usually leaves indeterminate whether it is merely

a stylistic device (abstracting from 'regular' humans) or intended as an element of the represented world itself.
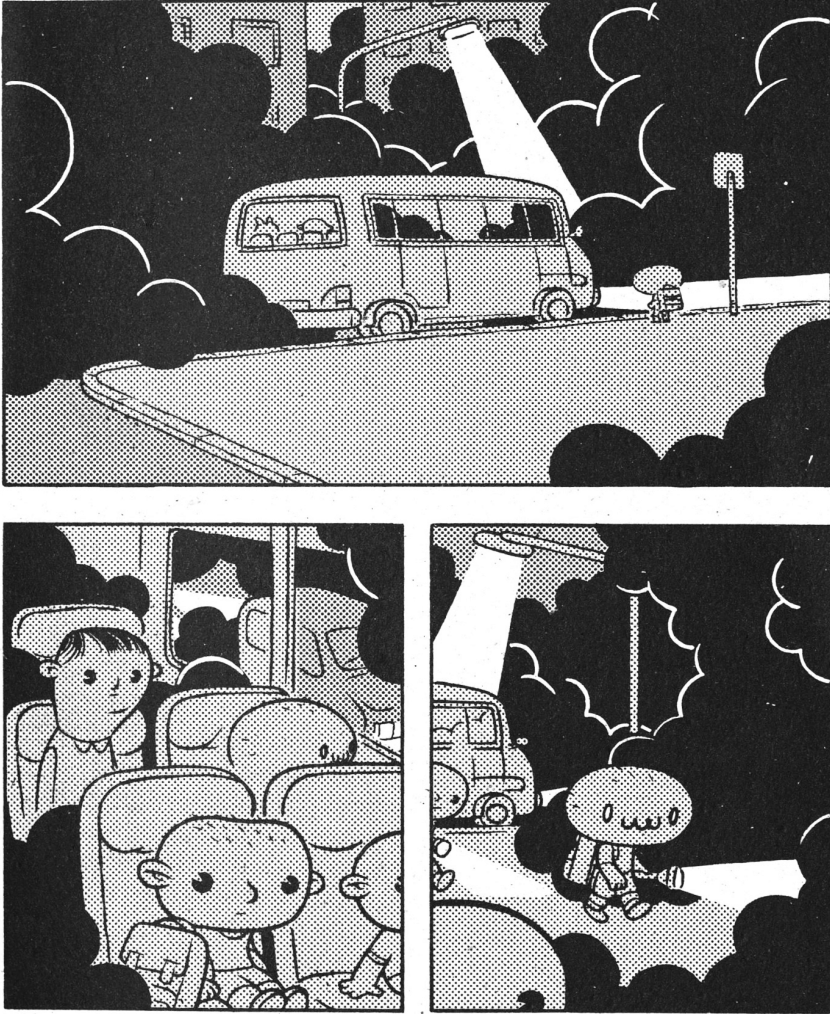
Fig. 4. Hommer (2006: 13).

Hommer does not stop there, however. In addition to the – stylistically interpreted or visually alien – humans, another species exists in the depicted world, to which our protagonist Pascal actually belongs, according to the verbal information: he is an "insect", as the title suggests, but not one that exists in our world, but a member of a fantastic, anthropomorphic species shrouded in myth ("Oh nonsense, that's just an old fairy tale with the insects, isn't it?," 2006: 24; "In truth, people don't even know what the insects look like – but they do exist!", 29; my translations). Pascal is not only of the same

size as humans and gifted with reason and speech, but also indistinguish-
able from the other inhabitants of the city due to the overall poor visibility
clouded by smog or haze. He himself does not even know that he belongs
to another species, and his diegetic environment has not yet recognized
this either. Like in a Platonic shadow world, the haze seems to inflict liter-
alized boundaries on the knowledge and awareness of all inhabitants. Only
those who leave the city and its veils of mist like the stranger leaves Pla-
to's cave, can truly perceive one another, and perceive one another as
members of a different species. Hommer thus uses the sense of sight as a
metaphor for knowledge and awareness. He builds on it a somewhat Kaf-
kaesque, certainly disturbing coming-of-age story about his protagonist,
whose youthful feelings of o t h e r n e s s, of not belonging, is literalized. As
an insect, he is indeed an alien being who disturbs and repels his child-
hood crush as soon as she looks into his eyes outside the confines of the
city walls (Fig. 5). Afterwards, he is bullied and brutally humiliated by his
classmates, forcing him to remove himself to his insect relatives outside
the city. The uncomfortable dynamics between the narrative and aesthetic
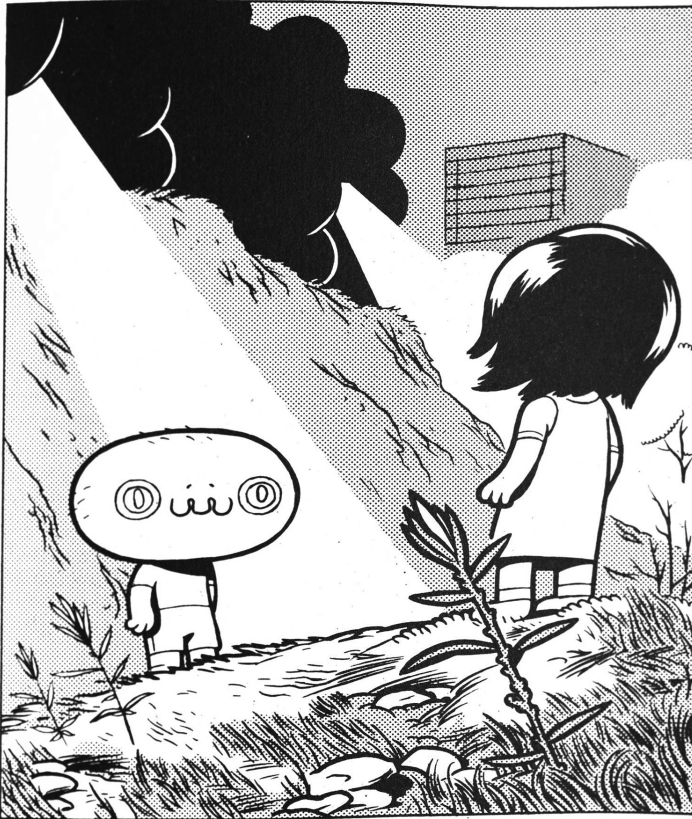level is regarded as a particular appeal of Hommer's work:

Fig. 5. Hommer (2006: 67).

The humiliations [Pascal's classmates] subject him to stand in shocking contrast to the childish scheme Hommer used (Welt.de 2007: para. 1, my translation).

Such symbolic interpretations – the implicit theme of the comic is accordingly Hommer's recurring motif of exclusion and alienation – certainly make up the actual relevance and appeal of the work. We can only discuss this motif through the multimodal construction of a primary, 'literal' diegetic world, in which an actual urban fog clouds all (self-)perception, where anthropomorphic insectoid beings co-exist alongside and with ordinary humans that are only (and barely) visually distinguishable from each other. Much about the actual perceptibility of *Insekt's* world must remain paradoxical, however. As readers, we can see even on the first page that Pascal looks different from the (Funko Pop) humans: The halftone film can hardly disguise his identifying contours. Its later absence will add little to the perceptual difference. Just as with theatrical conventions where a scene is intended to take place in complete darkness while the stage remains lit for the audience, we can never quite ignore a difference between what we see, and what it should represent. But what does his not-quite-girlfriend perceive at the moment when she catches sight of Pascal unveiled, especially in contrast to before (Fig. 5)? To what extent is it a matter of social projection, which is undoubtedly at play here: the insect creatures, after all, seem to be communicatively undistinguishable from humans, Pascal's identity went unnoticed for years. They are o t h e r e d for social reasons, perhaps not so differently from humans that are o t h e r e d and r a c i a l i z e d in our world (cf. Spivak 1985). The fictionality of the storyworld adds many questions about the alleged species-difference that cannot quite be resolved: Pascal's classmates are increasingly frightened by him, but not nearly to the degree that a supernatural shock would entail. Regular lessons are taken up again at school, and his classmates simply ask the teacher to replace Pascal as their class president. Is the entire story – as well as the difference between both species – merely a metaphor for entirely mundane, social forms of exclusion? While that seems entirely plausible, we cannot state where the metaphor begins or ends, just as with Gregor Samsa's transformation into an insect (cf. Kafka 2020):[1] Pascal's teacher even points out that what distinguishes insects from humans is their "balloon heads" (*Ballonköpfe*, Hommer 2006: 42; my translation) – exactly what we see depicted on top of the 'regular' humans as well, even if, perhaps, only due to aesthetic conventions. What is more, we only arrive at our thematic interpretations through inferences about Pascal's emotions and affects that we can see 'directly' from his face and posture. Just as in Kafka's most famous story, it remains programmatically open where exactly the boundary between 'literal' and 'metaphorical' meaning must be drawn, but Hommer achieves that effect through cartoonization. If the humans 'actually' have a 'photorealistic' appearance and Pascal a monstrous one, this would require readers to add many details to the schematic drawings (color, skin texture, etc.) and alter/replace others (head-body-proportions etc.), a combination of our earlier options 1) and 2). Or is this a c a r t o o n world in

which humans and insects are a l m o s t , but n o t   e n t i r e l y indistinguish-
able so that the othering rests on social, but mundane differences (option 3)?

4.   Close reading #2: *In China* (2016)

The same unreliability can be found in another comic by the same author,
even though this is taken as clearly non-fictional by most reviewers. *In China*
(2016) is based on autobiographical experiences by the artist. In 2011, he
helped a friend publish a magazine for foreigners in Chengdu for four months.
Chengdu comprises fourteen million inhabitants (at the time), the sky is
filled with bleak skyscrapers, there are continuous horns on the streets,
smog, again, clouds every street. The book has been compared to Delisle's
travel reports by reviewers (cf. Steinaecker 2016). Hommer similarly pres-
ents uneventful 'slice of life' episodes without much commentary. He medi-
ates his own everyday experiences abroad, his difficult search for an apart-
ment, taking Chinese classes, and his many conversations with other expats.
Meetings with locals remain the exception. If they take place, both parties
usually remain strangers to one another. The impossibility of truly 'encoun-
tering' a foreign culture like the Chinese is highlighted through intermedi-
ate chapters that explicate the thematic interest of the book. Each of its five
parts is accompanied by a drawn reproduction of some 'canonical' view on
China and on the cultural practice of travel, mostly from Western perspec-
tives or from the Western cultural imaginary (cf. Taylor 2004). Among them
are 'classics' like Marco Polo's travelogue, the VHS tape of *Bigbird in China*,
and the Tintin book *The Blue Lotus*.

Fig. 6. Hommer (2016: 68).

What strangely distinguishes the diegetic chapters, in contrast, is the fact that Hommer draws all the foreign visitors with animal and monster masks (Fig. 6). While Chinese residents are once again presented according to familiar manga conventions, the foreigners literally appear as aliens (while describing themselves so metaphorically on the verbal track, "we will always remain 'aliens' here", 2016: 53, my translation). At the same time, this makes it impossible – unlike the case of Pascal – to read any of their emotions, which adds a laconic and often depressing atmosphere to most events. But while the mask of Hommer's avatar appears as a static, physical object, his friend Linda has a highly stylized 'camel face' which d o e s change its cartoonish expressions. The effects of this technique on the overall mood of the story cannot be overstated, but the question of what these masks are supposed to m e a n, both on a mimetic (diegetic) and on a thematic level, is difficult to answer. Reviewers suggest that the masks "cleverly emphasize the role-like nature of one's own situation" in *Süddeutsche Zeitung* (Steinaecker 2016: para. 2, my translation) or that they indicate "skepticism about being able to encounter the foreign, to perceive the other" in *Taz* (Schirrmeister 2016: para. 2, my translation). If the discourse of exoticism in the intertexts makes it clear that a foreign culture can never be understood outside of metaphors or clichés, Hommer perhaps intends to turn this perspective around:

> They [the expats] have animal or even monster heads superimposed on them, which obscure their view of China and set them apart from the locals (Schirrmeister 2016: para. 2, my translation).

The Chengdu inhabitants, however, also remain indistinguishable, at least their pictorial representations. Pronounced manga conventions do not allow them any individuality, their faces remain template-like and devoid of distinguishing features (Fig. 6). While the 'masks' of the Chengdu residents are, at best, on the level of media conventions – concealing their perceivable features but making no claims on the visual ontology of the world – Hommer emphasizes at many points the l i t e r a l quality of his own's: In a key scene he will buy a new, different mask after watching a Sichuan opera performance, and continues to wear it from then on. When he directly asks another expat what the opera masks m e a n, that question could equally be directed towards the text as a whole. The answer remains equally unsatisfactory: "No idea. That's just the Sichuan variant of the famous Peking opera" (2016: 70, my translation). Later, at another expat party, Hommer even fails to recognize another friend, Markus, until the latter takes off his mask (in a lighting/coloring that makes it impossible for the readers to make out his face in turn) (Fig. 7).

Once again, the metaphor – if we even want to call it that at all – cannot be clearly deciphered, and once again the line separating it from a literal interpretation cannot be clearly drawn: should we take the entire work as generally autobiographical and imagine a storyworld that is 'disguised'

only where Hommer and his friends are concerned? Or, conversely, are we to imagine a diegesis in which all expats actually d o wear masks all the time, forming a more fictionalized metaphor for experiences abroad on a global textual level? What would both interpretations entail for the moment when Hommer decides to change masks to buy a new one, or when Markus reveals his face? What about the difference between his mask and Linda's? The added, aesthetic value seems to lie precisely in the fact that, once again, the line between the literal and the metaphorical cannot be drawn clearly – yet it must be drawn, somehow, in every reading, all the same! Again, we have to decide whether to enrich and replace certain aspects of the cartoonish representations towards a 'photorealistic' world, or whether to take them as closely at 'face value' as possible, merely in three dimensions instead of two, thus fictionalizing the overall storyworld.
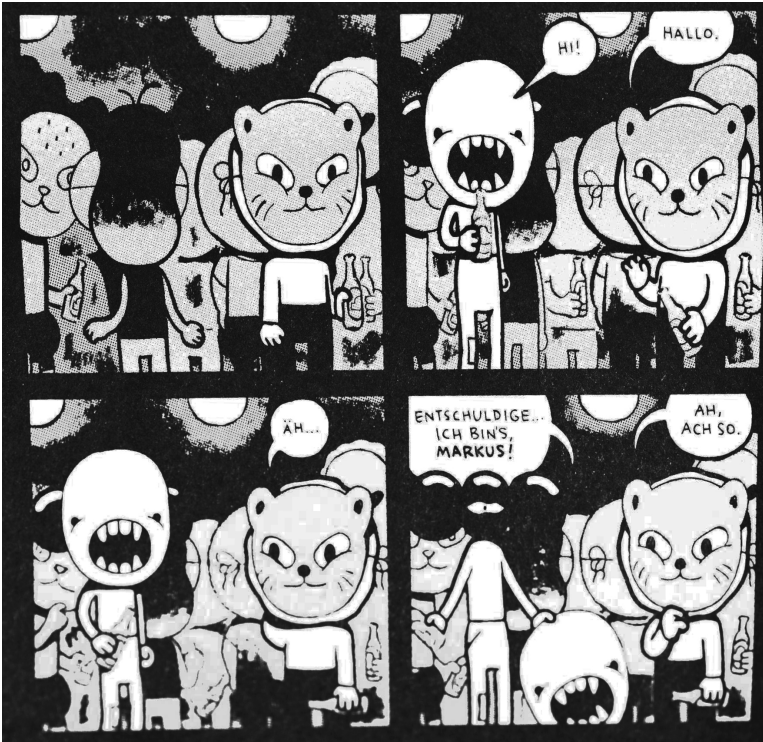

Fig. 7. Hommer (2016: 174).

5.   Proposals from comic studies: Referential meaning and the semiotic third space

Consider these two examples in the context of some other, perhaps more prominent works that have been discussed extensively in comic studies.

The genre of 'funny animal comics' is one of the most prominent and publicly visible within the medium, ranging back to groundbreaking works like George Herriman's *Krazy Kat* (1913–1944) and modern classics like Walt Kelly's *Pogo* (1949–1975) to newer genre subversions like Robert Crumb's *Fritz the Cat* (1965–1972) (cf., once again, Alaniz 2020 for a concise survey). The metaphorization and allegorization of pictures in comics through animal representations is certainly one of their most discussed aspects – usually, however, with respect to specific works (and their assumed strategies). The canonical example is Art Spiegelman's celebrated graphic novel *MAUS* (1980–1991; cf. Spiegelman 2003) retelling the autobiographical story of his father Vladek, an Auschwitz survivor. Although the work is clearly intended – and by now generally accepted – as non-fictional, Spiegelman represents Jewish people as mice, Germans as cats. Nevertheless, it should be clear to readers that these disguises are not to be taken literally, as Jan-Noël Thon had stressed in his comics narratology:

> [W]hat is represented here are not anthropomorphic animals but rather quite regular human beings whose affiliation with certain social groups is represented by more or less 'visible' but nevertheless exclusively metaphorical 'masks' (Thon 2016: 93).

Importantly, the meaning of this technique would be quite problematic if readers attributed it to Spiegelman, the author and artist of the book – whether taken as the actual, empirical creator, or as some overall, implied author of all multimodal (pictorial as well as verbal) elements. In fact, *MAUS* has initially been criticized for 'literalizing' a perceived ethnic difference, by reiterating "Hitler's racist thinking by casting groups as different species" (Spiegelman 2011: 131). After all, within Vladek's account, Germans and Jews can actually be perceived as biologically different races who cannot reproduce.

Fig. 8. Spiegelman (2003: 210).

Crucially, however, Spiegelman alternates his metaphorical or allegorical pictures of mice-humans with representations of humans wearing literal

masks attached to their foreheads with string. Some of the 'mice' within the historical narrative also disguise themselves by 'transforming' into other animals, a strategy that none of their co-protagonists ever see through – their animal representations are r e v i s e d (cf. Fig. 8). The reason for this is, of course, that it is not possible to tell whether someone actually 'looks Jewish'. Instead, the Jews-as-mice-metaphor must be understood as focalized, as bound to the ideological perspective of the characters, and the society in which they participate, to stress the point that the "racism [in National Socialist Germany] was all so arbitrary" (Spiegelman 2011: 132). The thematic meaning of the visible mice indicates that these differences are merely s o c i a l   p r o j e c t i o n s   about   a l l e g e d   ethnic differences. The aesthetics are hence employed as a media-specific means of characterization and subjectivization, to comment about the character Vlad and his experiences within his social surroundings. The fact that there have been extensive discussions about the contested meaning of these 'mice skins' shows two things: First, it is clear to most observers that Spiegelman does not want his readers to take these pictures literally, resulting in an ambiguity that can only be resolved on the thematic level once again: what does he (the actual or implied author, or perhaps 'the text' itself) want to convey here? The discussions result, secondly, in a number of   d i s t i n c t   proposals.

Fig. 9. Asano (2013: 54–55).

In other, perhaps even more challenging works (from a theoretical point of view), there are no identifiable options for metaphorical meanings to begin with, although most readers would still agree that the pictorial meanings cannot be taken literally. Inio Asano, for example, uses this technique to great effect in the manga series *Good Night, Punpun* (*Oyasumi punpun*, 2007–2013, cf. Asano 2016). The series, running to over 3,000 pages, tells of the depressing and often disturbing youth of a boy named Punpun, who is depicted as an abstract anthropomorphic configuration. The cartoon vaguely resembles a highly stylized bird (via the mark of a beak), while backgrounds and surroundings are rendered in the most detailed hyperrealism, with many pictures based on edited photographs (Fig. 9). Punpun's friends and teachers, in contrast, are drawn in manga aesthetics (especially their faces with huge eyes and almost absent noses), but nevertheless as clearly human. In selected key scenes, especially those connected to Punpun's first sexual experiences, individual body segments regain human traits in close-ups and angled shots. His bird-like regular appearance must thus be taken as another 'mask', leaving his actual perceivable identity completely indeterminate. The meaning of this device remains hard to pin down, however – it is hardly contingent on any connotations of 'birdness'. Punpun might be considered a center of subjectivity and unconstrained imagination which readers can easily empathize with, while the inexorability of his 'more intersubjective' lifeworld leaves little room for imagination or escape. In any case, calling this a 'pictorial metaphor' or 'allegory' – without being able to specify what it s t a n d s f o r – seems unsatisfactory.

While it would be tempting to connect this problem to Gilles Fauconnier's and Mark Turner's (2002) Conceptual Blending Theory, this would have to be done in a separate study. Instead I would like to stay within existing approaches of comic book aesthetics and semiotics, specifically Stephan Packard's (2017b) expansion of McCloud's concept of cartoonization that the former has developed into a powerful conceptual toolbox for analyzing comics and cartoons. It allows addressing comics' pictoriality on a global-textual, as well as on a local, character-bound (or domain-bound) level. Packard proposed a "semiotic third space" (cf. Packard 2017a; 2017b; Wilde 2020a) in which the 'duckness' of Donald Duck, the 'mouseness' of Vlad, and the 'birdness' of Punpun would be located if we were to decide that these characters were merely represented as animals, but 'actually' quite regular human beings in the storyworld (Gavaler discusses the same problems extensively as a problem of medial "transparency and non-transparency", 2022: 46–50). Comics' mimetic domain can be addressed as a cartoon's r e f e r e n t i a l m e a n i n g (cf. Persson 2003: 28; Thon 2016: 53; Wilde forthcoming). Since we can clearly see some stylized animal traits (although in dire need of enrichment) that would be neither part of our reality (where only lines on paper exist) nor of the referential meaning (where there would be a human being), we must assume a t h i r d d o m a i n distinct from both, as neither reality nor fiction seem to have a place for these aspects. That does not mean that the character traits in the third space are

redundant, or devoid of meaning. Often, there is meaning on the thematic, not the diegetic/referential level that can be interpreted through multimodal reasoning, as we have seen in Hommer's works (t o p i c s such as othering or exoticism). But sometimes the possible intentions behind the aesthetic device remain entirely undetermined. We can merely describe its effects in loosely defined 'aesthetic' terms. The thresholds between semiotic third space and referential meaning, however, are exactly what Hommer works with in his books, since we can never be sure about their precise delineations.

