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 narrative 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' cannot be considered similar to communication via language that more focused engagement with the issues is deemed unnecessary. Approaches to comics 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 'grammars' 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 linguistic-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 (Horstkotte 2015: 34–35).

In these few sentences Horstkotte conflates a host of theoretically questionable assumptions as if, for comics and graphic novels, they were self-evidently the case. 'Grammar' is equated with a requirement of 'linear reading', something allegedly inappropriate for readers' engagements with comics 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 Horstkotte 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 achievement 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 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 multimodal 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 important question for further analyses and frameworks becomes not 'if', but just what kinds of inference are necessary and when do they occur, operating on what kinds of premises. Some of these will be near to perception (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 phenomena in any artefacts under analyses and broader interpretations. This contribution of textuality will be a component of all the examples discussed 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 increasing application in analyses of comics and graphic novels (e.g., Kukkonen 2015) and is also taken as an inalienable component of the model of semiosis drawn upon here (cf. van Leeuwen 1999; Bateman 2019); further discussion 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 multi-modal 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, Bateman, Neuhaus, Mossakowski and Bhatt 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-
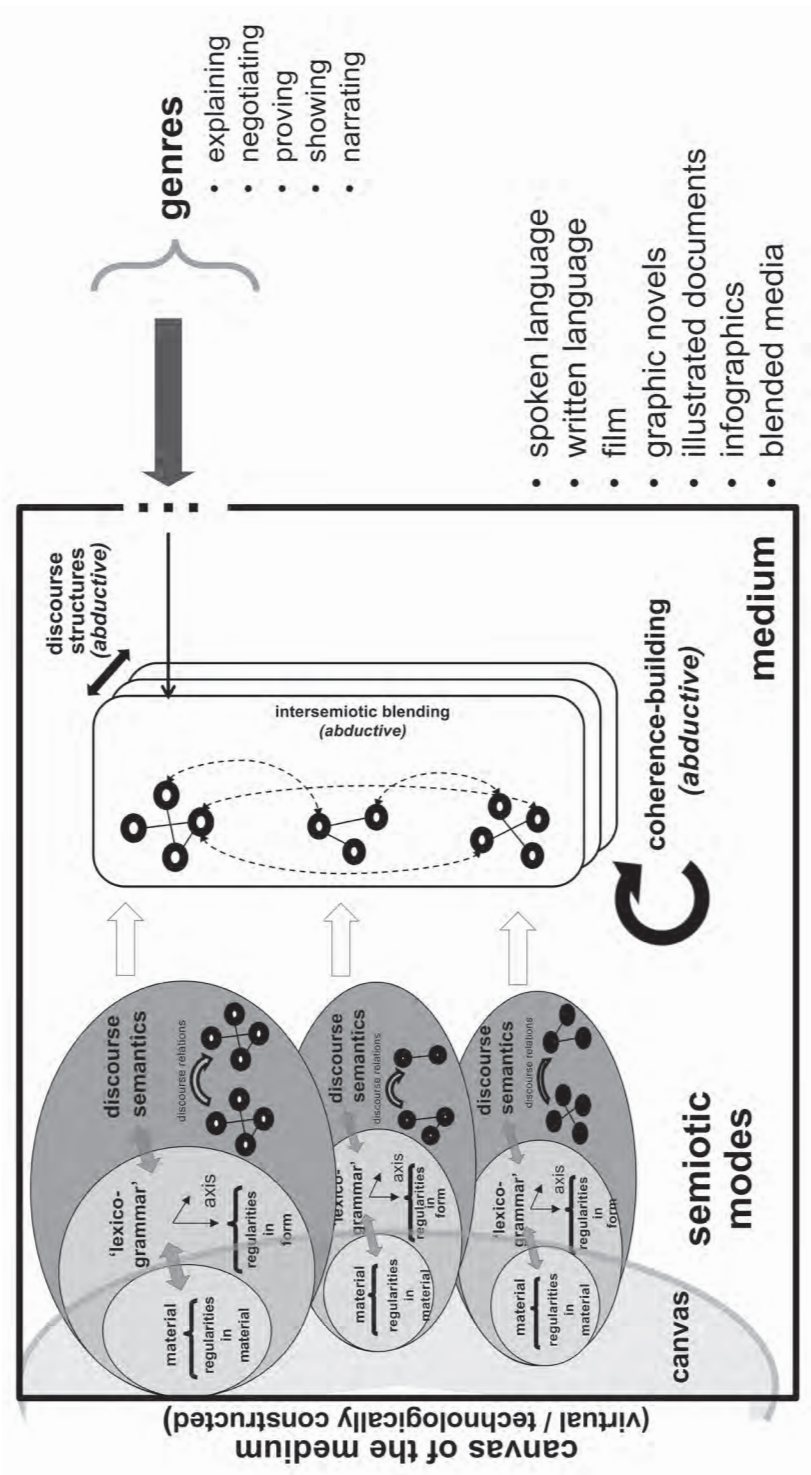
Fig. 1. Graphical summary of the model of multimodal communication defined by Bateman, Wildfeuer and Hiippala (Bateman et al. 2017: 124, Bateman 2016).

