From Narrative to Visual Narrative to Audiovisual Narrative: the Multimodal Discourse Theory Connection

John A. Bateman

Bremen Institute of Transmedial Textuality Research, Bremen University, Bremen, Germany
`bateman@uni-bremen.de`

───── **Abstract** ─────

Models of narrative have been proposed from many perspectives and most of these nowadays promote further the notion that narrative is a transmedial phenomenon: i.e., stories can be told making use of distinct and multiple forms of expressions. This raises a range of theoretical and practical questions, as well as rendering the task of providing computational models of narrative both more interesting and more challenging. Central to this endeavour are issues concerned with the potential mutual conditioning of narrative forms and the media employed. Methods are required for isolating narrative properties and mechanisms that may be generalised across media, while at the same time appropriately respecting differences in medial affordances. In this discussion paper I set out a corresponding approach to characterising narrative that draws on a fine-grained formal characterisation of multimodal discourse developed on the basis of both functional and formal linguistic models of discourse, generalised to the multimodal case. After briefly setting out the theoretical principles on which the account builds, I position narrative with respect to the framework and give an example of how audiovisual narratives such as film are accounted for. It will be suggested that a common anchoring in a well specified notion of discourse as an intrinsically multimodal phenomenon offers beneficial new angles on how narratives can be modelled, as well as establishing bridges between humanistic understandings of narrative and complementary computational accounts of narratives involving communicative goal-based planning.

1 Introduction

It has become relatively common to discuss models of narrative by drawing on various theoretical constructs and ideas proposed within the broad field of narratology. However, the narratologist David Herman [18, pp. 47–48] sets out what he describes as two 'great ironies' of narratological work on narrative as follows. First, although Barthes in his early considerations of accounts of narrative in the 1960s and 1970s had already called for a 'second linguistics' going beyond the limits of the sentence and structural linguistics at that time [4, p.83], narratology instead adopted many tenets of precisely that structural linguistics Barthes was urging moving beyond. And, second, despite the fact that many foundational works on narrative had begun by addressing oral storytelling, narratology has since been

largely dominated by an orientation to written literary texts. This shaping of the field has, it will be suggested in this discussion paper, led to a systematic lack of attention to areas of textual description that are now required in order to track narratives as they move across 'media' and purposes – as in particular when we see narrative aspects being drawn on in the contexts of games, argument, biography and so on.

Less well considered is the fact that the last 20 years has seen within various branches of linguistics precisely that 'second linguistics' that Barthes was looking for. I will call this for current purposes 'discourse linguistics', although there are several contributing areas that use a range of nomenclatures spanning functional linguistics, formal linguistics, psycholinguistics, computational linguistics and more. The central unifying feature drawn on here is that we now have a rich foundation of discourse-related linguistic methods for dealing with connected texts 'beyond' the sentence. These have only been applied fragmentarily to the issues typically discussed in narratological approaches to narrative despite the possibilities suggested of beginning to anchor more abstract interpretative schemes in empirically-grounded linguistic research on texts and discourse. This suggests a sustainable point of possible interchange between more humanistic approaches to 'text' and more concrete linguistic-feature oriented approaches from linguistics.

A development of this kind gains considerable further motivation by the fact that both narratology and the computational modelling of narrative are seeking to move beyond verbal narratives. There is considerable discussion of what is labelled 'transmedial narrative' (e.g., [29, 30, 32]) on the one hand, while computational models and formalisms are being suggested not only for connected text but also for film and virtual cinematography, for combinations of image and text as in graphic novels, and in interactional contexts and interactive narrative. Whereas narratology still has rather limited experience in this direction, recent developments in the formal and functional characterisation of *multimodal discourse* can be directly applied. This is then the main focus of this position paper: showing how new accounts of multimodal discourse may serve as a powerful intermediate level of modelling that would connect abstract narrative concerns with concrete, and thereby modellable, features of narrative artefacts and performances.

The structure of the paper is consequently as follows. First I set out how a linguistically-inflected model of discourse can be applied multimodally. Second, I suggest how this may relate to narrative concerns. And finally, I set out some directions of current and future work that may take us further, including applications to interactive narrative and games.

2   From discourse linguistics to multimodal discourse analysis

The essential property of discourse (and texts) that required the move beyond sentence-oriented accounts is the very different mechanisms of meaning construction that become relevant. Whereas sentence meaning can largely be characterised in terms of compositional semantics operating within monotonic logics of various kinds, this does not work for texts. As long stated in text linguistics of almost all persuasions, a text is not a kind of 'super-sentence' but belongs to a different class of linguistic behaviour.

Consequences of this qualitative change can be seen even within sentences, however, as in the trivial examples:

1. She went to the park and played football.

2. She played football and went to the park.

Regular interpretations of these two sentences would suggest differing temporal sequencing of the events described, despite the fact that a logical conjunction makes no such commitments.

Moreover, the temporal interpretations are *defeasible*: that is, it may well be, in particular contexts of use, that no such temporal assertion was intended. Such variations in meanings are often shunted to areas of pragmatics and contextualisation, drawing on aspects of world knowledge and additional reasoning mechanisms.