What we can see from these examples across different generic and cultural contexts is that various multimodal cues can guide our imagination in one direction or the other among the initial options 1., 2., and 3. (see p. 122):

a. Paratextual markers identifying a work as fiction or non-fiction: If we accept Delisle's work as non-fictional, it is clear that his line drawings are intended as vague, underdetermined abstractions from a more perceptually rich reality; although this does not dissolve all (or even many) questions with regard to *In China*, it is clear from the 'non-fiction' label that the manga aesthetics of his regular Chengdu inhabitants must be partly enriched, partly ignored, towards a more realistic visual ontology.
b. Generic traditions, including established 'reality principles' and rules of the imagination: In many parodic works, characters can be e x p e c t e d to perceive and manipulate 'extra-diegetic' signs (such as light bulb up-fixes) as if they were three-dimensional bodies in the diegetic space that look exactly like they are represented on the page; Superman can be expected to go entirely unrecognized as Clark Kent.
c. Specific textual cues, especially those concerning the perception of characters: When protagonists in *MAUS* put their animal 'masks' on or take them off without other characters noticing, these masks do not seem to exist within the referential meaning but within the semiotic third space; again, these cues must be related to paratextual markers and generic traditions, as we have to decide what sort of storyworlds we can expect in the first place and what seems ordinary or extraordinary within it (*Insekt*).

6.   Multimodality, iconicity, and diagrammatics

To account for the media-specific unreliability of cartoonization in comics, two different distinctions must be strengthened within theories of multimodality as provided by Forceville or Elleström earlier on. Arguing from relevance theory, a deeper consideration of Sperber and Wilson's (1995) distinction between the cognitive and the communicative principle of relevance turns crucial. This distinction, introduced for Forceville's overall theoretical design (2020: 33–40), remains largely neglected in his understanding of pictoriality. For Elleström's Peircean perspective, a more fundamental distinc-

tion between iconic and hypoiconic signs is helpful, which the author has provided himself in an earlier article (but not for narrative or fictional texts).

Generally, many discussions about the distinction between iconic, indexical, and symbolic signs isolate this trichotomy, while it is actually part of a much larger semiotic model (cf. Santaella Braga 1988, for instance). Without discussing Peirce's equally foundational trichotomies – that of representamen, object, and interpretant, of qualisigns, sinsigns, and legisigns, as well as that of rhemes, dicents, and arguments (cf. Nöth 1995: 42–45) – any discussion of iconicity must remain insufficient, because iconicity, indexicality, and symbolicity are "already situated somewhere in the 'middle' of the fuller account," as John Bateman has put it aptly (2018: 7). It is easy to get lost within these distinctions, however, as Peirce himself has revised his terminology (as well as his underlying concepts) many times during his career. For the context of multimodality, authors such as Elleström (2014) and more recently Bateman (2018) have provided thorough reconstructions and evaluations. Since I am only concerned with the specific problems posed by the iconic unreliability of cartoonization in comics – the distinctions provided by the semiotic third space in contrast to a cartoon's referential meaning – I am going to focus on my two proposed theoretical clarifications from above.

My overall suggestion is that 'pictoriality' in comics cannot be conceived of as a single mode, but as two distinct thresholds of pictorial comprehension and interpretation (cf. Wilde 2018: 89–213, more briefly Wilde 2020b): a mostly pre-attentional cognitive reconfiguration of two-dimensional lines on paper into three-dimensional bodies in space on the one hand, and a more conscious interpretational mapping (or d i s e n t a n g l e m e n t) of perceivable features to storyworld entities on the other, the latter guided by the aforementioned, often conflicting multimodal forces of specific textual cues, paratextual markers, genres and traditions. Iconicity is obviously not the same as similarity or resemblance, but semiosis b a s e d o n a n a p p e a r a n c e o f r e s e m b l a n c e (cf. Elleström 2013: 97). The 'regular' comprehension of more typical pictorial signs, such as photographs, is not merely – or not even primarily – a d e l i b e r a t e process of decoding, but "the result of interpretation also o n t h e s u b l i m i n a l l e v e l" (Elleström 2010: 22; my emphasis). Within relevance theory, communicative relevance is sharply distinguished from cognitive relevance, the everyday meaning-making according to schemata of sensorial perception. Cognitive semiotics and empirical research have shown that cognitive schemata allow us to perceive regular objects a s o b j e c t s and not merely as meaningless sensorial data, and that these schemata are stored according to principles of everyday relevance (for a species as a whole, within a specific historical or cultural context, but even for any one individual with all their past experiences, cf. Blanke 2003: 31). Cognitive types that have a higher relevance in the lifeworld will thus need fewer salient features to be recognized as such:

> There must be […] a Lifeworld hierarchy of most probable objects, beginning per-
> haps with the human body itself, in particular the human face (Sonesson 1989:
> 279).

Pictorial media can, at least to a certain degree, remediate the sensorial data provided by regular perception. Prototypical pictures can then be said to provide s u r r o g a t e   s t i m u l i of perception (cf. Eco 2000: 353–382). Since we can categorize pictorial signs at the same time as the actual shapes, lines, and colors that are materially present, iconic categorization has been conceived of as a d o u b l e   c a t e g o r i z a t i o n (cf. Blanke 2003: 62–70). The iconic categorization of surrogate stimuli allows three-dimensional objects to be 'seen in' (or rather: projected onto) two-dimensional surfaces. While the immediate pictorial comprehension above a certain iconic threshold of relevant sensorial data may be considered as 'purely' iconic, it does not allow us to make any claims or form any propositions yet, as Bateman has shown:

> [T]aking a painting, which is itself rhematic since it does not, by itself, assert, and
> adding a caption such as 'The Eiffel Tower', which, by itself, is also a rheme. The
> combination of rhemes may then lead to a dicent with content corresponding to
> 'This is what the Eiffel Tower looks like': only then can one respond to the painting
> and caption combination with 'truth' judgements and evaluations (2018: 13).

Such an indexical element does not need to point to any 'real' thing, however. In narrative (and especially in fictional) media texts it is primarily directed towards a storyworld which, by definition, is always already distinguished from its media and materials of representation. It is easy to see how our comprehension of sequential images necessarily transcends a more immediate, pre-attentive, 'purely iconic' understanding: If still images are read in a temporal sequence, this already adds a "symbolic element, namely the convention of sequential decoding" (Elleström 2019: 55). If this domain is established, aspects (or 'facets') of the material sign and its recognition can be transferred onto it, just as in Bateman's example of the picture of the Eiffel Tower. Elleström (2013) has reconstructed Peirce's distinctions between the more immediate, 'pure' iconicity from another form of hypoiconicity in greater detail (cf. Wilde 2018: 117– 129; Wilde 2019b; as well as Peirce 1932: CP 2.274–2.282). Hypoiconicity always entails a comparison, a process of mapping, because we are now transferring aspects or facets of the material sign and its recognition onto some other, distinct object. In the case of narrative representations, these can be entities within the storyworld. This comparison or transfer can be of three different kinds, distinguished by Peirce's internal trichotomy w i t h i n hypoiconicity, namely 'image/imagic', 'diagram/diagrammatic', and 'metaphor/metaphorical'. Even though this passage is notoriously difficult to interpret even for Peirce experts (cf. Braga 1996; Farias and Queiroz 2006; Colapietro 2011: 158), it precisely addresses the conundrum of cartoonization. If we consider a pictorial sign as (most-

ly) 'imagic', we transfer its perceivable qualities as closely as possible onto the represented object – 'something somewhere' is supposed to look like the material sign in front of us, to some degree. A distinction will always remain, however (a picture is usually flat, while the represented object is not; the picture is here 'with us', while the character exists in some other world).
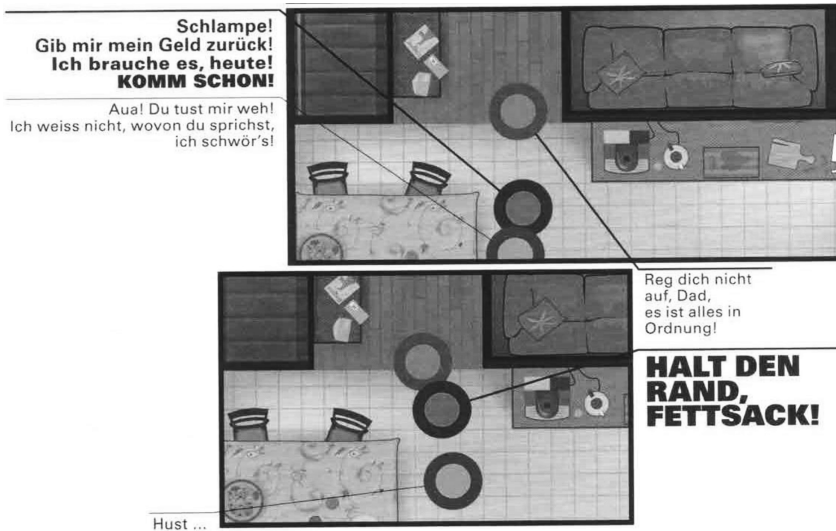
Fig. 10. Panchaud (2020: 48).

In contrast to such a (mostly) 'imagic' mapping, more diagrammatical forms of reasoning do not take any perceivable qualities into account (or considerably less so), but only the relations of the parts (cf. Stjernfelt 2007; Bauer and Ernst 2010; Krämer 2010 for different accounts of diagrammatics). Around the turn of the 20th century, Peirce described such a 'diagram' as a 'skeleton'. This evokes imaginations of a "skeleton diagram, or outline sketch" (Peirce 1932: CP 2.227). Elleström gives the following account: "[W]hile an image is a complete picture, a diagram is a sketch, characterized by the schematic relations of its parts" (2013: 101). The third form of hypoiconicity, 'metaphorical iconicity', does away with all physical resemblance – and all structural equivalence of physical resemblance – and only transfers aspects of the sign with regard to something else signified or implied by it. This is where iconic signs become charged with symbolism and conventionality. That is clearly the case for the 'mice' associations charged with notions of 'rodents' and 'vermin'. Cartoonization generally stays 'below' such clearly explicable metaphorical meanings, however, somewhere on two spectra of hypoiconicity incorporating varying degrees of both the imagic and the diagrammatic. We can see the most pronounced forms of diagrammatic hypoiconicity in works like *Die Farbe der Dinge* by Swiss artist Martin Panchaud (2020), wherein all the characters are represented by differently colored dots moving across a (much more 'imagic') map (Fig. 10).

The multimodal narrative d o e s refer to characters' body parts in verbal speech, however, and other diagrams even show their entirely human internal anatomy (Panchaud 2020: 77), or represent alcohol circulating through their bloodstream (2020: 131). Multimodal reasoning leaves no reason to doubt, then, that Simon Hope, Panchaud's protagonist, can be perceived as a three-dimensional human body in space, even though a gap between imagic and diagrammatic hypoiconicity seems to run exactly along the borders between characters and backgrounds. Interestingly, the semiotic third space closes in on the 'first' one here (the material means of representation), as the colored dots have no purely iconic meaning to begin with (if we do not wish to imagine them as spherical shapes or balls floating across space). They can be taken as strongly symbolic instead, perhaps as references to the digital location markers of GoogleMaps. They do retain s o m e hypoiconic (diagrammatic) aspects within the overall text, just like location markers do, by representing the exact spatial relations to other characters, their surroundings, and between each other. The s p e c t r a of (imagic as well as diagrammatic) hypoiconicity are thus entirely independent from material and physical properties (and 'looks') of the material signs. They are only determined by aspects we take into account in the process of the semiosis that constructs the storyworld alongside multimodal cues. Perceptually, characters can be based on images that we can comprehend 'subliminally' or pre-attentionally ('pure' iconicity), or they can be as abstract as Panchaud's. To what degree we decide to map what we see onto a represented plane, situation, or character (as a hypoicon) remains subject to various aspects of multimodal reasoning, guided by specific textual cues, paratextual markers, or generic traditions.

7.   Conclusions

If we decide to interpret a cartoon drawing alongside our initial option 1) (adding something), it has already entered semiosis as a diagrammatic hypoicon: we take it to represent only relations of body parts, facial expressions, differences between character sizes, and so on. The rest is added according to collateral experience or 'minimal departure'. The diagrammaticity can then be strengthened even further through option 2) (subtracting something), by 'blocking' not only many qualitative aspects such as colors or head and body contours, but disregarding a l m o s t e v e r y t h i n g of what the material sign is made of (and comprehended as in 'purely iconic' or 'imagic' terms). Perceivable features s h a r e d by Punpun, the cartoon bird, and Punpun, the human being, are almost entirely absent in terms of hypoiconicity. The most general spatial information is preserved, however (indicating where in the represented space the character is located), as are the body and face relations that signify posture, movement, affect, and emotion. This is where the semiotic third space opens up, 'swallowing' everything that does not enter into hypoiconicity and will thus not be mapped

onto the referential meaning. Additional metaphorical meanings c a n come into play (facilitated by the objects in the third space), but they can also remain vague, ambiguous, or entirely absent for most readers. If, on the other hand, we decide to close the third space and align our imagination as closely as possible to the material means of representation, the diagrammaticity is strongly deemphasized towards a more salient 'imagic' hypoiconic form of semiosis (although the diagrammatic information certainly remains equally 'valid'). We never take e v e r y t h i n g into account, however, as black and white flat pictures will usually be 'corrected' into a colorful, three-dimensional world. The spectra of hypoiconicity remain continuous scales, but there are discontinuous gaps and stitches even within one picture (between characters and backgrounds) and within one and the same cartoon (between facial expressions and aspects designated to the third space). According to relevance theory, this is decided according to communicative relevance (in comics through multimodal communication), distinguished from cognitive relevance (that is a mostly subconscious perception within the iconic mode).

Although the iconicity of a cartoon will hence always retain some ambiguity alongside the distinct options 1), 2), or 3), adding, subtracting, or aligning – an ambiguity that can be attributed to authors and artists as well as to individual characters and their subjective states themselves – it is far from arbitrary in any given context. As my examples have shown, this ambiguity can always be exploited artistically and thematically as a deliberate form of representational unreliability. We see in Hommer's works that the possibility to shift our attention across the hypoiconicity affordances between image and diagram (and, possibly, metaphor), between third space and referential meaning, offers one of the most powerful aesthetic resources of the medium of comics. Nevertheless, diagrammatic reasoning remains unguided in many cases, allowing readers to choose freely between options 1), 2), and 3). There is much to be done for a multimodal theory of cartoon comprehension, however. As I hope to have made clear, an account of multimodal reasoning could provide many tools to trace the aesthetic strategies of comics, especially across different semiotic modes beyond pictoriality. I could only indicate the broad range of specific textual cues, paratextual markers, and generic traditions that guide our attention along the hypoiconicity affordances and into, or out of, the semiotic third space. A more refined typology of reading instructions would be helpful, as would a r h e t - o r i c s of comic semiotics. Diagrammatics might be a suitable starting point for such endeavors, if it can be more clearly connected to theories of multimodality.

Notes

1   I would like to thank one of the anonymous reviewers of my article for this observation.

Bibliography

Alaniz, José (2020). Animals in graphic narrative. In: Frederick L. Aldama (ed.). *The Oxford Handbook of Comic Studies*. Oxford: Oxford University Press, 326–334.

Balzer, Jens and Lambert Wiesing (2010). *Outcault: Die Erfindung des Comic*. Bochum: Christian A. Bachmann.

Bartosch, Sebastian and Andreas Stuhlmann (2013). Reconsidering adaptation as translation: The comic in between. *Studies in Comics* 4, 1, 59–73.

Bateman, John A. (2018). Peircean semiotics and multimodality: Towards a new synthesis. *Multimodal Communication* 7, 1. URL: https://www.degruyter.com/document/doi/10.1515/mc-2017-0021/html [retrieved December 15, 2021].

Bauer, Matthias and Christoph Ernst (2010). *Diagrammatik: Einführung in ein kultur- und medienwissenschaftliches Forschungsfeld*. Bielefeld: transcript.

Berndt, Jaqueline (2013). Ghostly: 'Asian graphic narratives', nonnanba, and manga. In: Daniel Stein and Jan-Noël Thon (eds.). *From Comic Strips to Graphic Novels: Contributions to the Theory and History of Graphic Narrative*. Berlin: De Gruyter, 363–384.

Blanke, Börries (2003). *Vom Bild zum Sinn: Das ikonische Zeichen zwischen Semiotik und analytischer Philosophie*. Wiesbaden: Deutscher Universitätsverlag.

Braga, Lucia S. (1996). From pure icon to metaphor: Six degrees of iconicity. In: Vincent M. Colapietro and Thomas M. Olshewsky (eds.). *Peirce's Doctrine of Signs: Theory, Applications, and Connections*. Berlin: De Gruyter, 205–213.

Chandler, David (2017). *Semiotics: The Basics*. 3rd ed. London: Routledge.

Colapietro, Vincent M. (2011). Image, diagram, and metaphor: Unmined resource and unresolved questions. In: Pascal Michelucci, Olga Fischer and Christina Ljungberg (eds.). *Semblance and Signification*. Amsterdam: John Benjamins, 157–171.

Cook, Roy T. (2015). Judging a comic book by its cover: Marvel comics, photo-covers, and the objectivity of photograph. *Image & Narrative* 16, 2, 14–27.

Eco, Umberto (2000). *Kant and the Platypus: Essays on Language and Cognition*. London: Vintage.

Elleström, Lars (2010). The modalities of media: A model for understanding intermedial relations. In: Lars Elleström (ed.). *Media Borders, Multimodality and Intermediality*. Basingstoke: Palgrave Macmillan, 11–48.

Elleström, Lars (2013). Spatiotemporal aspects of iconicity. In: Lars Elleström and Olga Fischer (eds.). *Iconic Investigations*. Amsterdam: Benjamins, 95–118.

Elleström, Lars (2014). Material and mental representation: Peirce adapted to the study of media and arts. *The American Journal of Semiotics* 30, 1–2, 83–138.

Elleström, Lars (2019). *Transmedial Narration: Narratives and Stories in Different Media*. Cham: Palgrave Macmillan.

Farias, Priscila and Joao Queiroz (2006). Images, diagrams, and metaphors: Hypoicons in the context of Peirce's sixty-six-fold classification of signs. *Semiotica* 162, 1, 287–307.

Forceville, Charles (2014). Relevance Theory as model for analysing visual and multimodal communication. In: David Machin (ed.). *Visual Communication. Science*. Berlin: De Gruyter, 51–70.

Forceville, Charles (2016). Conceptual metaphor theory, blending theory and other cognitivist perspectives on comics. In: Neil Cohn (ed.). *The Visual Narrative Reader*. London: Bloomsbury, 89–114.

Forceville, Charles (2020). *Visual and Multimodal Communication: Applying the Relevance Principle*. Oxford: Oxford University Press.

Fauconnier, Gilles and Mark Turner (2002). *The Way we Think: Conceptual Blending and the Mind's Hidden Complexities*. New York: Basic Books.

Frahm, Ole (2011). Literatur und Graphik (Literature and Graphic Art). *Frankfurter Rundschau* 11 April, 2011, 12.

Fresnault-Deruelle, Pierre (1977). La Bande dessinée, ersatz de cinéma? In: Colloque international Education et bande dessinée (ed.). *Lecture et Bande Dessinée: Actes / Du 1er Colloque International Education et Bande Dessinée, La Roque-d'Antheron, 15-16 Janvier 1977*. La Roque d'Anthéron: Edisud, 25–31.

Gavaler, Chris (2022). *The Comics Form: The Art of Sequenced Images*. London: Bloomsbury Academic.

Herman, David (ed.) (2017). Animal Comics. Multispecies Storyworlds in Graphic Narratives. London: Bloomsbury.

Krämer, Sybille (2010). 'Epistemology of the line': Reflections on the diagrammatical mind. In: Olga Pombo and Alexander Gerner (eds.). *Studies in Diagrammatology and Diagram Praxis*. London: College, 13–38.

Kress, Gunther R. and Theo van Leeuwen (2006). *Reading Images: The Grammar of Visual Design*. London: Routledge.

Lamarre, Thomas (2008). Speciesism, Part I: Translating races into animals in wartime animation. *Mechademia* 3, 75–95.

Lefèvre, Pascal (2007). Incompatible visual ontologies: The problematic adaptation of drawn images. In: Ian Gordon, Mark Jancovich and Matthew P. McAllister (eds.). *Film and Comic Books*. Jackson, MS: University Press of Mississippi, 1–12.

Lefèvre, Pascal (2011). Some medium-specific qualities of graphic sequences. *Substance* 40, 1, 14–33.

Limoges, Jean-Marc (2011). Metalepsis in the cartoons of Tex Avery: Expanding the boundaries of transgression. In: Karin Kukkonen (ed.). *Metalepsis in Popular Culture*. *Narratologia*. Berlin: De Gruyter, 196–212.

Machin, David (2014). Multimodality and theories of the visual. In: Carey Jewitt (ed.). *The Routledge Handbook of Multimodal Analysis*. 2nd ed. London: Routledge, 181–190.

Molotiu, Andrei (2020). Cartooning. In: Charles Hatfield and Bart Beaty (eds.). *Comic Studies: A Guidebook*. New Brunswick: Rutgers University Press, 153–171.

Nöth, Winfried (1995). *Handbook of Semiotics: Advances in Semiotics*. Bloomington: Indiana University Press.

Nozawa, Shunsuke (2013). Characterization. *Semiotic Review: Open Issue* 3. URL: https://www.semioticreview.com/ojs/index.php/sr/article/view/16/15 [retrieved December 15, 2021].

Packard, Stephan (2017a). How factual are factual comics? Parasitic imaginations in referential cartoons. In: Reinhold R. Leinfelder, Alexandra Hamann, Jens Kirstein and Marc Schleunitz (eds.). *Science Meets Comics: Proceedings of the Symposium on Communicating and Designing the Future of Food in the Antropocene*. Berlin: Christian A. Bachmann, 19–27.

Packard, Stephan (2017b). The drawn-out gaze of the cartoon: A psychosemiotic look at subjectivity in comic book storytelling. In: Maike Sarah Reinerth and Jan-Noël

Thon (eds.). *Subjectivity across Media: Interdisciplinary and Transmedial Perspectives*. New York: Routledge, 111–124.

Packard, Stephan (2018). Der dramatische Reichtum der Multimodalität. Überlegungen zur Selbstorganisation semiotischer Ressourcen anhand von Sichtbarkeit und Visualität. *IMAGE. Zeitschrift für interdisziplinäre Bildwissenschaft* 18, 123–136.

Peirce, Charles S. (1931–1935, 1958). *Collected Papers*. 8 Volumes. Cambridge, MA: Harvard University Press.

Persson, Per (2003). *Understanding Cinema: A Psychological Theory of Moving Imagery*. Cambridge, MA: Cambridge University Press.