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-
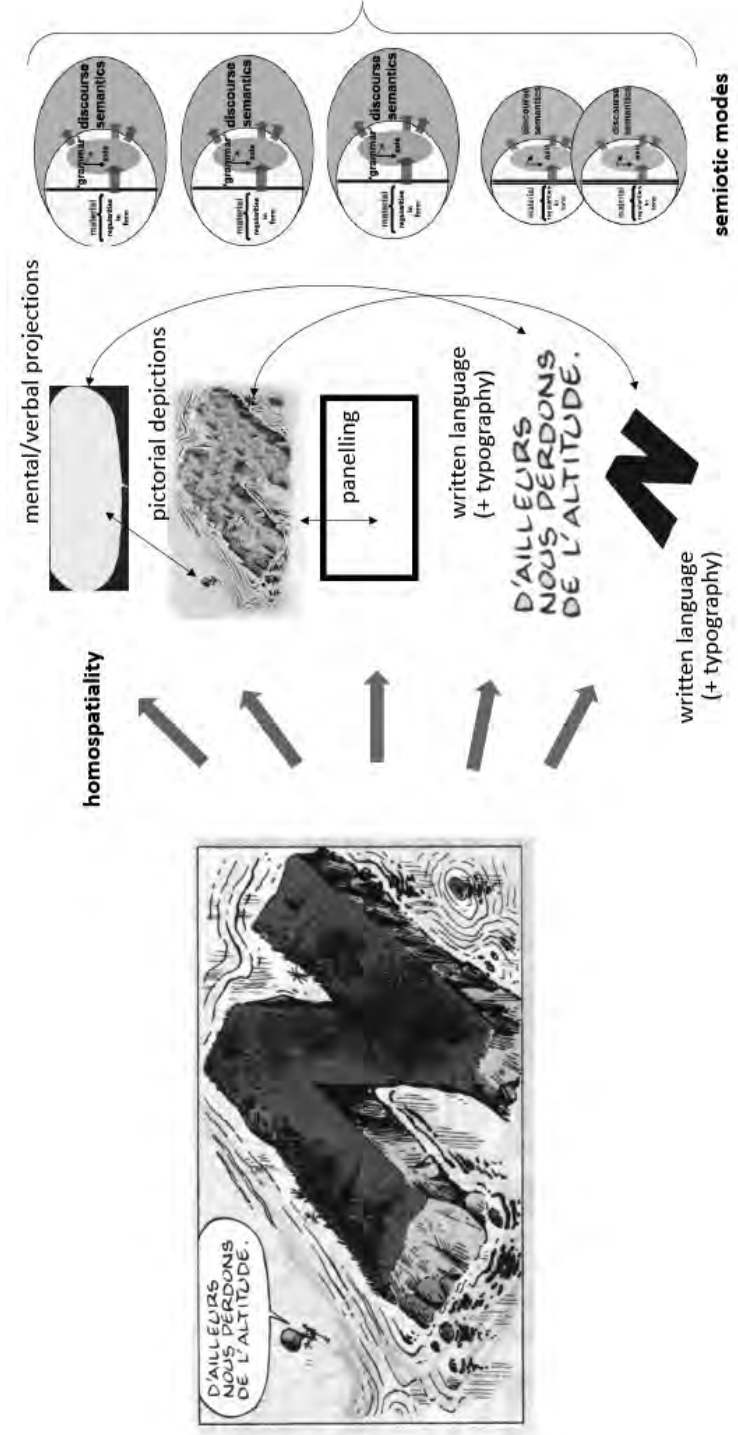
Fig. 2. Left: *Philémon: L'Intégral.* Fred (2011: 152). Right: decomposition according to contributing semiotic modes.