Rauscher, Andreas (2018). Die Cartoon-Modalitäten des Batman-Franchise und die Gemachtheit von Gotham City. In: Hans-Joachim Backe, Julia Eckel, Erwin Feyersinger, Veronique Sina and Jan-Noël Thon (eds.). *Ästhetik des Gemachten: Interdisziplinäre Beiträge zur Animations- und Comicforschung*. Berlin: De Gruyter, 305–334.

Ryan, Marie-Laure (1991). *Possible Worlds, Artificial Intelligence and Narrative Theory*. Bloomington: Indiana University Press.

Ryan, Marie-Laure (2014). Story/worlds/media: Tuning the instruments of a media-Conscious Narratology. In: Marie-Laure Ryan and Jan-Noël Thon (eds.). *Storyworlds across Media: Toward a Media-Conscious Narratology*. Lincoln: University of Nebraska Press, 25–49.

Santaella Braga, Maria L. (1988). For a classification of visual signs. *Semiotica* 70, 1–2, 59–78.

Schirrmeister, Benno (2016). Grafisch erzählt: Bibo in den traurigen Tropen. *Taz.de* March 31, 2016. URL: https://taz.de/Grafisch-erzaehlt/!5288462 [retrieved December 15, 2021].

Schüwer, Martin (2008). *Wie Comics erzählen: Grundriss einer intermedialen Erzähltheorie der grafischen Literatur*. Trier: Wissenschaftlicher Verlag Trier.

Shen, Dan (2013). Unreliability. *The Living Handbook of Narratology*. URL: https://www.lhn.uni-hamburg.de/node/66.html [retrieved December 15, 2021].

Sonesson, Göran (1989). *Pictorial Concepts: Inquiries into the Semiotic Heritage and its Relevance to the Interpretation of the Visual World*. Lund: Lund University Press.

Sperber, Dan and Deirdre Wilson (1995). *Relevance: Communication and Cognition*. 2nd ed. Oxford: Blackwell.

Spivak, Gayatari C. (1985). The Rani of Sirmur: An essay in reading the archives. *History and Theory* 24, 3, 247–272.

Steinaecker, Thomas (2016). Viel wunderschöner: Vier Monate hat Sascha Hommer 'In China' gelebt – Seine Eindrücke gibt es jetzt als Comic. *www.Sueddeutsche.de*. March 14, 2016. URL: https://www.sueddeutsche.de/kultur/comic-viel-wunderschoener-1.2897689 [retrieved December 15, 2021].

Stjernfelt, Frederik (2007). *Diagrammatology: An Investigation on the Borderlines of Phenomenology, Ontology, and Semiotics*. Dordrecht: Springer.

Taylor, Charles (2004). *Modern Social Imaginaries*. Durham: Duke University Press.

Thon, Jan-Noël (2016). *Transmedial Narratology and Contemporary Media Culture*. Lincoln: University of Nebraska Press.

Walton, Kendall (1993). *Mimesis as Make-Believe: On the Foundations of the Representational Arts*. Cambridge, MA: Harvard University Press.

Welt.de (2007). Herz für Loser – Der Zeichner Sascha Hommer. *Welt.de*. March 21, 2007. URL: https://www.welt.de/kultur/article769546/Herz-fuer-Loser-Der-Zeichner-Sascha-Hommer.html [retrieved December 15, 2021].

Wells, Paul (2009). *The Animated Bestiary: Animals, Cartoons, and Culture*. New Brunswick: Rutgers University Press.

Wilde, Lukas R.A. (2018). *Im Reich der Figuren: Meta-Narrative Kommunikationsfiguren und die 'Mangaisierung' des japanischen Alltags*. Köln: Herbert von Halem.

Wilde, Lukas R.A. (2019a). Kyara revisited: The pre-narrative character-state of Japanese character theory. *Frontiers of Narrative Studies* 5, 2, 220–247.

Wilde, Lukas R.A. (2019b). The epistemology of the drawn line: Abstract dimensions of narrative comics. In: Aarnoud Rommens, Benoît Crucifix, Björn-Olav Dozo, Pablo Turnes and Erwin Dejasse (eds.). *Abstraction and Comics/La BD et l'abstraction*, Vol. II. Brüssel and Liège: La 5e Couche and Presses Universitaires de Liège, 9–32.

Wilde, Lukas R.A. (2020a). Material conditions and semiotic affordances: Natsume Fusanosuke's many fascinations with the lines of manga. *Mechademia: Second Arc: Asian Materialities* 12, 2, 62–82.

Wilde, Lukas R.A. (2020b). The elephant in the room of emoji-research: Or, pictoriality, to what extent? In: Elena Giannoulis and Lukas R.A. Wilde (eds.). *Emoticons, Kaomoji, and Emoji: The Transformation of Communication in the Digital Age*. London: Routledge, 171–196.

Wilde, Lukas R.A. (2020c). Vom Bild zur Diegese und zurück: Bildtheoretische Rahmenüberlegungen zum narrativen Verstehen. In: Andreas Veits, Lukas R.A. Wilde and Klaus Sachs-Hombach (eds.). *Einzelbild & Narrativität: Theorien, Zugänge, offene Fragen*. Köln: Herbert von Halem, 88–123.

Wilde, Lukas R.A. (forthcoming). Die 'Cartoon'-Bildlichkeit von Comic, Manga und Animation: Referential Meaning, dritter Zeichenraum, heuristischer Zweifel. In: Margit Kern (ed.). *Visuelle Skepsis: Wie Bilder Zweifeln*. Berlin: De Gruyter. Preprint available at URL: https://www.researchgate.net/publication/374754896_Die_Cartoon-Bildlichkeit_von_Comic_Manga_und_Animation_Referential_meaning_dritter_Zeichenraum_heuristischer_Zweifel [retrieved April 11, 2024].

Wolk, Douglas (2007). *Reading Comics: How Graphic Novels Work and What They Mean*. Philadelphia: Da Capo Press.

Wulff, Hans-Jürgen (2007). Schichtenbau und Prozesshaftigkeit des Diegetischen: Zwei Anmerkungen. *Montage a/v* 16, 2, 39–51.

Sources

Asano, Inio (wri/art) (2013). Gute Nacht, *Punpun*. Vol. 1. Hamburg: Tokyopop.

Asano, Inio (wri/art) (2016). Good Night, *Punpun*. Vol. 1. San Francisco: Viz LLC.

Delisle, Guy (wri/art) (2005). *Pyongyang: A Journey through North Korea*. Montréal: Drawn and Quarterly.

Delisle, Guy (wri/art) (2008). *The Burma Chronicles*. Montréal: Drawn and Quarterly.

Gaiman, Neil (wri) and various (art) (1989–1993). *The Sandman*. New York: DC Comics.

Hommer, Sascha (wri/art) (2006). *Insekt*. Berlin: Reprodukt.

Hommer, Sascha (wri/art) (2016). *In China*. Berlin: Reprodukt.

Kafka, Franz (2020 [1912]). *Die Verwandlung*. Hamburg: Nikol Verlag.

McCloud, Scott (wri/art) (1993). *Understanding Comics: The Invisible Art*. New York: Harper Perennial.

Panchaud, Martin (wri/art) (2020). *Die Farbe der Dinge*. Zürich: Edition Moderne.

Rosa, Don (2014). Preface. In *The Don Rosa Library #1*. Seattle: Fantagraphics Books, 6–10.

Sattouf, Riad (wri/art) (2015). *The Arab of the Future 1: A Childhood in the Middle East 1978–1984*. New York: Metropolitan Books.

Spiegelman, Art (wri/art) (2003). *The Complete MAUS: A Survivor's Tale*. London: Penguin.

Spiegelman, Art (wri/art) (2011). *MetaMaus: A Look inside a Modern Classic, Maus*. New York: Pantheon Books.

The Peanuts Movie (USA 2015, directed by Steve Martino).

Who Framed Roger Rabbit? (USA 1988, directed by Robert Zemeckis).

Image Sources

Fig. 1. McCloud (1993: 36).

Fig. 2. Delisle (2008: 34).

Fig. 3. *The Peanuts Movie* (USA 2015, directed by Steve Martino), 24:30.

Fig. 4. Hommer (2006: 13).

Fig. 5. Hommer (2006: 67).

Fig. 6. Hommer (2016: 68).

Fig. 7. Hommer (2016: 174).

Fig. 8. Spiegelman (2003: 210).

Fig. 9. Asano (2013: 54–55).

Fig. 10. Panchaud (2020: 48).

Assoc.-Prof. Dr. Lukas R.A. Wilde
Norwegian University of Science and Technology (NTNU)
Department of Art and Media Studies
Edvard Bulls veg 1
Building 8, 8437, Dragvoll
7049 Trondheim
Norway
E-Mail: lukas.r.a.wilde@ntnu.no

Multimodal Cohesion in Panel Graphs:
A Pragmaticist Approach to the Gap Between Comics Grammar and Aesthetics

Stephan Packard, University of Cologne

Abstract. The essay presents a Peircean or pragmaticist approach to the contradiction between two kinds of theories about comics interpretation: One set which assumes the existence of a fundamentally lingual grammar in accordance with concepts of 'visual languages', and another set which insists on a free artistic interpretation as the force grounding interrelations between comics panels in sequence and sense-making in comics. The approach presented supports both positions and attempts to reconcile them by explaining in which two different senses comics do and do not have grammar. It examines the perceived gap between the grammar and the aesthetics of comics in a quotidian example taken from the pages of *Amazing Spider-Man*. While admitting a possible hermeneutic corridor between regular and singular, 'heautonomous' interpretations in the terminology of Romantic aesthetics, this position then historicizes such accounts and looks for alternative treatments of the same cognitive processes. By moving from a grammatical through a hermeneutical to a pragmaticist semiotic account of the resolution of ambiguities and irritations in interpreting a short panel sequence, the argument pivots on the Peircean continua of continued semiosis, of the interplay of prescision and abstraction, and of the gradual differences between simple and creative abduction, to outline three distinct conclusions for a multimodal model of cohesion in panel sequences that straddles the seeming divide between grammar and aesthetics: First, that such a hermeneutic corridor can be elaborated as a specific kind of r e v i s i o n a l   a t t i t u d e towards panel interpretation in sequence; second, that this allows the delegation of assumptions about conscious or reflective reading in favor of a comprehensive technical account of continuous interpretation; and third, that the historic contexts of previous interpretations have reasons to c o n f l a t e   d i s t i n c t i o n s between pictorial grammar and aesthetics with distinctions between conscious and automatic reading as well as between script and pictures, but that such conflations may no longer have to hold today.

Keywords. Comics, semiotics, hermeneutics, multimodality, pictorial grammar

Zusammenfassung. Der Aufsatz präsentiert einen Peirceschen, pragmatizistischen Ansatz zur Auflösung der Spannung zwischen zwei Theorien der Comiclektüre: Einer Gruppe, die eine grundsätzlich sprachähnliche visuelle Sprache mit entsprechenden Grammatiken voraussetzt, und einer anderen Gruppe, die auf einer freien künstlerischen Interpretation besteht, die die Beziehungen in Panelsequenzen und die Sinnstiftung in der Comiclektüre begründe. Der vorgestellte Ansatz will beide Positionen bejahen und miteinander vereinbar machen, indem er erklärt, in welchen zwei verschiedenen Sinnen Comics eine Grammatik haben und nicht haben. Er untersucht die wahrgenommene Kluft zwischen der Grammatik und der Ästhetik von Comics an einem alltäglichen Beispiel aus *Amazing Spider-Man*. Zwar lässt sich von einem hermeneutischen Korridor zwischen regelgebundenen und heautonomen Interpretationen sprechen. Eine Historisierung dieser Position erlaubt es jedoch, andere Theoriebildungen über die so beschriebenen Kognitionen in Betracht zu ziehen. In einer Bewegung von einer grammatischen und einer hermeneutischen zu einer allgemeineren, pragmatizistisch semiotischen Konzeption der Semiosen, mit denen Ambiguitäten und Irritationen in der Comiclektüre durch Revision bearbeitet werden, geht das Argument zu einer Betrachtung dreier Kontinua im Sinne der Peirceschen Semiotik über: fortgesetzte Semiose, 'Präszision' und Abstraktion, und der Übergang von einfachen zu kreativen Abduktionen. So ergeben sich drei Schlussfolgerungen für ein letztlich multimodales Modell der Kohäsionsbildung in Panelsequenzen, das die scheinbare Kluft zwischen Grammatik und Ästhetik übergreift: Erstens lässt sich die Spezifik einer r ü c k b l i c k e n d e n E i n s t e l l u n g in der Comiclektüre beschreiben, die sich zwar mit einem hermeneutischen Korridor vergleichen lässt, aber auch andere Deutungen zulässt. Zweitens lassen sich so Annahmen über die *b e w u s s t e o d e r r e f l e k t i e r t e* Qualität der Comiclektüre zugunsten einer umfassend technischen Beschreibung kontinuierlicher Semiose zurückstellen. Und drittens kann der historische Kontext anderer Deutungen die V e r m i s c h u n g konzeptuell zu trennender Distinktionen zwischen Grammatik und Ästhetik, zwischen bewusster und unreflektierter Lektüre sowie zwischen Schrift und Bild zwar motivieren, kann aber heute an Bindungskraft verlieren.

Schlüsselwörter. Comics, Semiotik, Hermeneutik, Multimodalität, Bildgrammatik

1. In what sense do comics have grammar?

Studies on the regularities of visual language, especially by Cohn (2013 and beyond), have shown conclusively that comics are subject to cohesion, i.e. unifying principles that bind together their elements, beyond the mere logical coherence of their semantic reference: In other words, as we look for connections between several panels and their various parts, we no more have to start reasoning in a vacuum about the meaningful structures employed in the pictorial sequence than we need to consciously reinvent language through cumbersome code-breaking every time we decode a sentence. There are grammatical rules that go beyond, and indeed before, those implications of world knowledge, artistic appraisal, and sensemaking that are involved in the

active and deliberate inference of connections between separate depictions in the anatomy of the comics pages' assemblies. While the latter aspects of coherence all contribute to a general sense of cohesion, I will distinguish the two terms here and limit 'cohesion' (cf. Halliday and Hasan 1976) to relationships of elements within a media artifact that are formally and explicitly established. I will mostly focus not on (the visual equivalent of) lexical but on grammatical coherence, i.e. referential chains established in accordance with rules governing the syntactical arrangements of signs, belonging to the domain of legisigns rather than symbols in Peirce's terminology: These relations are mostly at issue when claims to a visual language in comics are defended or disputed. Crucially for the ensuing argument on historicizing hermeneutics, I will look at grammatical cohesion proper as encompassing qualities of artifacts rather than as cognitive interactions (cf. e.g. Menzel et al. 2017), setting apart some of the elements considered under the same name in recent treatments in multimodal linguistics (cf. e.g. Tseng and Bateman 2018; Stainbrook 2016), with which this account should however remain compatible.

However, many appraisals of comics' specific aesthetic devices have emphasized a distinction between rule-bound grammar and comics' cohesive structures. They describe the experience and cognitive performance of meaning construction between panels and panel elements as essentially different from the lingual comprehension of written and spoken words. While the latter is considered to be more direct, opaque, and determined, the former is proposed to be always already an intelligent and intense engagement with art and meaning (e.g. Groensteen 1999; Grünewald 2014; and most explicitly Grennan 2017, chapter 1.2.2). As Groensteen puts it: "La bande dessinée repose […] sur un dispositif qui ne connaît pas d'usage familier." (1999: 23) – 'Comics depend on a dispositive without familiar usage.' In this view, comics should be opposed to language. For lingual art, literary or poetic modes can be distinguished from everyday usage by their differing linguistic qualities. This opposes them to comics, which are, in the same view, always already artistic. They depend upon and elicit a strong interpretation, which not only goes beyond, but is liberated from conventional comprehension. It by no means eschews rules, but constructs rules beyond or even contrary to those taken from grammar. If this conception in which poetic modes could escape from grammar were true, then we might be tempted to accept that comics are understood in a manner fundamentally different to language; that this difference is due to the specificity of their pictorial code; and that this specificity is about a lack of grammatical regularity. In this case, a specific tension connected to the multimodal combination of script and pictures would be essential for understanding comics, and a theory that repeats that distinction would be indispensable for understanding comics comprehension. But how can this view be reconciled with the many empirical demonstrations of the grammar of visual language?

On the following pages, I want to outline the requirements for a semiotic conception of comics that can operationalize this seeming contradiction by clarifying the two senses in which we may say that comics have gram-

mar on the one hand, and that their poetics lack a conventional grammar on the other hand. Because I find the evidence for both claims to be overwhelming both in generalized arguments and perhaps even more so in many individual readings of specific examples of comics art, I believe we need to find a way to affirm both positions at the same time. Obviously, this can only be possible if we are talking about two different things in each case. The challenge is that these two aspects nevertheless seem closely connected and can hardly be satisfactorily described in separation. What I will explain as the h e a u t o n o m o u s poetics of comics cannot be conceived of as some merely additional quality that linguists may ignore, because it apparently engages and even interferes directly with the same structures of sensemaking that are involved in comics' grammatical regularities. Vice versa, any artistic appraisal of comics' devices that ignores their grammatical conditions will miss at least large parts of the picture. The interplay between the two positions, or so I will argue, reflects an actual tension between grammar and aesthetics in the comic books themselves.

I want to discuss this semiotic cleft from four angles – grammatically in terms of some of the categories proposed for comics' visual language (2), hermeneutically in terms of the Romantic history of the ideas involved (3), strictly semiotically in terms of Charles Sanders Peirce' generalized theory of signs (4), and as a specific kind of staggered and hybrid modalities in terms of recent concepts of multimodality (5). Through the course of this argument, the most general approach from semiotics should work as a means of transference between the divided conceptions of empirically founded grammar and interpretatively liberated hermeneutics, and a unified model of multimodal comics comprehension. Crucially, the Romantic question of conscious or sophisticated interpretation as opposed to conventional and determined comprehension can be dissolved from the point of view of a semiotic theory for which consciousness is epiphenomenal (cf. Baltzer 1994; Colapietro 2014) and, for any individual case of semiosis, a facultative addition to an otherwise independent explanation of sign processing. We need not know whether a comics reader knows what they are doing in order to describe what they do.

2.   Grammar: Panel elements and revision

It follows that the phenomena that I want to look at here are of a lower or middle range in terms of their aesthetic quality. They occur across comic book syntax, and relate to specific devices employed more so than unique ruptures in comprehension. This separates them from another sense of aesthetics or poetics that focuses only on rules broken or flaunted, phenomena that do occur in comics as they probably do in all communication but are separate from the debate of grammaticality examined here. I thus avoid any especially unusual or marked moments of aesthetic interruption and foreground minor but pertinent devices that constitute poignant shifts in understanding as a part

of the ongoing progressive comprehension of a typical panel sequence. Crucially, their effect might or might not be explicitly conscious for readers.

Fig. 1. An emergent pipe. Spider-Man battling the Green Goblin in ASM #122: 4.

Consider this page from *Amazing Spider-Man* #122 from July, 1973 (Fig. 1, Conway and Kane: 4). The two triangular panels in the lower half describe a quite sudden reversal in the depicted combat between Spider-Man and his nemesis, the Green Goblin – exactly the kind of moment that is expected to occur in such combat scenes in superhero comics with some frequency, and with some flourish, but without breaking the format. If it first appears as if Spider-Man is gaining the upper hand, pummeling the Goblin as he sits on his shoulders while they ride in the air on the Goblin's hovercraft glider, this changes in the second panel, in which Spider-Man collides with a large steel beam and is thrown off the shoulders of his antagonist. What interests me about this sequence is the manner in which this reversal redefines the semiotic quality of the green beam. For Spider-Man, it suddenly appears – one might describe the immediate connection as an exemplary i n d e x i a l experience, in which the steal beam draws Spider-Man's attention by immediately connecting with his forehead, painfully. But we are interested in the readers' semiosis, not the hero's. If we look back at the previous panel, the beam is undeniably already there. In fact, upon closer examination of the first triangular panel, we might even be uncertain about the exact placement of the beam, Goblin's head, Spider-Man's fist, and Spider-Man's head. It is the important function of the beam in the second panel that encourages such a renewed consideration of the first panel. Upon such reflection, we might find the left triangular panel to be somewhat ambiguous in this regard, drawing on a vagueness in Gil Kane's artistic style that is demonstrably felicitous, though again not necessarily consciously intentional (although it very well might be).
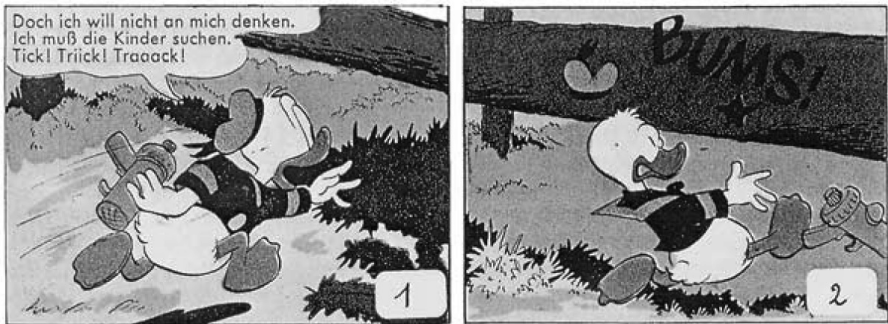
Fig. 2. An emergent branch. Donald running through a forest as cited in Krafft 1978: 48.

In one way, this example is certainly similar to a sequence that Ulrich Krafft, in his 1978 treatise on a textual linguistics of comic books, puts at the heart of his argument about different domains structuring consecutive panels. Here (Fig. 2, Krafft 1978: 48), Donald's head collides with a tree trunk in much the same fashion. Donald is surprised; but again, when we look back from the second to the first panel, we see that the trunk was clearly already there. In Krafft's model of panel domains governing referential movement,

this moves the trunk from the domain of spatial signs to object signs and even props, i.e. objects with which the characters directly interact. Carl Bark's art does not necessarily shift its depiction of the trunk from one panel to the next except for the fact that in the second panel, the contour of Donald's cartoonish character is directly connected to the circumference of the trunk; this is the main distinguishing stylistic marker for Krafft. Beyond it, the shift in domains in his theory is down to top-down ascriptions as readers make sense of the reappearance of elements across panels in continuous sequence. Functionally, the trunk in the second panel could not be erased without changing (or perhaps even destroying) the meaning of the picture; whereas spatial signs, according to Krafft, are specifically characterized by being subtractable. One might express a reversal of this idea (cf. Packard 2006: 207) by pointing out that comics spaces are productive, and allow for the emergence of object depictions in repeated spatial depictions.

The attention due the subsumption of graphical affordances under different panel elements reflects a number of claims about their functions in comics' comprehension, as well as possibly in their aesthetic appreciation. One first function throughout all similar theories is the precise description of different constituents of comics panels and pages. This would cover all attempts at a comics markup language, through to the most elaborate schemes for elemental annotation that can ground further corpus analysis (cf. Bateman et al. 2016). Krafft's model is more closely linked with one specific theoretical superstructure: distinguishing persons and objects across foreground and background, he arrives at four fundamental functions which are then taken up by the domains of action signs and spatial as well as canvas signs ('Handlungszeichen', 'Instrumente', 'Raumzeichen', 'Folienzeichen'; Krafft 1978: 15–41). While action signs are directly involved in marking the beginning, continuation, and suspension or replacement of phoric chains, spatial signs are only indirectly governed by such necessity. Krafft connects this logically to the way in which the reading process keeps track of signs: action signs can be enumerated, they are regularly repeated as long as the sequence keeps referring to their referents, and they are clearly delineated. Meanwhile, spatial signs can be added or eliminated with much greater liberty, their contours are often interrupted by panel borders, their internal structure of core elements and affixes is more liberally treated, and they are numerically indistinct: we know that there is only one Donald, or that there are only three of his nephews, in a given setup; but we do not expect the number of trees, bushes, or blades of grass to remain consistent. Surprisingly and hence persuasively, this even extends to architectural elements such as doors and windows, dissociating these rules of grammar from those of logic or even general concepts of salience. The specific salience of elements for each story told is upheld, tantamount to the concepts of focus or theme in other accounts of phoric movements. Any attempt at classifying domains within panels should be at least compatible with a concurrent theory of sequential construction or comprehension; additional reasons for this will become clear below.

A third possible use of panel domains as analytical categories bridges the intention to describe sequences and the goal of categorizing stylistic qualities. If action signs are more clearly and fully delineated and the contours of spatial signs do not need to be realized, even a single panel can suggest if not determine possible uses for different graphic elements. Here, Krafft's efforts to describe the compositional personhood of typical action signs (1978: 35) dovetails nicely with McCloud's famous emphasis on the differing drawing style of cartoons, usually employed as action signs, and less cartoonized backgrounds, which would usually be subsumed in Krafft's spatial sign (1994: 33–40). Both McCloud and Krafft, for differing reasons, consider an interim domain for objects which are taken up by, touched by, or interacted with by more fully realized agents depicted by action signs. Other models for the enumeration and distinction of types of panel elements and their interaction have suggested other motivations. A psychosemiotic model developed from Krafft's distinctions but founded on Peircean semiotics and Lacanian psychoanalysis has been presented (Packard 2006) for which cartoonized action signs, props, spatial signs, and canvas signs are all differentiated both by their function for the panel sequence and by their different interactions with the human gaze. Cohn has presented models founded on rich empirical accounts that continue the distinctions of Ray Jackendoff's parallel architecture model (Jackendoff and Audring 2020) into a six-part architecture comprised of external compositional structure, narrative structure, and event structure for the qualities of the sequence, corresponding respectively to a 'graphology', a 'visual morphology', and a lower level conceptual structure for the classification of units (e.g. Cohn and Schilperoord 2022). Notably, all of these models include a similar distinction between action and spatial signs, or more generally, activated signs and others. To the extent that visual language has a grammar beyond semantic comprehension, this distinction may be related to that between persons and objects, but cannot be reduced to it.

It is thus possible to think of the example in Krafft as a straightforward use of a comics grammar. In his model, the elements of each panel slot into specific domains, and the referential movement here is established precisely by shifting one element from one domain to another. All of this is also true for Kane's art on the *Spider-Man* page. However, on that page, several additional elements prepare and echo the reappraisal of that steel beam. First, the dynamic cut of the triangular panels assembled to form a rectangle on the lower half of the page draws attention to the different angles at which the beam appears. In fact, the two depictions of the beam and the diagonal panel gutter combine to roughly trace the shape of the 'Z-path' described by the gaze in a regular panel sequence comprehension in this genre (cf. Cohn and Campbell 2014). We rediscover that path, previously merely a transparent rule for grammatical comprehension, as a now-topical element of interpretation. This is not true for the trunk that stops Donald, whose shape and orientation in panel 2 is repeated almost identically from panel 1. Moreover, the three panels in the upper half of the *Spider-Man*

page distinctively and probably deliberately eschew clear spatial information. The third panel shows a close-up of the combatants' faces; the second panel arranges an overstated trail of smoke and condensation to fill the space of the panel with the path taken by the hovercraft, as opposed to filling in other markers that would place that movement in a relationship with the rest of the world. Crucially, the first panel takes up the previously established architecture from which the combat erupts briefly and in a compressed fashion, without giving any hint to the relative placement and later appearance of the beam.

One function of this first panel is to move the scene away from the dominance of the space used in the previous sequence(s). On several pages in this and the previous issue, the cliffhanger that connects the episodes revolves around Spider-Man kneeling on a platform, Gwen Stacy's dead body on his lap. This issue takes up the plot after the now-famous, then-surprising death of this central character in *Spider-Man*'s continuity. In a parallel break with narrative expectation, Spider-Man has vowed to kill the Green Goblin in response to the murder of Stacy. The scene examined here seems poised to end with that very killing. The movement away from the open and clearly laid out platform to the cluttered architecture in which a steel beam may suddenly appear thus parallels the movement of the plot, which suspends and retards the following action to give Spider-Man some time to struggle with his revenge. In the eventual resolution (Fig. 3, Conway and Kane 1973: 20), a similar play with spatial and object signifiers has the Goblin's glider surprise Spider-Man from behind and kill the Goblin just as Spider-Man is still trying to decide whether to kill the Goblin, saving Spider-Man from making that decision and resolving the plot of this revenge tragedy without a final commitment from the protagonist. Here, the movement of the hovercraft in the fourth panel encourages re-examination of the second panel, where a similar graphic ambiguity as before is at play regarding the distance between the hovercraft and Spider-Man's back, and the angle at which it approaches the Goblin.

In both cases, Kane's pages offer elements that connect to a re-appraisal of the beam or the hovercraft as spatial or object markers respectively. Where Bark's panels convey the suddenness with which Donald connects to the trunk by first hiding it in the spatial domain and then taking it into the object domain, Kane's pages are apt to be re-considered at length precisely with this question in mind: Where did the beam come from? What makes its appearance so sudden? And eventually, the answers to those questions lead into the deeper meaning of the plot. These ultimate consequences for the interpretation of the revenge tragedy are of course not within the scope of a panel grammar. In contrast, the distinction of spatial and object markers very likely is grammatical. But the revision invited by the elements paralleling the domain shift for the steel beam on Kane's page would be difficult to explain with any degree of precision without these grammatical categories, and yet goes beyond a straightforward referential movement. The panels have grammar, in the sense that the graphical element depicting the

beam shifts between grammatically defined domains, as does the depiction of the trunk. But their additional interrelations with surrounding elements contextualize this shift, and do so crucially on the same formal and elemental level, by operationalizing the vagueness involved in the assignment of pictorial elements to panel domains, as well as the subtractability or productivity of comic panel spaces. The panel sequence has a grammar in the sense that it regularly distinguishes spaces from objects; its poetics are heautonomous in that they recreate and reappraise that rule through different aspects of the 'iconic solidarity', as Groensteen (1999) describes it, that binds these images together.

Fig. 3. An emergent solution to a tragedy of revenge. Spider-Man failing to kill the Green Goblin in ASM #122: 20.

3.   A hermeneutic corridor in comics comprehension

If we want to connect this element of comics interpretation to established vocabulary in the humanities, we might consider it a h e r m e n e u t i c aspect that comes between the scopes of g r a m m a r on the one hand and either a e s t h e t i c s or r h e t o r i c a l n a r r a t i v e on the other. In his influential account, and in a context of interlacing concepts between Idealism and Romanticism, Schleiermacher (1809/10) has presented the practice of hermeneutics in relation to – but not completely convergent with – grammatical comprehension. Hermeneutics famously begins with the fact of not (yet) understanding, and hence consists in a conscious and reflective engagement with the material. At the same time, the grammaticality of the text interpreted is not just assumed, but becomes one of two structures – alongside the technical or psychological dimension following expectations shaped by genre or authorship – involved in that conscious reconstruction. Schleiermacher speaks of a g r a m m a t i c a l i n t e r p r e t a t i o n in as much as it r e - considers lexical and morphological elements not understood fully at once, in order to better understand them eventually. Thus hermeneutical interpretation is distinguished from direct grammatical comprehension, but not from grammar. In a similar sense, the specific devices pivotal to the appraisal of comics' art might be considered to be realized in a hermeneutic corridor between rule-bound grammar and 'heautonomous' aesthetics in the sense of Romantic and some Idealistic accounts of aesthetics that have been continued and evolved from 18th and 19th century notions. This notion of heautonomy denotes the autonomy of ordered structures that ground the rules of their order within the confines of their own architecture rather than merely instantiating external rules established elsewhere and holds for poetic language whose internal regularities go beyond and sometimes even against standard grammar (cf. Homann 1999). Crucially, neither half of the semiosis of comics can be fully explained without a reference to this corridor between them.

It might be useful to point out what this corridor, and the distinct parts of comics' semiosis between which it mediates, are not. Their distinction is not the same as the (nonetheless relevant) distinction between top-down and bottom-up subsumption of graphical elements to grammatical constituents. When Krafft traces the function of Donald's cartoon, the pertinent tree trunk, and the other flora depicted in the Barks sequence by separating subtractable from non-subtractable elements, he argues top-down; when he points out the continued contour of Donald's and the trunk's circumference as well as Donald's regular arrangement of bodily constituents, he argues bottom-up. Both arguments pertain to comics' grammar. Conversely, when we find the arrangement of elements in Kane's first triangular panel ambiguous, we argue bottom-up; when it is the arrangement of both triangles that draws our attention to this element, we argue top-down. Both movements are involved in the hermeneutic circle, which travels from the general to the particular and back as many times as needed, and this oscil-

lation further characterizes this part of the practice of reading comic books as properly hermeneutic.

Equally, while this interpretation agrees with accounts of comics' special aesthetics that foreground heautonomous semiotic modes rather than grammatical semiosis, the border between grammatical and heautonomous aspects does not coincide with that between pictures and script, or between mere script and a text-image-combination. This is in contrast to several accounts (e.g. Grünewald 2014), and especially some treatments in critical theory, that conflate the semiotic opposition between grammar and heautonomy with the distinction between text and images, and sometimes between semiotics and meaning altogether (e.g. Frahm 2010). Text and image are inarguably combined in all examples discussed here; but their combination is not intrinsic to the re-examination of pictorial elements proposed. Many comics employ no script whatsoever, but those 'pantomimic' comics would still be subject to the hermeneutic interplay between a visual grammar and aesthetics. Any model of multimodality involved in describing the use of comic books will have to separate the combination of text and image from the interplay of heautonomous and grammatical semiotic modes, finding both kinds of modes realized in the graphical elements of comics' pages (and likely also, independently, throughout the scriptoral elements and in their combinations).

Thirdly, this interpretation does not allow for a parallel reduction of grammar and aesthetics to two independent parts of reading practices – and hence, it also does not allow for a separation of analytical interest into a linguistic explanation of visual grammar on the one hand, and an aesthetic appraisal of further interpretability on the other. Instead, the elements constructed in the course of this hermeneutic semiosis make sense only in the interplay between reflexive grammatical and reflective heautonomous subsumption – the latter relating to the rhetorical problem of free hypotyposis as much as to issues of regular grammatical comprehension.

Finally, the gap described between comics grammar and aesthetics is not a distinction between rule-bound and irregular structures. While some irregular structures may have aesthetic effects, the central concept of aesthetics intended here, as in Groensteen's argument, is one of over-determination rather than indetermination, of an abundance of rules as opposed to an absence of rules. Accordingly, Jakobson has described this approach for the formalist and structuralist traditions when he referred to the 'poetical' as tantamount to a dominance of the 'grammatical' function (2010: 25). As the necessities of grammar demand relationships between elements of language uttered in linear sequence, and poetry creates additional, similar, but sometimes differently structured and always grammatically redundant similarities, the aesthetics of comics sequences and panel structures offer additional dependencies between their elements: in a structuralist sense, rendering the communication more grammatical than grammar demands. It is because of these additional interdependencies between elements that the hermeneutic look backwards from one panel towards a pre-

vious one is such a telling and typical movement for an aesthetic comprehension of comic book panels. It is after the second appearance of the trunk, the steel beam, and the glider that their first appearances are subjected to additional interpretative rules: Were they background objects or introduced as relevant props, was their placement in line with the later collision, did their presentation hint at, allow for, or exclude the way that the next panel continues their depiction?

In parallel, heautonomy, here, goes beyond any mere autonomy of art that would understand aesthetic objects as unregulated. It is intended not only to codify the fact that art is not entirely bound by the rules of everyday communication, but to describe the way in which individual pieces of art introduce new rules, sometimes by repetition within their own confines, but sometimes also by a constructed evidentiality in singular usage. Historically, the concept is equally connected to 18th and 19th centuries' Idealism's engagement with seeming intractabilities of aesthetics (cf. Homann 1999), in which it is the power of the genius, to which the sustained tension between having and not having rules is eventually ascribed. This occurs in its historical context, in a Romantic interpretation of a democratization of aesthetics for the bourgeoisie (cf. Rancière 2008: 27). Some of the emphatic statements of irregularity and its interpretation in comics art seem to connect to those evaluations of the worth of the art involved: If the comprehension of comics is down to an ultimately inexplicably irregular implementation of personal insight or taste, then the practice of that engagement becomes valuable in the sense of art's value for the *Bildung* of an individual personality, i.e. the free development of an individual mind and personality in accordance with the same Romantic conceptions. This conforms with the connection often drawn between this appraisal of comics and their use in didactics and pedagogics (e.g. Grünewald 2014).

The historical and political context of the semiotic tension can illuminate its structure and its historical evaluation, but does not deny its current potency. Perhaps we would view the relationship between grammar and aesthetics differently, if it were not for this place in history in which comics currently exist. But it is in this place that we find our work situated, and the analysis of the tension between grammar and aesthetics challenges a multimodal theory of comics formally as well as historically. One way to face this challenge, then, is to affirm both the existence of one or more grammars for comics and the importance of heautonomous aspects in individual comic books, and to posit a hermeneutic corridor as part of the reading practice in which the two are engaged in a circle of mutual revisions. In traditional terms, the distinction between the hermeneutic and the merely grammatical mode of interpretation would be understood as subjectivity or reflection, qualities of the Romantic concept of the consciousness. Notoriously difficult to prove alongside individual feats of cognition in empirical studies, or to describe alongside strict formal descriptions of grammatical elements in multimodal linguistic analysis, that concept might be less well suited to contemporary interdisciplinarity. If it is true that its importance in this context

stems from attempts to connect such conscious cognition to class rather than to ubiquitous individual abilities, we might have additional reason to deny it further emphasis. But semiotics allows for an alternative.

4. A semiotic approach: Reconsidering grammar, heautonomy, and abduction in a framework of multimodality

A complementary approach to face the same challenge moves beyond the historical language of its Romantic foundations by engaging with the continued work on a generalized theory of semiotics based in the ideas of Charles Sanders Peirce. As Bateman has recently argued (2018), Peircean semiotics are intrinsically multimodal. I believe that they excel through the precision with which they deal with multimodal differences and combinations. Not only is Peirce's semiotics generalized to the point where it can apply to pictures as easily as to script, but it allows for a description of each semiotic mode's differences and distinctions. Crucially, it allows for a definition of semiotic modes by which they are "generally far 'smaller', or more 'fine-grained', than those assumed in many other positions in multimodality studies, [while] [a]t the same time […being] also 'broader' in the sense that they are not restricted to lie within single *perceptual* modes" (Bateman 2018: 19).

This aptly frames the distinctions proposed in the previous sections: Assuming one or more grammatically structured, as well as one or more heautonomous semiotic modes, we will find that these modes are more finely grained, on the one hand, than the monolithic dichotomy between script and pictures, with grammatical and heautonomous elements appearing both in script and in pictorial elements without governing any of them completely. At the same time, they are broader in a different sense, as they do not depend upon the distinction drawn by modes of perception: whether the immediate sensual affordance of a material artifact entices us to view a picture or to read writing, we will still find ourselves with grammatical as well as heautonomous elements in either mode.

The distinction nicely mirrors the debate on how to define codes as well as the various levels of modes within the description of any modality (cf. the summary in Wildfeuer et al. 2020, esp.: 135–143). It is founded on the now well-established distinction between material and semiotic properties, by which a comic book printed on paper is materially different from its scan appearing on a computer screen, but both may offer many of the same important material affordances for semiotic use. In the examples that interest us, each comic offers material for the same perceptual modes, i.e. the graphic and the literal (and perhaps more). In addition, we might distinguish different semiotic modalities, which are not limited to, but include different explicit or implicit codes and their conventional rules. Simply, if a comic includes script in English as well as in German, we might say that the material affordance and the perceptual modes as well as the semiotic modes

are identical, but the conventional code employed respectively differs. Less simply, the same logic might apply to the graphic elements of a comic book. If Cohn is correct in distinguishing a North American from a Japanese visual language in comics and manga (2010 and 2016), one might imagine a combination of several such modalities within the same sequence of panels. This might be the case, for instance, when an American super-hero book employs an 'exploded' style to depict a sequence of movements and poses in great detail, integrating a device popularized by manga into an otherwise differently styled artifact.

In this sense, the claim that comics have grammar would refer to the two (or perhaps even more) grammars in action. This clearly separates the specific grammar in use from the hermeneutic distinctions made in the examples above, as it is not a difference between two different visual languages that separates the depiction of a trunk or steel beam as a spatial or as an object marker, but rather different subsumptions under the categories involved in any one such language. The fact that spatial and object markers are involved, and mutually distinguishable at all, might be a candidate for a convergence between different comics grammars, or point towards some universals of visual language.

One important consequence of this concept of grammaticality is that it easily and quickly leads any analysis into the distinction of different elements within panels, as opposed to the description of panel sequentiality alone. The scandal for classical philosophies of the image is the ease with which this introduces the otherwise controversial idea of an intra-pictorial syntax. As Sachs-Hombach (2013: 125) has argued in dialogue with the more general semiotic vocabulary of Goodman, images might well be considered syntactically 'dense' in a sense that invites the ascription of several different functions to one image, but discourages the disassembly and precise delimitation of several different functional elements within a panel. It is worth noting that Sachs-Hombach immediately connects this presumed quality of pictures to the fact that they do not engage in specific syntactical combinations in context, reserving the option for more complex combinatorics and complexes of graphical elements beyond single images, and its assorted analogues of lingual grammaticality (2013: 140). But it is the existence of exactly such syntactical relations between separate images within the panel sequence as well as more far-reaching, and still more complex, inter-panel relationships, that drive the top-down disassembly of panels in Krafft's text-linguistic model of comics comprehension and lead him to identify some bottom-up qualities typical for elements with specific functions in a next step. It is precisely by pulling apart the seemingly dense interrelations between elements w i t h i n each panel that he grounds his explanation of the interaction b e t w e e n panels.

This might prove important, as empirical work to demonstrate the existence of a grammar in comics has hitherto been more successful for page or sequence structures between panels than for the internal structures of panels (cf. Cohn et al. 2012, Cohn and Campbell 2014; Bateman et al. 2018).

This suggests that more detailed work on grammaticality within panels and across panel borders is needed in addition to the existing insights into panel inter-relations, and that the hermeneutic gap which invites conscious reconsideration of previous panels' elements might be co-emergent with aesthetic, as well as grammatical codes that dissect panels into such functional elements in the first place.

Perhaps the most striking proof so far for the existence of grammaticality between whole panels, as separate from logical semantic inferences between their contents, lies in readers' ability to distinguish between re-arrangements of panels, e.g. from *Peanuts* strips (Cohn et al. 2012: 6), which are altogether scrambled, from those re-arrangements in which each panel has been taken from a different strip, thus destroying all semantic cohesion, but in which the relative placement of the panels is maintained. This in turn sustains a syntactical cohesion that need not make sense, but constitutes the panel equivalent of those 'colorless green ideas sleeping furiously' – a sentence that is famously cohesive grammatically while logically incoherent (famously employed in and since Chomsky 1956: 116). The example also happens to showcase one stark difference between lingual and visual languages: While readers show the passive ability to sanction violations of grammaticality in the strips they are shown, they are not necessarily able to produce the obvious correct alternative. *Colorlessly greenly ideas sleeps* is easily corrected to *Colorless green ideas sleep*, but the corrected version of a scrambled panel sequence is less easily determined in theory and much less easily drawn by most readers in practice. Similarly, once corrected, the sentence *Colorless green ideas sleep* might be translated into German or French with a high albeit not absolute expectation for convergence among translations, whereas an attempt at translating a structurally correctly formed *Peanuts* strip from North American into Japanese visual language would present a much less convergent result, if any. Certainly one may come up with parallel constructions to a Schultz strip while employing Japanese visual language, but there are far more such constructions and there is far less determination as to the one obvious solution than in comparably simple lingual translation tasks, owing only in part to the density of graphical realization beyond the grammatically necessitated aspects, and the strength of iconic semantic references.[1]

If this points to a relative lack of productive determinedness for the grammar of comics as opposed to typical grammars in spoken and written language, this might still give further weight to the importance of the hermeneutical corridor, in which irritation is interpreted rather than corrected. This would also conform with empirical results that show that backtracking is more prevalent in comics than in written narrative, suggesting that revision upon irritation – the core hermeneutic practice – plays an especially large role (cf. Foulsham et al. 2016; Kirtley et al. 2023). In the two examples from Kane's art above, the vagueness of the relative arrangement of Goblin, Spider-Man, and the steel beam, or of Spider-Man's back and the Goblin's hovercraft, in the respective earlier panels is turned into an operative ambigu-

ity to motivate the ensuing shift in panel domains, and from there, the pivotal moments for the plot. An alternative would of course have been, to criticize and perhaps correct an infelicitous drawing that fails to clearly convey the relative positions of objects in space in each depicted scene. But the sustained abstract grammaticality of the panel sequence, quite independently from its representative function, allows, and the hermeneutic interpretability motivates, an acceptance and appraisal for the artistic choice.