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 arte-fact 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 particu-larly interesting cases. These show well the benefits of adopting a more sys-tematic 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 signi-fier-signified pairs or Hjelmslevian form-expression ensembles. Semiotic systems of these kinds are generally insufficient for characterising commu-nication, however, precisely because of their lack of provision for interpre-tation. The semiotic mode construct introduced in the previous section addresses this concern by explicitly complementing the relating of materi-al regularities to qualitative categories with the further inferential compo-nent now labelled discourse semantics. This is the concrete analytic mech-anism 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 inter-pretation:

> 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 descrip-tion 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, tech-nical 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 pur-poses 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 materi-al 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 (2023) 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
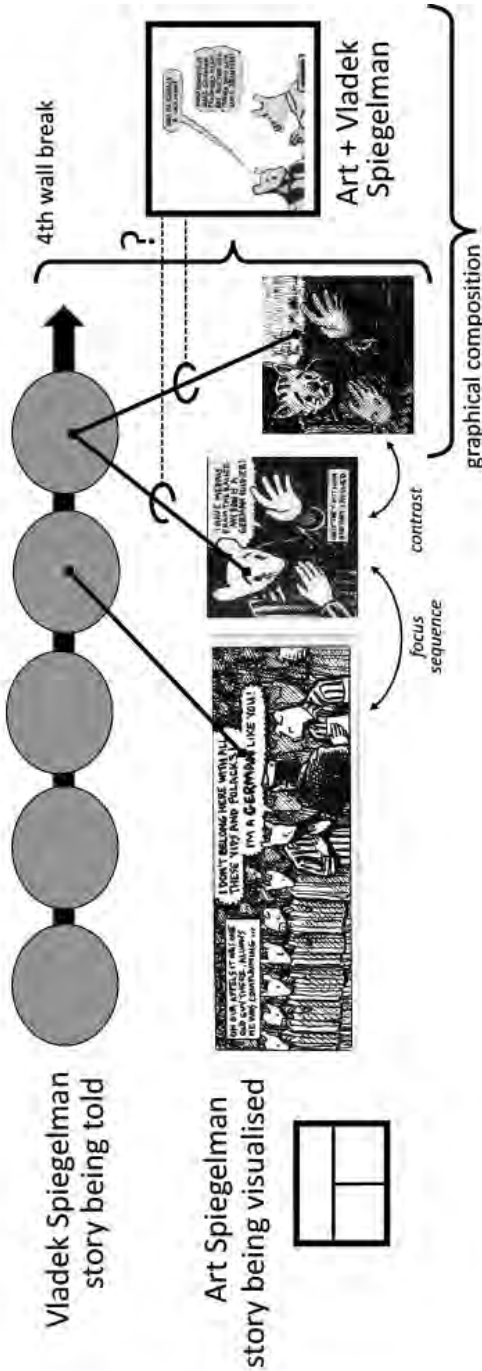
Fig. 4. Discourse relations among the components of the panels from Art Spiegelman's (1991: 50). *Maus: A Survivor's Tale. II: And here my troubles began*.

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

Jews and Germans in doubt. Their questionable status is expressed graph-
ically by backgrounding a depiction of that alternative (by placing it literal-
ly 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 con-
trast, but the motivation for that contrast would be lost. Instead, the relation
is cued explicitly and unambiguously by breaking the visualisation and turn-
ing 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 rejec-
tion of a neat resolution to the issue is then maintained graphically as both
visual renditions are offered to the reader. Thus, uncertainty itself is depict-
ed graphically as well as verbally in a mutually supportive fashion. The con-
trasting 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 inter-
pretation 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 interpre-
tations 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 interpretation 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 precisely 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 employing pictorial semiotic modes, where it may not be possible to assign context-free interpretations to individual formal 'items', such as lines, brushstrokes, 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 interpretative guidance is designed into what is shown. Moreover, even when such strategies are deployed in a narrative, this is nothing particularly specific 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 mechanisms 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 distributed 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 organisation in comics and graphic novels remains strongly sequential in organisation. 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 provide 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 cre-
ated 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 high-
ly 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 extend-
ed narrative (and other) sequences. Consequently, one of the primary aims
of this contribution has been to show how a more contemporary and fine-
ly articulated account of the semiotics of multimodality can be used to dis-
tinguish 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 consti-
tution (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 approach-
es 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 cog-
nitive 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 can-
not be 'in' a text as this would violate basic Peircean understandings of semi-
osis (i.e., wrongly attributing Thirdness to Secondness, i.e., what is 'pres-
ent'); this is also the traditional distinction drawn, for example, between cohe-
sion and coherence in text linguistics (cf. Halliday and Hasan 1976; de Bea-
ugrande 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 solu-
tion to this, we have proposed in this article a thorough reorientation to semi-
otic 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 ana-
lytic access to highly complex graphic materials such as comics and graph-
ic 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 a 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: Universiade 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 Rheory. 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. Springer, 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 Pres, 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 (2023/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