All of this allows for a translation of the general attitude of hermeneutics into a specific semiotic aspect: Foregrounding the need for revision by backtracking, reviewing, and comparing material backwards as well as forwards constitutes a circular movement that introduces and revises candidates for heautonomous rules as well as hypothetical subsumptions of elements under conventional rules of grammars. Backtracking, reviewing, and comparing are of course elements entertained in most grammatical accounts of comprehension, and some of the most pertinent examples for the aesthetic devices under scrutiny here might also be described as 'garden-pathing', as 'discourse pops' (cf. Wildfeuer 2013), or in similar fashions. What is at issue here, is whether an account of the specifics of this movement in comics' aesthetics allows for a purely grammatical explanation that nevertheless remains compatible with the ascription of aesthetic qualities to the same structures by some readers. This requires attention to at least three distinct continua as proposed by Peircean semiotics: As a framework, a Peircean view must insist on a continuum of ongoing reasoning as part of any semiosis, as ideas about connections and meanings are progressively 'filled in' on the path from an immediate through a dynamic to a final interpretant. Then, more specifically, a second continuum must be assumed between granular and summary treatments of parts of semiosis, in which each element may be broken down into several smaller elements or integrated into greater wholes as needed, a balance which Peirce has referred to as the 'prescision' and 'abstraction' of ideas. Finally, both of these continua also draw our attention to modes of conclusion other than fully determined deductions, especially by induction and the continuum between simple and creative abduction.

I will return to the three continua in a moment. A necessary preliminary logical step to realize such a fine-grained approach to rules across all three, is to redifferentiate what it means for rules to be conventional. Heautonomous rules by definition are not previously established conventions. But even rules of grammar may be conventional only in some of several conceivable ways. We should at the very least distinguish conventional syntax for the shape and arrangement of signifying elements (regularly formed and arranged 'representamens', i.e. 'legisigns') from the most commonly discussed conventional connection between signs and their semantics (regular object relations, i.e. 'symbols'), and each again from the perhaps most important conventionally rule-bound integration of signs into greater wholes (regular 'interpretants', i.e. 'arguments' in Peirce' diction). Crucially, as Bateman (2018) also emphasizes, the trichotomy of representamens and

the trichotomy of object relations are not entirely independent from one another – and hence, symbolicity and conventionality are n o t entirely orthogonal.[2] That is to say, while iconic signifying relations that operate through some similarity between a representamen and their content may be grounded on qualisigns, which convey a pure sensational quality, equally well as on legisigns, which are formed in accordance with a syntactical rule of the code employed, the reverse is not true: A symbolic sign, which refers to its object by some conventional rule, can only do so if it is a legisign, i.e. formed such that it may be recognized as the signifying element that satisfies the rule in the first place. In other words, while symbolicity is conventionality for sign-object relations, it is 'legisignality' that constitutes accordance with rules on a purely sensual or aesthetic level; and symbolicity always entails legisigns, even if legisigns need not be used as symbols.

A purely aesthetic approach to comics, if taken to an extreme, might focus on the same graphical elements merely as 'sinsigns', i.e. singular tokens in a usage that does not consider their regular repetition.[3] Similarly, the integration of an icon or a symbol into a fully determinate interpretable unit might include that icon or symbol in what Peirce refers to as a 'rhema', a 'dicent,' or an 'argument', depending on whether a statement is made as hypothetical across possible cases, as indicative of a given case, or as accordant to a general rule for all pertinent cases. Once again, any symbol and any legisign can be integrated into any of these semiotic triads, from rhema through dicent to argument. But if the interpretation is to arrive at generalized statements or arguments, the signs involved have to be legisigns and symbols, not mere qualisigns or icons: Only when a sign is recognizably formed in accordance with a certain rule and then connected to a certain signified object in accordance with its rule of interpretation, can it express a proposition about that interpretation that is itself rule-bound in a generalizable fashion.

This can mean forming a legisign in accordance with an exact model, such as is the case with even a singular replica of a letter from an established alphabet, used to express what that letter should look like. In the cases of legisignality in the examples discussed so far, it is not the exact shape of a picture of a trunk or beam that is predetermined, but it is its relational usage across arrangements of panel elements that accords to one or more visual languages. The shifting functions ascribed to tree trunk, steel beam, and hovercraft, as each moves from a spatial to an object marker, all assume recognizability for these elements (which is why we can name the graphical units by their semantic lingual equivalents in the first place). To say that tree, beam, or craft are in the situative background in one panel but interact with the plot-driving characters in the next panel, is to first be able to r e c o g n i z e these elements across panels by their r e g u l a r shape and arrangement, i.e. because they function as legisigns, which allows for but does not demand their iconic relationships to the objects denoted.

The salience of this point might become even more clear when we consider the alternative. An alternatively constructed semiosis might connect the two instances of tree trunks, steel beams, or hovercrafts across two panels not as two tokens of the same type of legisigns, but merely as the icons which they doubtlessly also constitute. This would read the second occurrence of each element as an iconic sign for the first and/or an iconic sign of the same referent, i.e. associate the two depictions of the same thing in two panels by their similarity without reducing that similarity to an implicit rule. In other words, we know that the trunk in panel 2 is the trunk in panel 1 because they each look like a trunk and like each other. And, of course, that is exactly the case. But because each picture of a tree trunk is similar to the other depiction of the same trunk, and also to an actual tree trunk, this runs the risk of conflating the repetition of signs with their iconic motivation, i.e. confusing syntax with semantics. We a l s o recognize that the trunk repeats syntactically in that place and in a pertinently shifted function. Iconicity is a possible, but not the only involved relationship: In each specific case, an element such as a tree trunk will be partially recognizable across panels because each instance is iconically similar to the appearance of an actual tree trunk, and yet the certainty with which the element is recognized as the s a m e tree trunk will draw on similarities between each depiction, which are not merely iconic relationships, but qualities of legisigns by which the first instance imposes limitations to the variability of further instances.

The semiotic principle involved here echoes the distinction between arguing that some lexical elements of language are iconically motivated, and that lexicalization does not govern them: *Cuckoo* and *Kuckuck* sound the way they do for a reason, but each language has recognizably different conventionalizations for their sound. Each sound might be interpreted as iconic by a person unacquainted with English or German, but to recognize the German or English word is an additional feat, establishing that a certain legisign is used, which in turn may carry a conventional, i.e. symbolic meaning. In parallel, each depiction of a trunk is motivated by the similarity in its appearance, but comics' syntax identifies the subsequent elements in subsequent panels through additional conventions.

What is at stake here for semiotic theory becomes clear when we compare this approach to the far more limited view of semiotics that considers only sign-object relations, and hence only recognizes iconic references from depiction to referenced object in panel sequences, denying the possibility of a visual grammar built on legisigns as opposed to mere icons. Strong versions of this latter view appear in critiques of semiotics in several phenomenological accounts. The most prominent example in comics studies might be in Lambert Wiesing's repeated (e.g. 2008) emphasis on the aesthetics of the speech bubble as the seminal moment in the history of sequential graphic art. In this account, comics become their own art form when Outcault interrupts the densely iconic depiction of scenes with 'rifts' in the canvas that contain written words. Having chosen the limited framework for

pictorial elements' interrelations, Wiesing has no choice but to once more return to the combination of text and images as comics' defining aesthetic feature: his denial of internal semiotic structures and intra-pictorial syntax has taken regular relationships between pictorial elements out of the argument. On the contrary, a visual morphology that allows for but is not limited to the iconic mode for graphical elements will assume "open-class lexical items" which "easily allow for new patterns to be created," and which can explain why in "visual form, these are typically," but n o t necessarily,

> iconic representations: it is easy to create a novel schematic pattern for iconic elements, based on the way they look (Cohn 2018b: 3).

The first continuum involved in such a semiosis is then built upon, and can be analyzed as, a series of conclusions, of deductions as well as more fragile induction and abductions. For our first example, one possible d e d u c - t i o n would be to recognize the unfinished contour of the steel beam as typical of what Krafft considers a spatial sign, and in a second step come to the conclusion that this is a spatial sign by a b d u c t i o n : If the steel beam is a spatial element, its contours need not be finished; its contours are unfinished; that might be the result of it being a case of a spatial element. Note how the parallel operation could be described as a consecutive deduction if the rules were certain. Assuming, then, that unfinished contours are only allowed for spatial signs, and recognizing that the contours of the steel beam are unfinished, we would h a v e to conclude that the steel beam is a spatial sign. But the actual uncertainty of the assignment negotiates grammatical relations considered across panel domains. This marks the previously described difference between the vagueness of comics' rules as opposed to language: *Colourless green ideas sleeps* contains a definite incongruity between the plural *ideas* and the predicate *sleeps*. Comic books' panels propose more probable assignments alongside definitive ones.

This fits Cohn's and others' employment – not of a strict generative grammar in a Chomskian sense – but of a probabilistic construction grammar whose empirical foundation does not expect that each competent comics reader will arrive at completely identical conclusions in every case, and in which lexicon and grammar are not entirely separated in their functions (e.g. Cohn 2018a). Crucially again, comics readers not only may arrive at different hypotheses, but they know that, and in which regards, they might do so, and can thus allow for ambiguity as a fact about a given communication, rather than a problem to be dissolved. As the visual linguist engages in an induction across as many possible readings of as many possible instances as they can empirically analyze, this account assumes that in hermeneutic reconsideration, each individual reader will engage in a similar inductive movement, considering real or imagined cases of parallel arrangements of panels and panel elements, to double-check their preferred interpretation against each individual case in the mode of the grammatical interpretation in Schleiermacher's hermeneutic circle.

A Romantic interpretation would make this consideration the hallmark of an educated interaction with the material. However, the second continuum taken from Peirce's descriptions offers an account that need not be explicit about assumptions of conscious or unconscious processes in this feat of sensemaking. As described in his seminal paper *On a New List of Categories* (1868), detailing the continuum of semiotic inference allows and even necessitates the distinction of two movements which summarize all that the hermeneutic circle accomplishes, in our instance without having to introduce the idea of consciousness as their specifying quality. Rather than assuming any one proper segmentation of ideas into propositional units, from which only reflective reasoning can hypothetically deviate by reconsidering whether what has been understood subliminally has been understood properly, it is on the contrary the very "function of conceptions to reduce the manifold of sensuous impressions to unity" (Peirce 1868: 278), and that function can be fulfilled in different parallel ways and is a l w a y s open to reconsideration with or without conscious reflection.

The general reason for this is that the granularity of the reduction is shifting and reversible. To adapt Peirce's example, I might consider a black stove a unit and add to its idea its placement at the far end of the kitchen; or consider the unit a stove, to which its color is added by predication; or consider the black stove at the far end a unit, for which I may then predicate its owner, etc. The far reaches of this movement are congruent to the general philosophical accounts of substance and being in Peirce's view, but conceiving of their opposition as continuous allows for a clear understanding of the logical movements between them: Movements which Peirce terms 'prescision' when they occur in one direction, and 'abstraction' in the other. If we follow his terminology,

> the terms 'prescision' and 'abstraction,' which were formerly applied to every kind of separation, are now limited, not merely to mental separation, but to that which arises from *attention to* one element and *neglect of* the other (Peirce 1868: 289).

The attention, or lack thereof, is the echo of the mind at work in more essentialist accounts of cognition, and becomes one dimension of secondness or volition in the semiotic account. But that volition is at once removed from the limitations to genuine human consciousness or 'mental' operations of separation. Rather, it allows us to think of prescision and abstraction as movements in any progressive semiotic process. Whether the steel beam, trunk, or hovercraft is to be summarized among the elements of the spatial sign, or grouped as a circumscribed unity of its own among the objects interacting with agents, is then due to just such a mutual movement of possible prescision and abstraction. It is a process open to conscious reflection, but not dependent upon it; the intricately staged emergence of the steel beam from one triangular panel to the next requires backtracking and is open to its conscious performance, but does not require consciousness. Readers will, as a general rule, appreciate the beam's sud-

den importance; they might, but need not, appreciate how that transformation has been signaled.

The third continuum of conclusions – deduction, induction, and abduction – has already been touched upon to describe the gradual process of semiosis. But in addition to the continuous movement from one conclusion to another in logical sequence, they also entail a continuum of more or less certitude. Not only are induction and abduction uncertain syllogisms as opposed to the determinedness of full argument of deduction (cf. Peirce 1878), but their conjunction allows for a yet more detailed typology. If deduction combines an established rule with an established case of that rule to determine the necessary result, abduction proceeds from a result and an established and possibly pertinent rule to a hypothetical case; and induction from a case and a result, or – as is more usually assumed – many such combinations of cases and their results, to a possible rule. But later versions of this theory distinguish a creative from a simple abduction such that a creative abduction introduces a rule that is less firmly established, and might be hypothetical itself: As Umberto Eco argues, it is in these cases, that

> the law must be invented *ex novo*. To invent a law is not so difficult, provided our mind is 'creative' enough (Eco 1988: 207).

Eco, here, is primarily considering the creativity in scientific reasoning that introduces new categories, focusing on the example of those populating a Linnean taxonomy in biology. However, the combination of all three continua allows us to conceptualize some important aspects of the creativity involved in heautonomous comics comprehension: that revision may often be gradual and subject to revision as we move from panel to panel; it might be open to prescision and abstraction in regrouping, summarizing, and redifferentiating elements and their subsumptions under panel domains: and it might either be ascribed to an author or reader, or none of them, as the concept of this creativity emerges from rather than determines the interrelations between signs.

This leaves us with a firm conception and a greater degree of freedom in explaining the phenomena from which we took our departure. In order to describe that the connection between Spider-Man and the steel beam might be interpreted as a categorical shift and then further connected to the interpretation of the whole episode as an interrupted and eventually avoided tragedy of revenge, we need to engage with the way in which pictorial elements are subsumed in panel domains; we need n o t make assumptions about whether or not the revision between two panels is conscious, or whether its creativity lies with the artist or the reader; but we need to admit room for the creativity of the abductions involved in either case. It is the interrelation of signs described as such, that allows for the possibility of these creative abductions[4], and in order to understand that, we need to consider the pictorial elements involved not merely as icons of denoted objects, but as legisigns governed by visual languages' grammars.

5. A revisional attitude: Conclusions for multimodal panel segmentation

This leaves us with three major conclusions for any multimodal model of comics comprehension that can deal with formal cohesion as interrelated with, but not reducible to, logical narrative or aesthetic coherence.

First, the importance of r e v i s i o n as an active part in the comprehension of panel sequences has become obvious, and can be explained in specific detail. While revision is usually, perhaps always, possible in any kind of somewhat linear processes of comprehension along the sequence of consecutively arranged elements within a media artifact, the revision involved in the comprehension of comics' panel sequences allows for a more specific descriptions of its properties and functions, with five qualities that I want to emphasize: It is (1) expressible as abductions about the subsumptions of panel elements under panel domains. This assumes (2) that some regularity for the formation of panel domains and their appropriate tokens has to exist in the first place: Grammar has to ground these aesthetics. By the same measure, this introduces (3) a direction for the productivity as opposed to the receptivity of comics' visual language that is mirrored in such a backward movement: For comics' grammar, revision surpasses correction or production. At least in the phenomena discussed here, the ability to resolve irritation by backtracking takes prominence against abilities more readily prevalent in other grammars, especially the ability to correct malformed units or to produce likely continuations. Stylistically, (4) devices of ambiguity in previous panels allows for productive resolutions to irritation in subsequent panels, and that ambiguity plays out in an oscillation between prescision and abstraction as complementary movements associating and separating elements. In terms of (5) a formal aesthetics, this foregrounds the importance of the contour, realized or implied, as an orienting device in the segmentation of the intra-pictorial syntax of comic panels.

Secondly, in terms of a politics of aesthetics, this allows for, but does not necessitate a connection to the Romantic interpretation of c o n s c i o u s h e r m e n e u t i c s as a hallmark of taste and education. To the extent that comics and their accompanying meta-discourses make that connection, they remind us of a historical context in which this aspect was foregrounded beyond its abstract semiotic necessity. We might consider the gap between aesthetics and grammar a hermeneutical corridor, if we want to frame it in the concepts of Romantic theory, and doing so tells us something about comic books and the history of this art form. But we might as easily specify a revisional attitude integral to the semiotics of comics, which is less of a conundrum about the tension between grammar and aesthetics, and more a unified interplay of abstraction and prescision.

In other words, when Spider-Man's head connects with that steel beam, the rearrangement of elements makes sense not merely because the imagined fictional world allows for such accidents, but mostly because the ambiguity and invited revision of the pictorial representation motivates such an

interpretation despite its tension with the previous depiction. We might value such revision and interpret it as a deeper appraisal of aesthetic or narrative consequence (cf. Grünewald 2014). A discourse that pursues this angle will create metaphors of immediacy and reflection, or automation and liberal thought. It will assume that we are given each picture, but that the closure between them is not given, and constructing it promises great cognitive challenges and feats; or that we are given meanings of words, but have to create meanings for images. It is certainly possible to fill the gradual directedness of hermeneutic time with such an imagery of gradual gain. But at the same time, this gradual process relates to any other kind of continued semiosis, and the continuum of subsequent semioses is not specific to these cognitive actions. To foreground some of the cognition involved as conscious, reflective, or sophisticated to distinguish it from others, is to echo a likewise possible but unnecessary social interpretation distinguishing accomplished from other readers by the quality of their thought, as opposed to their trained literacy in the art form. In educational discourses especially, this focus seems to mirror and narrow down a more general tendency to elevate the objects of our study by emphasizing the hardship of their interpretation. But revisional attitudes are not hard; after all, they have their own grammars. They can be learned.

One of the things we learn are the limits of determined deduction as opposed to hypothetical, creative abduction – a better way of distinguishing what is given, to what remains open for discussion. Which is why thirdly, it might be that the historical context of Romantic interpretation and its role in distinguishing an educated social class from other media users is a better explanation than any abstract semiotic theory for the c o n f l a t i o n and even confusion that has been going on in parallel accounts of these devices. They far too readily identify the revisional reading attitude towards comics panels' sequential combination with other dichotomies: If one is the distinction of c o n s c i o u s   o r   r e f l e c t i v e   r e a d i n g from automatic comprehension, another is the combination of pictures and script within and across them. The latter is no more identical to the revisional attitude than the former. The process of revising assignments of graphical elements to panel domains by backtracking revision hardly seems congruent with that other semiotic interrelation typical of but not specific to comics which connects images to the written word. The combination of semiotic modes that happens between writing and pictures in most comic books is of an entirely different nature than the specific type of a semiotic mode that is characterized by the conditions for the revisional attitude. When Frahm, for instance, insists that the heterogeneity between script and pictures in comics underlines their self-referentiality and thus undermines a metaphysical belief in semiotic relations (2010: 146), he at once argues for a well-deserved attention to the parodistic elements of a fundamentally caricatural art form – but distracts from the very pictorial elements and interrelations that define it, to instead discuss their opposition to writing. Ironically, it is this hastily drawn connection to the distinction between pictures and language in comics that may obscure the grammatical features of the pictorial dimensions themselves.

This paper took its point of departure from the tension between two accounts of comics comprehension: One that emphasizes regularities of a grammatical nature, another that foregrounds aesthetic judgments. I hope to have shown that in contrast to many such accounts in recent debates, these views can be integrated. One reward of the effort to do so might lie in the specifications of some uses of backtracking enumerated above.

A second reward might be constituted by a better understanding of the historicity both of those two research positions, and the qualities of the art form that give rise to them. Here, different theories of how we understand comics point to different kinds of social imaginations, expressible by different assumptions about how a theory of comic comprehension should deal with individual differences in comics interpretation. What if you read those panels from *Spider-Man* differently to me?

Some of your interpretation might be separable from the specific functions of detailed panel elements. In my reading, Spider-Man's tragedy of revenge is resolved without saving or condemning the protagonist by the sudden elation of the Goblin's glider from a spatial marker to an interacting object or perhaps even an autonomous agent. A traditional view of the tension between pictures and script will look for a resolution to a possible ambiguity in the pictures in words: those spoken in the comic itself or those offered by readers. In the case at hand, the words spoken by the characters certainly mirror this interpretation of the plot. But that interrelation comes later than the immediate correspondence of lingual and pictorial markers in each panel, and much later than the interplay of the graphical elements, which exploit the productivity of the panel spaces involved, to motivate a sudden resolution that is congruent with the iconically represented world, but does not take its cohesion from that congruence. This might build up a parody of justice, even as the cartoonish representation of the actors implies a parody of iconicity. But it does not force us to identify an aesthetic tension between image and language with a political tension between parody and symbolic order, or of unconnected aesthetic images with grammatically well-ordered panel sequences and compositions. To do so repeats a social imagination of the artform grounded in the hierarchies of previous centuries and their continued quest for the genius interpreter that recognizes, reflects, and resolves or retains these tensions.

This leaves a broad scope for differences in resolving abductions even for the grammatical interrelations between purely pictorial elements on the page. As we know that creative abduction remains uncertain, we cannot presume that other readers will agree with ours. One aspect of the Romantic tradition, repeated in the insistence on a disruptive aesthetics for comics, is its translation of this fact into an unboundedness proper to aesthetics: We might only hope that others agree with our aesthetic judgments, a hope summarized by the *Ansinnen* in Kantian aesthetic theory, as despite evidence to the contrary, we turn to one another with an expectation or even demand, that they might agree with what we have read. But the Peircean concept of abduction does not force resolution; instead, it allows for accu-

mulation. We are equally free to take the openness of that interpretation as an invitation to turn to one another and ask: Here, in this panel, did you see that too?

Notes

1   This fact seemed obvious to me, but the peer review has shown that others do not necessarily think it so. So we need more empirical evidence: Can we elicit convergent solutions from persons fluent both in North American and Japanese visual language for translating a sequence of panels from one to the other? I am not aware of previous work that answers this exact question.
2   I am indebted to Neil Cohn for clearly pinpointing this claim in the discussion of his presentation at the 2021 annual *ComFor* conference, even though I am here taking the opposite stance to his on this detail.
3   I am grateful to the peer review for this important point.
4   In general, the Peircean account of the possibility of a certain semiosis might leave questions about the choice of one of these possibilities, its interestingness, salience, or fit, unanswered. In other words, it explains why what happens is possible, but not why this possibility among many is realized. Cf. Packard (2006) for an attempt to combine these hypotheses with a concept of semiotic desire that motivates such choices while remaining equally non-committal about the ascription of these states to individual minds.

Bibliography

Baltzer, Ulrich (1994). Selbstbewußtsein ist ein Epiphänomen des Zeichenprozesses: Die Landkartenparabel von Ch. S. Peirce. *Zeitschrift für Semiotik* 16, 357–372.

Bateman, John A. (2018). Peircean semiotics and multimodality: Towards a new synthesis. *Multimodal Communication* 7, 1, 1–24.

Bateman, John A., Francisco O. D. Veloso, Janina Wildfeuer, Felix HiuLaam Cheung and Nancy Songdan Guo (2016). An open multilevel classification scheme for the visual layout of comics and graphic novels: motivation and design. *Digital Scholarship in the Humanities* 32, 2, 476–510.

Bateman, John A., Annika Beckmann and Rocío Inés Varela (2018). From empirical studies to visual narrative organization. Exploring page composition. In: Alexander Dunst, Jochen Laubrock and Janina Wildfeuer (eds.). *Empirical Comics Research*. New York: Routledge, 127–153.

Chomsky, Noam (1956). Three models for the description of language. *IRE Transactions on Information Theory* 2, 3, 113–124.

Cohn, Neil (2010). Japanese visual language: The structure of manga. In: Toni Johnson-Woods (ed.). *Manga – An Anthology of Global and Cultural Perspectives*. New York: Continuum, 187–204.

Cohn, Neil (2013). *The Visual Language of Comics*. New York: Bloomsbury.

Cohn, Neil (2016). The vocabulary of manga: Visual morphology in dialects of Japanese Visual Langue. *Journal of Pragmatics* 92, 17–29.

Cohn, Neil (2018a). In defense of a 'grammar' in the visual language of comics. *Journal of Pragmatics* 127, 1–19.

Cohn, Neil (2018b). Combinatorial morphology in visual languages. In: Geert Booij (ed.). *The Construction of Words*. London: Springer, 175–199.

Cohn, Neil, Martin Paczynski, Ray Jackendoff, Phillip J Holcomb and Gina R Kuperberg (2012). (Pea)nuts and bolts of visual language: Structure and meaning in sequential image composition. *Cognitive Psychology* 65, 1–38.

Cohn, Neil and Hannah Campbell (2014). Navigating comics II: Constraints on the reading order of comic page layouts. *Applied Cognitive Psychology* 29, 2, 193–199.

Cohn, Neil and Joost Schilperoord (2022). Remarks on multimodality: Grammatical interactions in the parallel architecture. *Frontiers in Artificial Intelligence* 4 (no page indication).

Colapietro, Vincent (2014). Notizen zu einer Skizze einer peirceanischen Theorie des Unbewussten. *Journal für Psychoanalyse* 55, 119–140.

Eco, Umberto (1988). Horns, Hooves, and Insteps. Some hypotheses on three types of abduction. In: Umberto Eco and Thomas A. Sebeok (eds.). *The sign of three – Dupin, Holmes and Peirce*. Bloomington: Indian University Press, 198–220.

Foulsham, Tom, Dean Wybrow and Neil Cohn (2016). Reading without words: Eye movements in the comprehension of comic strips. *Applied Cognitive Psychology* 30, 4, 566–579.

Frahm, Ole (2010). *Die Sprache des Comics*. Hamburg: Philo.

Grennan, Simon (2017). *A Theory of Narrative Drawing*. New York: Palgrave Macmillan.

Groensteen, Theirry (1999). *Système de la bande dessinée*. Paris: Presses Universitaires de France.

Grünewald, Dietrich (2014). Das Prinzip Bildgeschichte: Comics als eigenständige Kunstform und die Notwendigkeit, den sachgemäßen Umgang mit ihr zu erlernen. In: Barbara Lutz-Sterzenbach, Maria Peters and Frank Schulz (eds.). *Bild und Bildung. Praxis, Reflexion, Wissen im Kontext von Kunst und Medien*. München: kopaed, 453–80.

Halliday, Michael A. K. and Ruqaiya Hasan (1976). *Cohesion in English*. London: Longman.

Homann, Renate (1999). *Theorie der Lyrik. Heautonome Autopoiesis als Paradigma der Moderne*. Frankfurt a.M.: Suhrkamp.

Jackendoff, Ray and Jenny Audring (2020). *The Texture of the Lexicon. Relational Morphology and the Parallel Architecture*. Oxford: University Press.

Jakobson, Roman (2010). Linguistics and poetics. In: *Selected Writings III: Poetry of Grammar and Grammar of Poetry*. Berlin: De Gruyter, 18–51.

Kirtley, Clare, Christopher Murray, Phillip B. Vaughan and Benjamin W. Tatler (2023). Navigating the narrative: An eye-tracking study of readers' strategies when reading comic page layouts. *Applied Cognitive Psychology* 37, 1, 52–70.

Krafft, Ullrich (1978). Comics lesen. Untersuchungen zur Textualität von Comics. Stuttgart: Klett-Cotta.

McCloud, Scott (1994). *Understanding Comics. The Invisible Art*. New York: Harper Collins.

Menzel, Katrin, Ekaterina Lapshinova-Koltunski and Kerstin Kunz (2017). Cohesion and coherence in multilingual contexts. In: Katrin Menzel, Ekaterina Lapshinova-Koltunski and Kerstin Kunz (eds.). *New Perspectives on Cohesion and Coherence*. Berlin: Language Science Press, 1–11.

Packard, Stephan (2006). *Anatomie des Comics. Psychosemiotische Medienanalyse*. Göttingen: Wallstein.

Peirce, Charles Sanders (1868). On a new list of categories. *Proceedings of the American Academy of Arts and Sciences* 7, 287–298.

Peirce, Charles Sanders (1878). Deduction, induction, and hypothesis. *Popular Science Monthly* 12, 470–482.

Rancière, Jacques (2008). *Ist Kunst widerständig?* Edited and translated by Frank Ruda and Jan Völker. Berlin: Merve.

Sachs-Hombach, Klaus (2013). *Das Bild als kommunikatives Medium: Elemente einer allgemeinen Bildwissenschaft*. 3rd ed. Köln: Herbert von Halem.

Schleiermacher, Friedrich (1809/10). Allgemeine Hermeneutik von 1809/10. Edited by Wolfgang Virmond. In: Kurt-Victor Selge (ed.). *Internationaler Schleiermacher-Kongreß 1984*. Berlin: De Gruyter, 1269–1310.

Stainbrook, Eric (2016). A little cohesion between friends; Or, we're just exploring our textuality: Reconciling cohesion in written language and visual language. In: Neil Cohn (ed.). *The Visual Narrative Reader*. London and New York: Bloomsbury, 129–157.

Tseng, Chiao-I and John A. Bateman (2018). Cohesion in comics and graphic novels: An empirical comparative approach to transmedia adaptation in *City of Glass. Adaptation* 11, 2, 122–143.

Wiesing, Lambert (2008). Die Sprechblase: Reale Schrift im Bild. In: Alexandra Kleihues (ed.). *Realitätseffekte. Ästhetische Repräsentation des Alltäglichen im 20. Jahrhundert*. München: Fink, 25–46.

Wildfeuer, Janina (2013). Formale Zugänge zur Diskursanalyse. *Zeitschrift für Semiotik* 35, 3–4, 393–417.

Wildfeuer, Janina, John A. Bateman and Tuomo Hiippala (2020). *Multimodalität. Grundlagen, Forschung und Analyse. Eine problemorientierte Einführung*. Berlin: De Gruyter.

Sources

Conway, Gerry (wri) and Gil Kane (art) (1973). *Amazing Spider-Man 122: The Green Goblin's Last Stand!* New York: Marvel.

Image Sources

Fig. 1. Conway (wri) and Kane (art) (1973: 4).
Fig. 2. Krafft (1978: 48).
Fig. 3. Conway (wri) and Kane (art) (1973: 20).

Prof. Dr. Stephan Packard
Universität zu Köln
Institut für Medienkultur und Theater
Meister-Ekkehart-Str. 11
50937 Köln
Germany
E-Mail: packard@uni-koeln.de

The Logics and Semiotics of Discourse Relations in Comics

Janina Wildfeuer, University of Groningen

Summary. This article aims to point out some of the specificities of an approach that understands discourse relations in comics within a broader conceptualization of multimodal discourse. It lays out the theoretical foundations of these relations in philosophy, semiotics, and logic, and showcases how a small and concise set of these discourse relations can explain processes of reasoning in the interpretation of a comic.

Keywords. Discourse relations, logics, semiotics, multimodal discourse, reasoning

Zusammenfassung. Dieser Beitrag diskutiert Besonderheiten des analytischen Zugangs zu Diskursrelationen in Comic im Kontext multimodaler Diskursanalyse. Er legt dafür zunächst die theoretischen Grundlagen dieser Diskursrelationen in Philosophie, Semiotik und Logik dar und veranschaulicht dann, wie ein relativ kleines, gebündeltes Set von Diskursrelationen es ermöglicht, unterschiedlichste Interpretationen und Schlussfolgerungen in der Interpretation von Comics nachzuvollziehen.

Schlüsselwörter. Diskursrelationen, Logik, Semiotik, multimodaler Diskurs, Schlussfolgerung

1. Introduction

In my previous work on comics and graphic novels together with John Bateman (Bateman and Wildfeuer 2014a, b), we have characterized sequential visual narratives as a form of multimodal discourse which is interpreted dynamically by inferring d i s c o u r s e   r e l a t i o n s between segments. From a linguistic and multimodal analytical point of view, these discourse relations are usually treated as semantic relationships that build coherence between different entities within a panel or across several panels and so account for the discursive (and in many cases narrative) structure of the comic strip or page.

Many different approaches in comics studies have theorized about similarly meaningful relations between units of image and text, including Scott McCloud's fundamental work from 1994 as well as several other, diverse perspectives (e.g. Cohn and Bender 2017; Harris 2007; Saraceni 2000; Varnum and Gibbons 2001; Wartenberg 2012). In fact, the aim to describe relationships between elements in a comic has been a "central theoretical concern of comics studies" (Spanjers 2021: 81), and Spanjers' recent overview has demonstrated once again the diversity of approaches that have been employed to address it. Whereas the author himself goes back to the much discussed 'Laokoon' by Gotthold Ephraim Lessing (1990 [1766]) as a central starting point and draws connections to Roland Barthes as well as W.J.T. Mitchell, particularly d i s c o u r s e -oriented approaches to this debate that put a focus on the textual and discursive characteristics of comics are rarely mentioned in this and other overviews.

For a multimodal analysis of comics, this is interesting, especially given that the immensely growing context of multimodality studies has provided many different frameworks for the analysis of visual artifacts as discourse, both from a functional as well as a formal perspective. These works often go beyond discourse theories into philosophy proper, not only taking into account advancements from discourse studies and discourse semantics, but also integrating some fundamental features and concepts from semiotics and logics (see also Bateman 2021; Wildfeuer 2021).

This article connects to that latter kind of comic theory. It aims to point out some of the specificities of an approach that understands comics discourse relations within a broader conceptualization of multimodal discourse. It lays out the theoretical foundations of these relations in philosophy, semiotics, and logic, and provides an integrated view of multimodal communication in general and in comics in particular. While the main part of our previous work has been focused on the notion of multimodal discourse (not least in contrast to the concept of a visual or multimodal grammar, see Bateman and Wildfeuer 2014a) and has approached the theory of this discourse from a linguistic point of view, this paper goes back to its origins explaining the connections drawn between logic, semiotics, and discourse theory. Beginning by highlighting the specific characteristics of a logical conceptualization of discourse relations and their semiotic treatment as a result of abductive inferences about the meaning of multimodal artifacts, I trace these ideas back to general philosophical theories of understanding and meaning in order to provide insights into how comics work similarly to verbal discourses – namely through basic principles of connections between thoughts.

An important aspect on which I aim to elaborate in this discussion is the fact that the approach of analyzing discourse relations in comics is not only applicable to purely 'narrative' types of this media genre, but equally allows for the examination and explanation of other communicative purposes to which comics are employed as discourses. This is particularly interesting and important with regard to the current trend of establishing comics in many different educational, persuasive, and instructional contexts. Bram-

lett et al. (2016) for example list eleven different, narrative and non-narrative comics genres, including autobiographical and journalistic comics; so-called 'metacomics' that play with their self-referentiality (cf. Cook 2012) or 'instruction comics' that give first-aid instructions in the form of a poster (cf. Wildfeuer et al. 2022) are just two more example genres to be mentioned here. A multimodal analysis of these different comics genres aims at a relatively neutral and thorough analysis, examining foundational and general properties of the communication form as such. Whether these genres fulfill more specific functions in the form of a narrative, an argument, or an instruction, for example, can then be examined more precisely with specific methods such as the identification of discourse relations. With the case study below, I aim to analyze an example from the educational context of health communication in order to demonstrate the breadth of the possible applications of such analyses of discourse relations.

The centerpiece of this paper will therefore consist of a small and concise set of discourse relations that are seen to hold between several elements and structures in comics (see section 2). The remaining parts of this paper will then provide a discussion of the basic principles underlying the conceptualization of discourse relations as logical and semiotic entities (see section 3) as well as a demonstration of the straightforward applicability of this set and the relations as a fruitful, hands-on method for the analysis of comics and graphic novels (see section 4). The conclusion in section 5, finally, will briefly discuss how this analysis contributes to comic theory as a whole.

2. Discourse relations in comics and graphic novels: a general overview

In his discussion of the general relationship between text and image(s) in many different communicative artifacts, Bateman (2014: 205–222) gives an overview of the main models of discourse and resulting classification schemes for image-text relations. He demonstrates applicability to diverse multimodal documents, including comics and graphic novels. Bateman also explains the general principles underlying this analytical approach:

> Regardless of framework, discourse is generally considered to be made up of 'discourse moves' of various kinds that serve to advance the communicative goals pursued by a speaker or writer. Since communicative goals are rather abstract, it is then natural to think of whether linguistic expressions are the only ways of achieving them. [...] Multimodal accounts of discourse consider whether particular proposals for discourse organisation can be extended with images taking on some of the roles of discourse moves [...] (Bateman 2014: 206).

The ability of images to take on roles of discourse moves covers many processes initiated by a discourse in the course of constructing meaning, including some that are not visibly or explicitly expressed in the artifact itself. This

is taken up in more explicit accounts of comics as multimodal discourse (see also Jacobs 2013; Feng and O'Halloran 2012; Forceville et al. 2014; Tseng and Bateman 2018). They develop the idea from verbal discourse analysis further and adjust it for all visual units in comics discourse. Our work provides an overview of how these discursive principles are at work on different levels of comics, and how they keep the comprehension and further processes of interpreting a comic going (Bateman and Wildfeuer 2014a, b). In the following, I want to illustrate these discursive principles through one short example.

Fig. 1. From *Mental Load* (Emma 2017: 106).

Figure 1 shows a passage from Emma's *Mental Load* comic (2017) in which several discourse relations can be identified between the various segments, constructing a coherent piece of communication. Discourse analysis can show where exactly these discourse relations can be inferred, for example between several verbal units or between verbal and visual units. For instance, the sentences in this extract are connected by so-called 'Narration' relations, since the events they express (going to the office, sitting down, getting a coffee, coming back) can be seen as happening in a temporal and spatial sequence. The final sentence, in contrast, can be seen as presenting some sort of E n h a n c e m e n t of the tasks that are mentioned in the previous sentence, and the discourse relation inferred as holding between the two sentences is then also usually called E n h a n c e m e n t or E l a b - o r a t i o n . The visual part of the extract, the image of the woman with her bag, can also be connected to some parts of the text: The image illustrates details of the events of 'getting to the office' and 'taking one's place,' and details presented in this manner are usually inferred to be a part of some-

thing else, so holding a part-whole relationship, which is usually called P a r t - o f. Quite similarly, the speech bubble to the right of the woman represents an utterance that is related to the event of sitting down and, possibly, greeting some colleagues, although the latters' presence is not expressed verbally. It thus adds even further details and may also be seen as an E n h a n c e m e n t / E l a b o r a t i o n. So we have already identified several discourse relations between the various segments that render the whole a coherent piece of communication. That whole can then also be related to preceding and following parts of the comic as a larger narrative.

This short and rather informal discourse analysis has shown how meaning can be constructed, or 'made', out of relations between various elements in comics. While several frameworks and approaches are available that define discourse relations in comics and visual artifacts more specifically, also including non-narrative discourses (see, e.g., Martinec and Salway 2005; Liu and O'Halloran 2009; Feng and O'Halloran 2012), the analysis here has described only a few of these and on the basis of some very general principles of understanding. For instance, we usually expect stories to be told in a temporal sequence and we therefore expect that things and events in the story connect in a sense-making, logical, and temporal manner. At the same time, we often look for causes and purposes or oppositions and contrasts holding between two different ideas, and we know that sometimes two things or ideas are juxtaposed in order to show their similarities. Even if specific frameworks label the relations differently[1], there are some general principles at work, certain 'logical relationships' that are usually used in communicative artifacts to connect two ideas.

Some very general descriptions of such relationships are often doubled in writing tutorials, where the connection between ideas plays an important role. Writers are then asked to use logical connectors to combine sentences in a text or to build the structure of their texts on these logical relations, even if they are not explicitly expressed. Connections between logical and discourse relations have been most systematically developed in text and discourse linguistics, and especially in formal or functional discourse analysis. On this basis, we have previously presented (Bateman and Wildfeuer 2014a) a set of formalized discourse relations specifically defined for the semiotic mode of comics and with the purpose of analyzing meaning-making processes between units in comics and other visual artifacts. Similar, but in most cases less formalized overviews of these relations have been provided in other works in the context of multimodality research.

In Table 1, a broad summary of the main discourse relations that have been identified in these accounts over the years is given. The list only features the most frequently used relations that are commonly defined as being identified within several types of static artifacts. It builds on the sets of relations provided in Bateman and Wildfeuer (2014a), an even smaller set given in Packard et al. (2019) as well as the basic set of relations that has been defined by Asher and Lascarides (2003: 146) for both narrative and expository texts. The latter, for example, justify this limited set of relations by vir-

tue of the relations' truth conditional effects. Their choice is therefore always motivated by semantic interpretation. The set also includes parts of the main relations identified by van Leeuwen (1991) for film and by Liu and O'Halloran (2009) for text-image combinations more generally. Most of the artifacts analyzed so far have been fictional stories, mainly with a narrative function, telling a clear story and not serving any other communicative purpose. However, earlier accounts of verbal discourse analysis have already highlighted the applicability of this set to other text types, for example in the broad framework by Asher and Lascarides (2003) that also includes the analysis of professional conversations and disputes.

Tab. 1. Discourse relations in comics and graphic novels.

RELATION	NECESSARY CONDITIONS
Enhancement/ Elaboration	specification of an event/state/object by another event/state/object
Property	indicated property of one event/state/object is applied to the other event/state/object
Part-of	part-whole relationship between two events/states/objects in a sequence
Contrast	structural similarity, but semantic dissimilarity between two events/states/objects
Parallel	Structural and semantic similarity between two events/states/objects
Narration	Spatio-temporal consequence and shared topic among two events/states/objects
Result	Spatio-temporal consequence between two events/states/objects; cause

The description in the right column provides information about the conditions that should apply to the specific context in which the relation is identified. In the example analysis above, some of these conditions are paraphrased in very similar ways. What is not described further in this table is the definition of the units themselves, here given as either 'events', 'states', or 'objects'. Depending on the specific focus of the analysis (for example the discursive structure of several pages in a comic book, the relation between panels on a single comic page, or the interplay of even smaller

elements within individual panels), these units are not fixed beforehand, but need to be identified dynamically and on the basis of further inferential work.[2]

Comparing the different approaches and their definitions, the description given here is intended to be as inclusive and as general as possible for various levels of discourse in comics and graphic novels – be they narrative, argumentative or instructional. The relations described represent highly general semantic principles and logical relationships that are commonly assumed in many contexts in our daily lives. For instance, it is common knowledge that several events, states or objects stand in a causal relationship and lead to certain effects (e.g. pushing someone usually leads to the other person falling or stumbling). When these events or actions are described or shown in a discourse, this cause-effect relationship is usually analyzed as a discourse relation labeled R e s u l t (or sometimes also E x p l a n a - t i o n), and a necessary condition for the inference of such a relation is that the discourse also includes a cause which explains the connection between the two events or objects. When things are put next to each other for comparison, the relationships inferred between these two things are usually showing a C o n t r a s t, a S i m i l a r i t y, or a P a r a l l e l. This is equally true for two segments in a discourse, be they verbal or visual.

The analysis of these relations in a discourse, thus, directly builds on common knowledge about how things relate to each other. For this, however, not all details of this knowledge and the underlying structures have to be explicit in the discourse. This becomes even clearer when looking at the multimodal character of comics and other discourses in which even more information than in most written formats is given only implicitly, and much of the basic semantic content has to be inferred from non-verbal entities. One example in the passage above is, that no explicit information is given to identify the visually depicted character, such that the drawing has to be recognized as representing the same character that has been shown before in the comic, and which is the speaker or narrator of a written text that uses the first-person pronoun 'I'. While general relationships between events in verbal discourse can often also be expressed by so-called logical connectors (for instance conjunctions such as 'because', 'although', etc.), (audio-)visual elements often do not include any direct, explicit indication of such a connection. The little tail or pointer usually represents a direct relation to the character speaking, but that the picture shows something that is also described in the captions has to be logically concluded. These conclusions and relationships can be made visible by multimodal discourse analytical tools, a fruitful endeavor to explain meaning-making processes.

3.   Connecting ideas: Origins of discourse relations

Interestingly, in most accounts of discourse relations, the specific types of knowledge and their underlying structures as shortly explained above are not systematized any further. Whereas some approaches identify 'logical

relations' (see below for details), they do not discuss their theoretical background in logics and the accompanying principles of reasoning. Similarly, the semiotic aspects of the processes of inferential reasoning about the relations are often not explained comprehensively, although the "'system-internal' (i.e., logico-philosophical) organizations of such processes" (Bateman 2017: 21) build an important foundation for a fruitful connection to cognitive approaches to comprehension. In the following, therefore, the basic foundations of approaches to discourse relations in logic and semiotics will be discussed in further detail.

3.1 Discourse relations as logical relations

In an overview and classification of the various linguistic accounts of discourse relations, which are also often called 'coherence relations', Bateman and Rondhuis (1997: 3) state that it is "commonly assumed that one essential part of comprehending and creating discourse is the recognition of intended relations". Similarly to many other discussions in the field of linguistics and discourse analysis, the concept of 'discourse' is no further explained nor is there any elaboration of what Bateman and Rondhuis see as a 'common assumption' about the comprehension of discourses. In most cases, the aspect of 'understanding' is given as a, if not the, strongest motivation, as it is directly mirrored in the description of discourse relations as "the cornerstone of comprehension" (Graesser et al. 2003: 82). However, the linguistic focus in most accounts is then often 'only' placed on the creation of coherence as a basic principle of texts and discourses.

In the realm of systemic-functional linguistics, for example, explicit connections between processes inherent to language and those of human thinking are generally rather scarce. One important exception is given in Halliday and Hasan's (1979) notable work on cohesion and, more particularly, on conjunction as a specific type of cohesive relation:

> There are certain elementary logical relations inherent in ordinary language; doubtless these derive ultimately from the categories of human experience, and they figure importantly on the sociolinguistic construction of reality, the process whereby a model of the universe is gradually built up over countless generations in the course of semiotic interaction. (They can be regarded as departures from the idealized norm represented by formal logic; but it is worth remembering that in the history of human thought the concepts of formal logic derive, however indirectly, from the logic of natural language.) These logical relations are embodied in linguistic structure, in the form of coordination, apposition, modification, etc. (Halliday and Hasan 1979: 320).

The authors, here, make explicit what is often seen as an obvious fact: discourse relations build on more general logical relations that again derive from processes of human experience. The short analysis in section 2 has broadly exemplified these general logical relationships. Their discussion

from a more systematic perspective, however, rarely receives a lot of attention. For instance, in Halliday and Hasan's work, neither are the 'categories of human experience' nor is the close connection to logics explained further – and the reference quoted above is in fact the only concrete mention of the notion of logic in the entire book.

In Rhetorical Structure Theory, the definition and recognition of relations between units "rests on functional and semantic judgments" and "illustrate[s] a diverse range of textual effects", such as "interpersonal or social effects, ideational or argumentation effects, and textual or presentation effects" (Mann and Thompson 1988: 250). The semantic judgments listed are nowhere explained. An important detail in this theoretical conceptualization of relations, however, is the requirement that the definitions apply "only if it is plausible to the analyst that the writer wanted to use the spanned portion of the text to achieve the Effect" (Mann and Thompson 1988: 258).

In contrast, a much more explicit connection between discourse relations and underlying logical principles is usually given in approaches in formal discourse analysis which aim at describing coherence from the broader perspective of a formal model of comprehension within a theory of communication. Logical concepts are then used to explain the activities involved in comprehension in terms of drawing inferences. Hobbs (1979, 1983), for example, in his important early works on 'coherence relations', states that

> the sense we have that a discourse is 'about' some entity or set of entities is frequently just the conscious trace of the deeper processes of coherence [which is] the mortar with which extended discourse is constructed (Hobbs 1979: 68–69).

These deeper processes of coherence are then usually described as principles of inferential reasoning and are thus also connected to processes of understanding or comprehension:

> Comprehension is not simply a matter of the Speaker depositing a proposition in the Listener's heard. It involves an active inference process, in which, among other things, the Listener must infer the specific from the general or the general from the specific, in order to zero in on the Speaker's full intended meaning. By choosing or ordering his utterances in a particular fashion, the Speaker can exercise some control over this inference process by supplying or modifying the appropriate framework for their interpretation (Hobbs 2004: 14).[3]

Hobbs himself and several other accounts in the context of formal discourse semantics give overviews of some of the most general inferential principles within the context of nonmonotonic logic, which is seen as driving commonsense knowledge:

> Virtually all commonsense knowledge beyond mathematics is uncertain or defeasible. Whatever general principles we have are usually only true most of the time or true with high probability or true unless we discover evidence to the contrary. It

is almost always possible that we may have to change what we believed to be the truth value of a statement upon gaining more information. Almost all commonsense knowledge should be tagged with 'insofar as I have been able to determine with my limited access to the facts and my limited resources for reasoning'. The logic of commonsense knowledge must be nonmonotonic (Hobbs 2004: 2–3).

One specific variety of this nonmonotonic logic is abductive reasoning, which stands in the center of Hobbs's argument – and has later been taken up by many others, both in verbal discourse analysis as well as in semiotic accounts for media analyses (cf. Moriarty 1996; Jappy 2013). Hobbs himself bases his 'Interpretation as Abduction' framework on early works such as Newton's *Principia* (1934 [1686]), Christian Wolff's (1963 [1728]) understanding of philosophical hypotheses, as well as Peirce's explicit introduction of abduction as the third principle of reasoning (see more below). With this, it becomes an important starting point for several other works on discourse relations in the context of formal discourse semantics and artificial intelligence (see also Hartung and Cimiano 2007).

Asher and Lascarides (2003: 98), for example, criticize Hobbs's work insofar as his "abductive account misses important generalizations about the organization of different knowledge sources and their interactions during interpretation". In their work, the authors focus on exactly these knowledge sources and distinguish them in terms of several logics that are at work in a meaning-making process:

> One of our main claims here is that discourse interpretation should result from several interacting but separate logics rather than via a single, 'all singing all dancing' logic. Each logic is designed to do a distinct specific task: e.g., there is a logic in which you construct a representation of what is said, another logic in which you evaluate (the consequences of) that representation, a logic in which you reason about lexical polysemy, a logic in which you reason about another person's cognitive state on the evidence of his utterance and the assumption that he's rational and cooperative, and so on (Asher and Lascarides 2003: 430).

Much more so than other works published in this context, Asher and Lascarides make it evident that different knowledge sources as well as contextual conditions are at play when meaning is made in a (verbal) discourse interpretation. While the authors do not identify how this world knowledge results in the rhetorical relations they provide in their account, they still give a general architecture of these various logics in which they not only include the broad aspect of 'world knowledge', but also the aspect of the representation of such knowledge by another logic, the 'logic of information content' (see Asher and Lascarides 2003: Chapter 4). This part of their framework, in which formal representations are used to systematically identify the semantics, i.e. the meaning of the discourse, has its origins in the beginnings of formal logic and the general aim of representing knowledge and making beliefs, desires, etc. computable, i.e. processable.

Hans Kamp and Uwe Reyle's Discourse Representation Theory (DRT), a direct precursor on which Asher and Lascarides build their theory (SDRT), explains the idea of approaching meaning from a formal logical perspective with the following general remarks:

> One of the central features of cognitively complex beings like ourselves is that they reason. They move, with greater or lesser confidence, from beliefs they hold, hypotheses they entertain, desires they harbour, and intentions they have, to new beliefs and new intentions, and in this way they arrive at new ways of seeing the world and are propelled into new ways of acting upon it (Kamp and Reyle 1993: 9).

The authors, here, describe reasoning as some form of movement from one idea to another, which is very similar to how Bateman describes discourses as being made up of discourse moves, i.e. similar connections between ideas and segments (see the quote in section 2 above). Kamp and Reyle insist that the general reasoning process can only be properly systematized when the ideas and premises are available in some sort of representation, or language:

> However, the processes of reasoning cannot be understood [...] unless we assume that both beliefs, desires, etc. which act as the premises of mental inferences and the conclusions that are drawn from them have some kind of formal, language-like, representation structure within which the particular inference drawn instantiates a general formal inferential pattern, defined in terms of the structural relations between premises and conclusions as they appear within that mode of representation (Kamp and Reyle 1993: 9).

With this, Kamp and Reyle not only explain the strong connection between semantics as a theory of meaning and logic as the science of inference, but they also conceptualize the basic processes of reasoning as connecting new and old ideas. This is an important aspect of explaining the processes that lie behind the idea of discourse relations.

Kamp and Reyle give a few more details of the origins of formal logic and its connection to semantics when they refer back to Aristotle's syllogistic logic, in which classes of argument patterns are explained. They briefly talk about the further development into predicate logic as introduced by Frege (1980) and the general logical and semiotic processes as introduced in Peirce's work – and end up summarizing:

> Inference and deduction are activities in which human beings engaged long before logical theory began and which they engage in irrespective of whether the theory of logic is known to them or not. Logical theory must explain the nature of this activity (and, where possible, but only via this explanation, provide canons which might help us to improve it) (Kamp and Reyle 1993: 21).

With this plea for the close connection between reasoning in general and language and, more particularly, logic as a way of explaining processes of

reasoning, they then develop a model for the structural interpretation of sentences and texts (see also Wildfeuer 2014). The latter is in fact the intrinsic aim of all discourse analytical accounts, as has become apparent in the overview so far: theories of discourse comprehension usually describe links between language utterances (or discourses) and the world, but they also, and more importantly so, say something about how these links and discourses can and should be understood and interpreted (see also Kamp and Reyle 1993: 13). As a consequence, discourse relations as they appear in these theories may be seen as indications of how the link between discourse and world, between something communicated and something thought, can be described (and understood).

3.2 Discourse relations as 'signs' of knowledge

In the preceding section, some of the philosophical and semiotic origins to which discourse analytical approaches often recur have already been mentioned. While the earliest systematic thoughts reach as far back as Aristotle's conceptualization of logic, Peirce's foundational ideas about logic and semiotics play a similarly important role. In this section, I will discuss these in greater detail.

Interestingly, Peirce's more general ideas for a theory of communication are rarely taken into consideration within works that try to compute discourse relations more formally. As indicated above, abduction as a reasoning process plays an important role in these theories, and is indeed traced back to Peirce's conceptualization as a third principle of inference. However, no closer connections between this conceptualization and more general ideas in logic and semiotics are usually made. Nevertheless, Peirce's general thinking about knowledge, reasoning and the association of ideas are of particular interest and importance for the discussion here.

For instance, Peirce's conceptualization of logic as a general "science of the forms of thought" is already manifest in his early writings from 1860 (Peirce 1865: MS 921), which, according to Emily Michael (1978), is separated from psychology by its metaphysical approach:

> The study of how we think is the domain of psychology, according to Peirce. Logic, on the other hand, does not use our psychological processes as its data; it is unconcerned with our processes of thinking. The subject matter of logic, we are told, is: (1) 'logical reactions' of conceptions and (2) 'the thoughts as they present themselves in their logical form' (Michael 1978: 177; quotes from MS 921, published in Peirce 1865).

This distinction between logic and psychology as two different branches dealing with thoughts, one conceptually and formally and the other cognitively, is an important foundation for the understanding of discourse relations as specific logical relations that describe meanings and ideas in com-

municative artifacts. It is also an important basis for the differentiation between several types of comic theory to which we will come back in the conclusion of this paper.

Michael (1978: 179) states that "the findings of a science of logic will apply to all thoughts, but need not start with a study of thought." And the author further elaborates that

> we make the assumption that 'meaning resides in words and other material representations though these representations be understood or not, and whether they be actually written or fashioned or not' (Michael 1978: 179; quotes from Peirce 1865: MS 726).

By describing this understanding of logic as descriptive (and not normative), Michael then builds the bridge to logic as a semiotic study of

> the laws of linguistic symbols in terms of their necessary relations, quite independently of their derivation from or application to the mind (Michael 1978: 181).

Building on Peirce's general conviction that thinking always happens in signs, she concludes:

> That is, the logical relations of symbols in valid arguments are also descriptive of thoughts when thoughts are viewed as symbols and considered in terms of their logical relations (their logical form) (Michael 1978: 182).

This is in line with Peirce's later consideration of logic as a general 'semeiotic' (see Bellucci 2014) that sees the fundamental relation between premise and conclusion, and the process of reasoning along with it, as signs.[4] The important aspect here is that 'sign' must be understood as what Peirce sees as an 'external sign', an expression or instantiation, which Bellucci (2014: 525), in reference to Peirce (1893–1913: 544), calls "a sign materialiter":

> Materialiter, a piece of reasoning may be studied as instantiated in someone's thinking or as expressed through external signs (words, symbols, diagrams, etc.) (Bellucci 2014: 525).

Although Bellucci highlights that to avoid conflating this with psychology and cognitive aspects of reasoning (see above), the study of reasoning should concentrate on external signs and not on other instantiations of thinking, this general conceptualization of reasoning as a sign is another crucial point for our understanding of discourse relations – and it is very similar to what has been pointed out in formal-logical approaches to discourse as the need for a representation of the processes of reasoning (see section 3.1).

From a semiotic point of view, processes of reasoning and the resulting logical relations are seen as descriptions or representations of thoughts. This can then also apply to discourse relations as a specific type of logical

relations (see section 2.1), namely those instantiations of thoughts or think-
ing that implicitly connect external signs with each other (in a discourse).
Due to Peirce's universal understanding of signs as not only verbal but also
other material types of signs, it is then also reasonable to talk about those
external signs and the instantiations of thoughts in multimodal and/or non-ver-
bal discourses.

   Although there is no concrete discussion of a similar notion of discourse
in Peirce's work, his concept of a 'universe of discourse' (Peirce and
Ladd-Franklin 1902) is closely related to the points made here. As Hugh
Joswick (1996: 99) explains, this 'universe of discourse' is a

> common stock of knowledge between the utterer and interpreter that makes com-
> munication possible. It is not something that can be adequately described: it can only
> be indicated as something familiar to both speaker and auditor (Joswick 1996: 99).

The author further quotes from Peirce's own words:

> The parties of semiosis must thoroughly understand that they are talking about 'objects
> of a collection with which both have some familiarity … A certain amount of truth
> about this universe is taken for granted between them. So far as they have the same
> idea of the universe, upon that universe the attention of both is fixed; and when makes
> any assertion to the other, and the other assents to it, what happens is simply that
> their common idea of the universe becomes more definite; for their whole discourse
> is about that and nothing else' (Peirce 1976: NEM 3: 407; Joswick 1996: 99).

In an application of these principles, discourse relations in the semiotic
understanding discussed above might be thought of as instantiations of
thoughts in a common discourse on the basis of shared knowledge between
the producer and the recipient. A discourse can therefore only be commu-
nicative and effective if both producer and recipient have a similar under-
standing of how the segments of the discourse relate to each other – and
this understanding is based on a shared knowledge. Or, as Hobbs puts it:

> [C]oherence relations are conventionalized ways of being reminded of things. They
> are those ways of traveling through our mental maps that we can reasonably expect
> a listener to follow (Hobbs 1983: 10).

Making these discourse relations explicit in a discourse analysis is then
also a process of making this shared knowledge explicit – though not com-
parable to the explicitness that cognitive experiments can bring about.
Although the concept of discourse relations is often used to explain the
meaning-making, inferential reasoning processes that recipients undertake
to understand the discourse, the resulting descriptions of this analysis do
not explain the brain processes that cognitive approaches aim at revealing.
Instead, a logical and semiotic analysis of the discourse relations provides
insight into the system-internal, i.e. the media specific, logico-philosophi-

cal organizations of such processes (see also Bateman 2017). How these system-internal organizations work on several different levels of comics and graphic novels, and how the set of discourse relations provided in section 2 can straightforwardly be implemented to show these organizations at work, will be demonstrated in the following section.

4. Discourse relations in use

In our previous work, we have already demonstrated the applicability of a small set of discourse relations to several levels of comics and graphic novels. We have shown that they can be identified to hold between relatively small units within a panel, such as, for example, between motion lines and the visual representation of a character (see Bateman and Wildfeuer 2014a: 189–194) or between so-called 'upfixes' and the head of a character (see Bateman and Wildfeuer 2014b: 381–383). We have demonstrated that they similarly hold between individual panels in a sequence, both within smaller comic strips as well as on larger pages from a comic or graphic novel (see also Wildfeuer 2014). For this, we have also discussed controversial examples from Cohn's approach to the analysis of narrative visual structures (e.g., Bateman and Wildfeuer 2014a: 196–199). Not least, we have shown that discourse relations have the potential to indicate spatial resources and layout strategies in large scale panel organizations (Bateman and Wildfeuer 2014a: 200–203; Bateman and Wildfeuer 2014b: 398; Bateman et al. 2017: Chapter 13; Wildfeuer 2021).

Almost all of the examples we have discussed so far are mainly from fictional genres and we have argued that an analysis of the discourse relations "helps to identify and describe the overall narrative structure of comics pages or more complex sequences of images" (Bateman and Wildfeuer 2014a: 196). With the following example, I want to argue that the basic set of relations provided is also applicable to other comic genres which are not primarily fictional or narrative, but particularly those that follow an educational and instructive purpose and for which the analysis of discourse relations helps to outline the particular teaching and instructive aspects of the comic.[5]

The analytical example in question is the so-called 'comic education module' examined in a study by Hanson et al. (2017) which tested the comic's fitness for educational use in a US hospital, where young patients complaining of pain were instructed on the application of pain medication at home. A survey of children and caregivers was conducted to find whether the comic influenced children's behavior at home. This education module consists of a comic spread featuring two pages with somewhat regular table grids, designed in black and white and with a relatively high proportion of speech bubbles spread over the various panels (see Fig. 2). According to Hanson et al. (2017), the predefined teaching goals included encouraging young patients to talk about their pain at home and to accept it as something normal after an injury, which can be treated with pain medication.

To support these teaching points, the authors created a simple story normalizing pain after injury, modeling a child asking for pain medication, and showing pain relief after medication. The teaching points are reinforced by two of the characters, and the comic includes a teach-back portion that prompts the reader to fill in a speech bubbles of a learning each of the three teaching points (Hanson et al. 2017: 530).
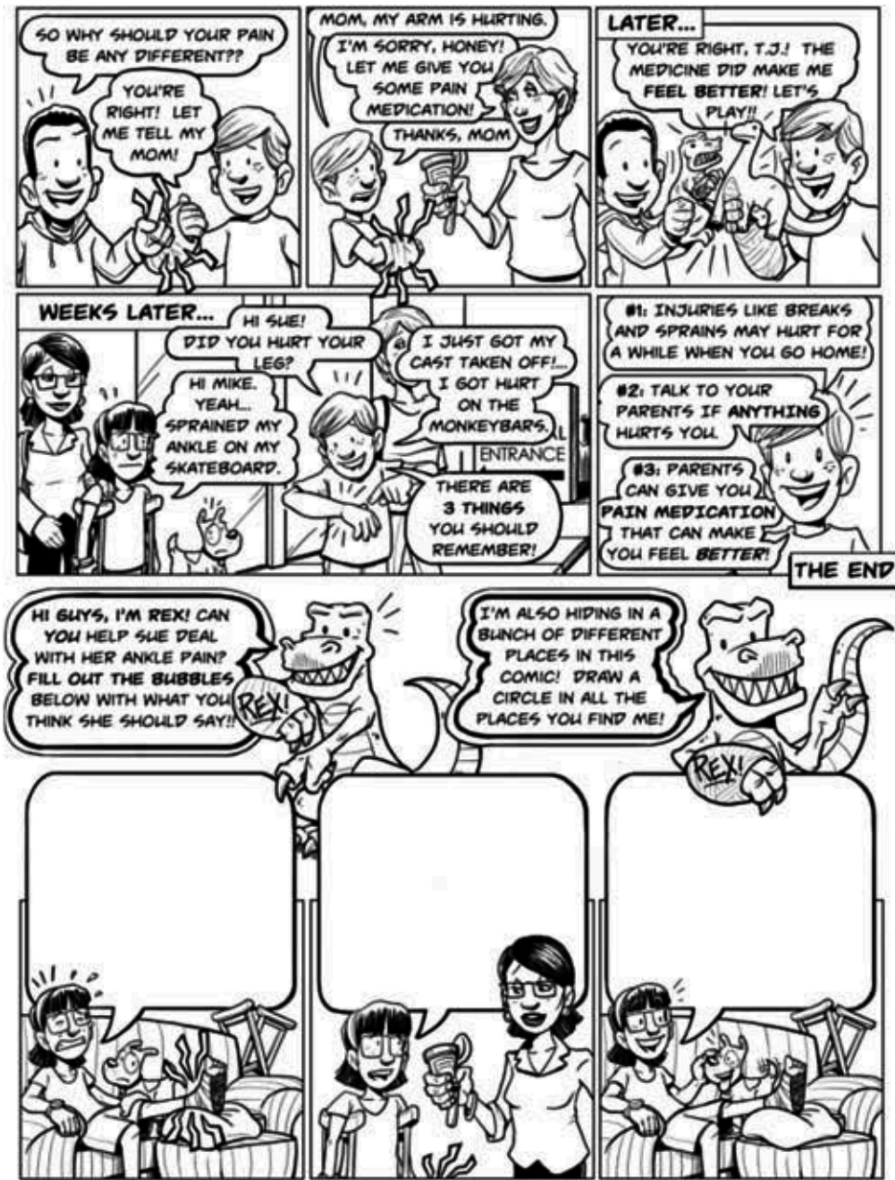
Fig. 2. Comic education module as used in the study by Hanson et al. (2017: 530).

The authors explicitly highlight the comics' potential to "communicate more than just information, because the medium can visualize the fear and anxieties that patients may also feel" (Hanson et al. 2017: 529). In the comic, this visualization becomes clear in the facial expressions of the characters that are prominently displayed in almost all panels of the comic. They do not only show the injured kid's discomfort in general, but lend a visual appearance to their pain.

This becomes very clear in the second panel in line 3 on the first page (see Fig. 2), in which the character on the left, T.J., asks the other character, Mike, whether their arm is hurting, and they reply: 'Yeah'. The logical connection between the pain that is mentioned by T.J. and the facial expression drawn for Mike can be made explicit by analyzing the discourse relation Enhancement/Elaboration between the units in this panel: The small visual details of the facial expression that is showing a sad, unwell kid specifies the verbally expressed 'hurting' even more.

Interestingly, Hanson et al. (2017) do not take into consideration any of the larger structural techniques of constructing the comic's story or the layout of the two pages on the spread. But they play as much of a role for the enhancement of the patients' understanding. This may similarly be made explicit by constructing the overall discourse structure of the comic by analyzing the discourse relations between the panels and larger units. For this, the spread and the overall unfolding story can first be divided into 4 different parts:

1. the first line of the first page constitutes the general description of the situation resulting from an injury, alongside the educational takeaway of three things that one should remember before going home – this is embedded in a conversation and can thus be seen as part of the overall narrative in which several characters experience some events in a specific setting;
2. the second part of the first page as well as the first line of the right page show a setting at home ("a few days later", as the caption says), wherein the two characters play together, eat, and find a solution for the kid's pain;
3. the second line on the second page shows yet another setting ("weeks later…"), in which the injured kid is now teaching the same three things to another friend who is also injured;
4. the bottom part of the second page is the "teach-back portion" (Hanson et al. 2017: 530) that lets the reader fill in speech bubbles to use the gained knowledge immediately for yet another narrative setting with two characters.

All four larger parts can be connected by discourse relations of the N a r -r a t i o n kind, because temporal and spatial sequences between the events that are shown and told become clear in the captions or the setting in the panels. In addition, all four parts share the same topic. Even the final teach-back portion is embedded in one part of the story, since the two characters displayed there, Sue and her mother, are introduced earlier as experiencing a situation that is very similar to the one Mike and T.J. experienced before. The construction of such an overall narrative helps guide readers through the whole page, and the results of Hanson et al.'s survey confirmed that the comic was experienced as "likable, easy to read, and providing important information" (Hanson et al. 2017: 531). Hence, the instructive

genre of the comic as a whole uses a narrative story as its organizational strategy for the combination of both storytelling and educating elements (see Bateman et al. 2017: 314–315). It is exactly here that the potential of a broad discourse relations analysis comes to the fore: The relatively neutral approach to diverse forms of discourse with a very general set of relations makes it possible to describe the different genre patterns and elements simultaneously. This not only makes the complexity of the comics' discourse structure visible but also provides guidance for the interpretation of these specific patterns.

While the individual parts of the overall narrative structure are also mainly constructed by N a r r a t i o n relations holding between the panels, zooming in on individual parts shows what can further be revealed through such an analysis of relations, in this case holding between smaller elements, namely between and within the panels or dialogue parts in several panels. For example, in part 2 of the overall story, a R e s u l t relation can be inferred as holding between the event of asking the mom for a snack (in the right panel in the second row on the second page) and the event of bringing the snack in the next panel (on the left of the third row). This causal relation is exactly what is taken up in the dialogue to show the logical coherence between a cause and an effect in the two panels in the bottom row on the left page: "Think about it, Mike! What did you do when you were hungry? – I told my mom… – And what did she do? – She brought us a snack… – And are you still hungry? – Hmm… no!" Instead of actually explaining that there is a cause-effect relationship between hunger and eating a snack, the next panel then only points to the fact: "So why should your pain be any different?"

A very similar relation can also be inferred in the next panel (second in the top row on the second page) between the event of telling the mom that the arm is hurting and receiving some pain medication. Interestingly, these two events are, here, displayed in the same panel (and not in two subsequent panels), but the drawings are very similar to the ones used in the panel on the first page (and in fact, the event of expressing hunger is repeated in the panel on the left page). There is thus also a discourse relation holding between these two panels (one on the left and the other on the right page), which can be identified as a very strong P a r a l l e l, because both panels show a strong semantic as well as structural similarity. In the fourth part, in the "teach-back portion", this visual design of the specific panel is repeated again (in the middle panel in the last row on the right page) in order to create the same cause-effect relationship for the reader.

This kind of analysis demonstrates nicely how the actual cause-effect relation between having pain, getting pain medication, and feeling better (as is then also shown in the panel in the top right corner on the right page) is made clear for readers in several parts of the comic. The verbal text alone does not make this relationship explicit; it does not, for instance, outright state that pain medication works similar to food in that it somehow fulfills a specific need. Instead, it only implicitly shows this analogy by having a snack brought in when someone is hungry (which in fact establishes yet another

parallel) and having another character ask the question: "Why should your pain be any different?" The analytical identification of causal discourse relations between those panels can more strongly support this interpretation and further explain the reasoning processes. The relations here work as descriptions and representations of thoughts: As signs of knowledge (see section 3.2) that are activated by the rather implicit argumentation in the comic. The analysis can also hypothetically identify the discursive structure in which the R e s u l t and P a r a l l e l relations are embedded as having an educational purpose and, in comparison to the structures that evolve from N a r r a t i o n relations, not a purely narrative nature. This identification is in fact still hypothetical because an empirical analysis of a larger corpus of such structures would have to verify this particular pattern. The analysis as demonstrated here provides the methodological basis for this empirical verification.
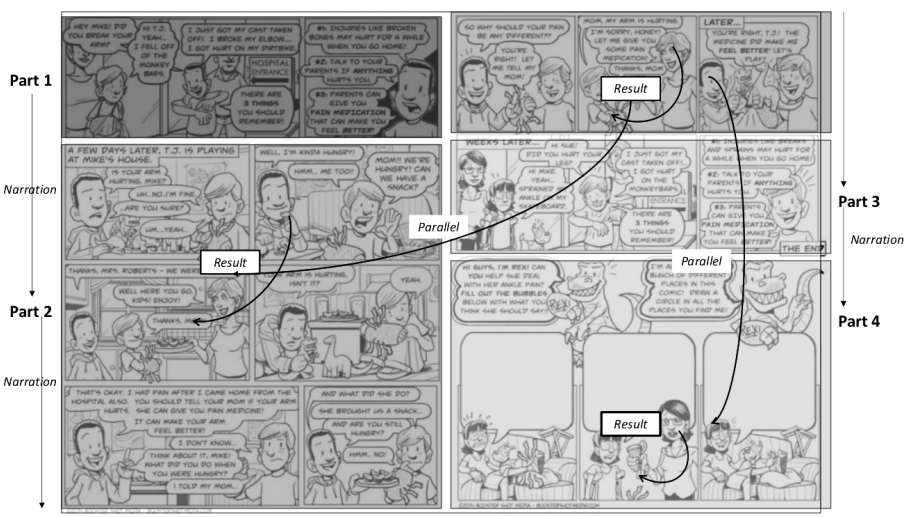
Fig. 3. Graphical illustration of the discourse relations in the comic spread used in Hanson et al. (2017: 530).

Figure 3 represents a graphical illustration of the relations holding between the different units and panels in this comic spread, and figure 4 gives a more abstract representation of the resulting discourse structure of the spread with the various relations between the different parts identified before. As described in section 3, outlining these various discourse relations showcases how they render instantiations of thoughts explicit, on the basis of certain knowledge. In this particular case, it is mainly the cause-effect relationships between several events and processes in the comic that not only show inferable links between specific entities of the discourse alongside corresponding entities in the world, but also explain how these links are meant to be understood and explained to others. An examination of these discourse relations then enumerates the basic processes of reasoning that
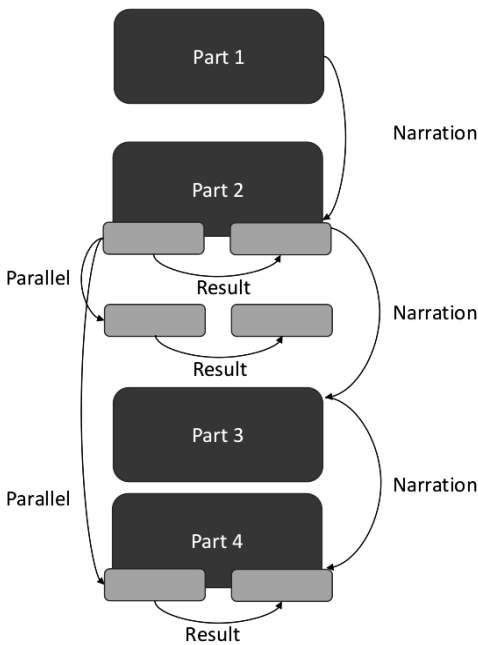
Fig. 4. Abstract representation of the resulting discourse structure of the comic spread.

are expected to happen when children and their caregivers read this comic – and Hanson et al. (2017) indirectly report on these processes when they summarize the results of their study (see also above). Their conclusion highlights that "patients need structured content, presented verbally, with written and visual cues to enhance recall", but provides no further details about this 'structured content' (Hanson et al. 2017: 531). The analysis of the discourse relations involved in constructing such structured content can therefore be seen as an important addition to the overall point that "[c]omic instructions are an innovative medium of communication that can be used to achieve these goals" (Hanson et al. 2017: 531).[6]

5. Conclusion

Understanding how comics work and how readers make meaning out of visual and verbal units has always been a central aim in comics studies. One way of approaching these processes of comprehension theoretically is the analysis of discourse relations between comics units, and this paper gives a broadly oriented overview of this kind of analysis on the basis of previous work in the context of multimodality research. As a result from this previous work, a concise set of discourse relations has been presented. By setting out these relations as logical relations and as signs of knowledge, the discussion has provided ways of making thoughts and knowledge structures for the construction and design of comics explicit. A short example analysis of discourse relations in an educational comic has demonstrated the applicability of this set to a specific comic genre. Similar to some of the previous work, the paper first of all showcases yet another case of building analytical hypotheses for the understanding of comics on the basis of semiotic and discourse analysis.

Hence, this account is clearly to be distinguished from cognitively oriented approaches in the realm of psychology and neuroscience. Many other approaches have similarly addressed and are still pushing a theoretical-meth-

odological account to meaning-making in comics, and our own account also strongly connects to advancements in cognitive psychology and the study of discourse comprehension (such as Asher and Lascarides 2003; Kamp and Reyle 1993). Nonetheless, criticism has continuously been raised against this type of research, particularly with regard to the lack of "adequate evidence" and the need to "be verifiable through some sort of testing" (Cohn 2014: 57, 68). This is understandable insofar as empirical work on comics through psychological experimentation providing such evidence presents an important enrichment for the discipline of comics studies. However, theoretical advancements as well as consolidation and combination of existing theoretical approaches are similarly needed in this dynamically evolving field, which is often challenged by transdisciplinary discussions and sometimes loses sight of important theoretical foundations.

Consequently, the approach presented in this paper contributes to the aim of building a 'better comic theory' (cf. Cohn 2014) by providing arguments for and explanations of the "necessary 'system-internal' (logico-philosophical) organizations of […] processes of signification" (Bateman 2017: 21) by delivering starting points for the empirical analysis of the brain processes following up on, or activated by, these signification processes. As Bateman further points out, "establishing connections between these levels of description constitutes a challenging and worthwhile research task of its own" (Bateman 2017: 21), and, as explained in the introduction, the discussion in this paper explicitly aims at contributing such a connection. Moreover, and with particular regard to the development of multimodal methods for the analysis of comics, the kind of theoretical foundation offered in this paper also serves the need for more qualitative, or discursive, accounts that complement the trend towards a stronger use of experimental methods in interdisciplinary environments (see the discussion in Bateman 2022 as one example).

Notes

1   There is for example a far-reaching consensus about the relation that indicates a spatio-temporal sequence, which is called Narration, but there are many different forms of Elaboration, e.g. expansion, enhancement, etc.

2   I thank the anonymous reviewer for their careful consideration of this analytical step. Parts of our previous work deal with this topic more substantially, see, e.g., Bateman and Wildfeuer 2014b and Wildfeuer 2019.

3   A further example of a very explicit connection to earlier philosophical works is given by Asr and Demberg (2012) in their introduction to a paper on the implicitness of discourse relations: "David Hume, in his prominent work 'An enquiry concerning human understanding' proposed that ideas in the human mind were associated according to at least three types of relations: resemblance, contiguity in time or place, and causality (Hume, 1784). Since then, many language scientists have tried to adapt this idea about human general reasoning to the world of language [...]" (Asr and Demberg 2012: 2670).

4   As part of his broader conceptualization of logics in general, Peirce also introduc-
     es (first in his work "Short Logic"; Peirce 1893–1913: 11) the notion of a 'specula-
     tive rhetoric' which he later defines as "the theory of advancements of knowledge
     of all kinds" (Peirce 1893–1913: 256) and which Rellstab (2008: 327) paraphras-
     es as "an investigation of the human mind" in order to "know just what the process-
     es are whereby an idea can be conveyed to a human mind and become embed-
     ded in its habits" (Peirce 1893–1913: 330). A prerequisite for this process of under-
     standing is a 'common ground', "a set of believes taken for granted as part of the
     background of conversation. Communication, then, can be seen as the re-consti-
     tution of this common ground" (Rellstab 2008: 328).
5   Interestingly, many of these genres often additionally use a narrative structure
     when giving instructions or aiming at enhancing readers' understanding. We dis-
     cuss this further in the use case chapter on comics and graphic novels in Bateman
     et al. (2017: Chapter 12) and also examine a meta-comic which still uses a narra-
     tive structure as its organizational strategy.
6   And subsequently, it is of course also of particular interest to test empirically wheth-
     er the children and caregivers actually process this kind of reasoning – a question
     that should then be addressed by more experimental, user-oriented studies.

Bibliography

Asher, Nicholas and Alex Lascarides (2003). *Logics of Conversation*. Cambridge: Cam-
     bridge University Press.
Asr, Fatemeh T. and Vera Demberg (2012). Implicitness of discourse relations. *Proceed-
     ings of the 25th International Conference on Computational Linguistics COLING,
     Mumbai, India*, 2669–2684.
Bateman, John A. (2014). *Text and Image. A Critical Introduction to the Visual/Verbal
     Divide*. London: Routledge.
Bateman, John A. (2017). Peircean semiotics and multimodality: Towards a new syn-
     thesis. *Multimodal Communication* 7, 1.
Bateman, John A. (2021). The foundational role of discourse semantics beyond lan-
     guage. In: Michele Zappavigna and Shoshana Dreyfus (eds.). *Discourses of Hope
     and Reconciliation. On J.R. Martin's Contribution to Systemic Functional Linguistics*.
     London: Bloomsbury, 39–55.
Bateman, John A. (2022). Multimodality – where next? Some meta-methodological con-
     siderations. *Multimodality and Society* 2, 1, 41–63.
Bateman, John A. and Klaas Rondhuis (1997). Coherence relations: Towards a gener-
     al specification. *Discourse Processes* 24, 3–49.
Bateman, John A. and Janina Wildfeuer (2014a). A multimodal discourse theory of visu-
     al narrative. *Journal of Pragmatics* 74, 180–208.
Bateman, John A. and Janina Wildfeuer (2014b). Defining units of analysis for the sys-
     tematic analysis of comics: A discourse-based approach. *Studies in Comics* 5, 2,
     373–403.
Bateman, John A., Janina Wildfeuer and Tuomo Hiippala (2017). *Multimodality. Foun-
     dations, Research, Analysis. A Problem-Oriented Introduction*. Berlin: De Gruyter.

Bellucci, Francesco (2014). "Logic, considered as semeiotic": On Peirce's philosophy of logic. *Transactions of the Charles S. Peirce Society: A Quarterly Journal in American Philosophy* 50, 4, 523–547.

Bramlett, Frank, Roy T. Cook and Aaron Meskin (eds.) (2016). *The Routledge Companion to Comics*. London and New York: Routledge.

Cohn, Neil (2014). Building a better 'comic theory': Shortcomings of theoretical research on comics and how to overcome them. *Studies in Comics* 5, 1, 57–75.

Cohn, Neil and Patrick Bender (2017). Drawing the line between constituent structure and coherence relations in visual narratives. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 43, 2, 289–301.

Cook, Roy T. (2012). Why comics are not films. Metacomics and medium-specific conventions. In: Aaron Meskin and Roy T. Cook (eds.). *The Art of Comics: A Philosophical Approach*. Blackwell, Maiden and Oxford: Wiley, 165–187.

Feng, Dezheng and Kay L. O'Halloran (2012). Representing emotive meaning in visual images: A social semiotic approach. *Journal of Pragmatics* 44, 14, 2067–2084.

Frege, Gottlob (1980). *Translations from the Philosophical Writings of Gottlob Frege*. Edited by Peter Geach and Max Black. London: Blackwell.

Forceville, Charles, Elisabeth El Refaie and Geert Meesters (2014). Stylistics and comics. In: Michael Burke (eds.). *The Routledge Handbook of Stylistics*. Oxon and New York: Routledge 485–499.

Graesser, Arthur C., Danielle S. McNamara and Max Louwerse (2003). What do readers need to learn in order to process coherence relations in narrative and expository text? In: Andrew Sweet and Christopher Snow (eds.). *Rethinking Reading Comprehension*. New York and London: The Guilford Press, 82–98.

Halliday, Michael A. K. and Ruqaiya Hasan (1979). *Cohesion in English*. London: Longman.

Hanson, Aaron, Amy L. Drendel, Gary Ashwal and Alex Thomas (2017). The feasibility of utilizing a comic for education in the emergency department setting. *Health Communication* 32, 5, 529–532.

Harris, Miriam (2007). Cartoonists as matchmakers: The vibrant relationship of text and image in the work of Lynda Barry. *Elective Affinities: Testing Word and Image Relationships* 6, 6, 129–146.

Hartung, Matthias and Philipp Cimiano (2007). To abduce or not to abduce, that is the question. *On the Role of Abductive Inference within Discourse Construction*. Heidelberg: Universität Heidelberg.

Hobbs, Jerry R. (1979). Coherence and coreference. *Cognitive Science* 3, 1, 67–90.

Hobbs, Jerry R. (1983). Why is discourse coherent? In: Fritz Neubauer (eds.). *Coherence in Natural Language Texts*. Hamburg: Buske, 29–70.

Hobbs, Jerry R. (2004). Abduction in natural language understanding. In: Laurence R. Horn and Gregory Ward (eds.). *Handbook of Pragmatics*. London: Blackwell, 724–741.

Hume, David (1784). *An Enquiry Concerning Human Understanding*. New York: The Liberal Arts Press, 1955 edition.

Jacobs, Dale (2013). "There are no rules. And here they are": Scott McCloud's making comics as a multimodal rhetoric. *Journal of Teaching Writing* 29,1, 1–20.

Jappy, Tony (2013). *Introduction to Peircean Visual Semiotics*. London: Bloomsbury.

Joswick, Hugo (1996). The object of semeiotic. In: Vincent M. Colapietro and Thomas M. Olshewsky (eds.). *Peirce's Doctrine of Signs. Theory, Applications, and Connections*. Berlin: De Gruyter, 93–102.

Kamp, Hans and Uwe Reyle (1993). *From Discourse to Logic. Introduction to Model-theoretic Semantics of Natural Language, Formal Logic and Discourse Representation Theory*. Amsterdam: Kluwer Academic Publishers.

Lessing, Gotthold Ephraim (1990 [1766]). *Werke 1766–1769: Laokoon*. Frankfurt a.M.: Deutscher Klassiker Verlag.

Liu, Yu and Kay L. O'Halloran (2009). Intersemiotic texture: Analyzing cohesive devices between language and images. *Social Semiotics* 19, 4, 367–388.

Mann, William C. and Sandra A. Thompson (1988). Rhetorical structure theory: Toward a functional theory of text organization. *Text* 8, 243–282.

Martinec, Radan and Andrew Salway (2005). A system for image-text relations in new (and old) media. *Visual Communication* 4, 3, 337–371.

McCloud, Scott (1994). *Understanding Comics: The Invisible Art*. New York: Harper Perennial.

Michael, Emily (1978). Peirce's adaptation of Kant's definition of logic: The early manuscripts. *Transactions of the Charles S. Peirce Society* 14, 3, 176–183.

Moriarty, Sandra E. (1996). Abduction: A theory of visual interpretation. *Communication Theory* 6, 2, 167–187.

Newton, Isaac (1934 [1686]). *Mathematical Principles of Natural Philosophy, Vol. 1: The Motion of Bodies, and Vol. 2: The System of the World*. English translation by Andrew Motte and Florian Cajori. Berkeley, CA: University of California Press.

Packard, Stephan, Andreas Rauscher, Véronique Sina, Jan-Noël Thon, Lukas R.A. Wilde and Janina Wildfeuer (2019). *Comicanalyse. Eine Einführung*. Cham: Springer.

Peirce, Charles S. (1865). Logic of the Sciences. MS [W] 113; MS [R] 769, 734, 729, 921, 922, 728, S 66.

Peirce, Charles S. (1893–1913). Short logic. Chapter 1: Of reasoning in general. In: *The Essential Peirce: Selected Philosophical Writings: Volume 2*. London: Bloomington, 11–26.

Peirce, Charles S. (1976). *The New Elements of Mathematics by Charles S. Peirce*. Edited by Carolyn Eisele. The Hague and Paris: Mouton & Co.

Peirce, Charles S. and Christine Ladd-Franklin (1902). Universe of discourse. *Dictionary of Philosophy and Psychology*, volume 2, The Macmillan Company, Macmillan and Co, 742, URL: https://archive.org/stream/philopsych02balduoft [retrieved May 7, 2024].

Rellstab, Daniel H. (2008). Peirce for linguistic pragmaticists. *Transactions of the Charles S. Peirce Society* 44, 2, 312–345.

Saraceni, Mario (2000). Relatedness: Aspects of textual connectivity in comics. In: Jan Baetens (ed.). *The Graphic Novel*. Leuven: Leuven University Press, 167–179.

Spanjers, Rik (2021). Text-image relations. In: Sebastian Domsch, Dan Hassler-Forest and Dirk Vanderbeke (eds.). *Handbook of Comics and Graphic Novels*. Berlin and Boston: De Gruyter, 81–98.

Tseng, Chiao-I and John A. Bateman (2018). Cohesion in comics and graphic novels: An empirical comparative approach to transmedia adaptation in *City of Glass*. *Adaptation* 11, 2, 122–143.

Van Leeuwen, Theo (1991). Conjunctive structure in documentary film and television. *Continuum* 5, 1, 76–114.

Varnum, Robin and Christina Gibbons (2001). *The Language of Comics: Word and Image*. Mississippi: University Press of Mississippi.

Wartenberg, Thomas E. (2012). Wordy pictures: Theorizing the relationship between image and text in comics. In: Aaron Meskin and Roy T. Cook (eds.). *The Art of Comics: A Philosophical Approach*. Blackwell Publishing Ltd., 87–104.

Wildfeuer, Janina (2014). Formale Zugänge zur Diskursanalyse. *Zeitschrift für Semiotik* 35, 3–4, 393–417.

Wildfeuer, Janina (2019). The inferential semantics of comics: Panels and their meaning. *Poetics Today* 40, 2, 215–234.

Wildfeuer, Janina (2021). Discourse semantics and textual logic: Methodological considerations for multimodal analysis. *OBST – Osnabrücker Beiträge zur Sprachtheorie* 99. Special Issue on Linguistics and Multimodality edited by Arne Krause and Ulrich Schmitz, 87–114.

Wildfeuer, Janina, Ielka van der Sluis, Gisela Redeker and Nina van der Velden (2022). „No Laughing Matter!? Analyzing the page layout of instruction comics". *Journal of Graphic Novels and Comics* 14, 2, 186–207.

Wolff, Christian (1963 [1728]). *Preliminary Discourse on Philosophy*. Indianapolis: Bobbs-Merrill.

Sources

Emma (wri/art) (2017). *The Mental Load: A Feminist Comic*. New York: Seven Stories Press.

Image Sources

Fig. 1. Emma (wri/art) (2017: 106).
Fig. 2. Hanson et al. (2017: 530).
Fig. 3. Hanson et al. (2017: 530).

Asst.-Prof. Dr. Janina Wildfeuer
Rijksuniversiteit Groningen
Faculteit Letteren – CIW
Oude Kijk in 't Jatstraat 26
9712 GK Groningen
The Netherlands
E-Mail: j.wildfeuer@rug.nl

Vorschau auf den Thementeil der nächsten Hefte

Nachfolgend sind die geplanten Themenhefte der Zeitschrift für Semiotik aufgeführt. Autor/-innen mit Interesse zur Abfassung von Beiträgen, Einlagen und Institutionsberichten können sich über die Adresse zsem.redaktion@tu-chemnitz.de direkt an die Redaktion der Zeitschrift für Semiotik wenden.

Selbstoptimierung
Marcel W. Lemmes, Benedikt Matt und Klaus Sachs-Hombach

Themenoffenes Heft
Ellen Fricke und Martin Siefkes