'Classical' theories of text interpretation generally attempt to characterise this process of making meaning from texts as an inference process where particular propositions are derived from the logical propositional content of the utterances of the verbal text. These are then further developed by the application of cultural and background knowledge of various kinds. Problematic with this family of approaches is that it is generally difficult to characterise precisely *which* facets of background/cultural knowledge are required in any way apart from simply describing the text. Since as competent language users and text producers we generally know what texts mean, the 'analysis' comes to describe what we already knew the text means. This is not an adequate theory of meaning construction in texts for many reasons, and particularly not for any attempt to provide computational models.

Two objections that are particularly important are the following. First, pulling in information as required from world knowledge is a potential blackhole in terms of the extent of the reasoning that is required: this is not realistic in that hearers evidently come to interpretations very quickly and appear *not* to pull in arbitrary quantities of background knowledge – methods must be provided which indicate more specifically just what knowledge may be required and in what detail. Second, the interpretation process is in many respects backwards in that it assumes detailed world knowledge in order to understand the text. In fact, the text as text *asserts* world knowledge, i.e., it presupposes that certain relationships hold, regardless of whether the interpreter has the necessary knowledge already or not. Only by this means can text also function as constitutive of meanings and connections in a culture rather than simply echoing them.

Such implicit assertion of relationships occurs at various levels of generality and abstraction: commonly the implicit assertions of a text are sufficiently general, relating to issues of temporal order, causality, explanations, etc., that it is not necessary for an interpreter to proceed to further levels of testing *unless necessary for the specific communicative goals of the text*. This critical distinction between pursuing text understanding in terms of world knowledge and pursuing text understanding in terms of a text's implicit assertions is the basis underlying several current dynamic models of discourse. Within formal discourse semantics such *dynamic* properties of textual interpretation first received focused attention and formalisation with the development of inherently dynamic semantics, within which the semantics of an entire sequence of contributing blocks of information depends crucially on the ordering of that sequence (e.g., [20, 17]). This has been progressively extended since that time.

Currently the most prominent linguistic account of the dynamic properties of discourse semantics is that couched in Asher and Lascarides' Segmented Discourse Representation Theory (SDRT: [21, 2]). Similar emphases of the dynamic nature of discourse interpretation and production occur across most approaches to discourse, including both formal and functional linguistic schools (cf. [22]). The level of interpretation of discourse semantics that such models provide forms a mediating buffer between general world and cultural knowledge and the information that is directly required to interpret a text. Discourse interpretation then operates by constructing a semantic representation for each incoming discourse contribution, traditionally a sentence or utterance, which is then linked by means of discourse relations into a growing discourse structure. Discourse relations are defined so that both their applicability to particular semantic representations and the requirements they make of context are made explicit. They thus look both 'downwards' towards concrete linguistic forms (and their

compositional semantics) and 'upwards' towards context. The requirements made of context define precisely the ways in which identifiable 'gaps' in interpretation are both created and resolved.

Important for us here is that it is possible to extend such a treatment of dynamic discourse semantics *multimodally* so that it can be applied to any similarly dynamic 'textual' artefact or performance, regardless of the modalities, or forms of expressive resources, that are employed – by these means we can draw connections across verbal texts and audiovisual discourses as realised in film as well as conjoined verbal-visual texts such as graphic novels and comics. In several papers, our group in Bremen has set out how such an extension of modelling capabilities can be applied to a diverse range of media (e.g., [5, 9, 11, 34]); a detailed introduction to the foundations of the approach is given in Bateman and Wildfeuer [10], while the particular characterisation of the *modes* of multimodality, and their relation to discourse semantics on the one hand and to more traditional notions of 'media' on the other, is given in Bateman [7, 8].

A brief example showing a discourse semantic view of what happens during the interpretation of a fragment of film will help clarify this position. It can quickly be established that viewers are making assumptions of meanings for such fragments over and above an appreciation of what happens in each shot and these assumptions show striking similarities to effects observed in verbal discourse interpretation requiring both dynamic interpretation and defeasible reasoning. A sequence decomposition of a fragment from Michelangelo Antonioni's *Blow Up* (1966) taken from towards the end of the film is shown in Table 1. During the entire fragment there is no dialogue or non-diegetic sound (simplifying our task); we only hear leaves being blown in the wind; frame numbers run from the beginning of the extract and are approximate. Although *Blow Up* is a classic film for the purposes of interpretation from many theoretical and analytic perspectives, including cultural and narrative concerns, our interpretation here will focus more precisely on how a well specified multimodal discourse analysis can pick out a film fragment's operation in terms of generating expectations. Such expectations are also driven partially by convention, which may then be intended in the film's composition in order to create particular opportunities for filmic meaning articulation.

Following the sequence through, we have the following interpretation; for readers unfamiliar with the fragment, the effect to be described can probably be gained (although significantly weakened) simply by examining the images presented in the table in sequence before reading the next paragraph.

First, we see the main character (simply named 'the Photographer') looking for something that he had formerly seen under a tree (shot 1). We focus in on him from above and behind (shot 2) as he continues to look on, now aware that what he is looking for will not be found. At the end of the shot, he looks up. We then see the branches of a tree, blowing in the wind (shot 3). At this point, viewers have a strong expectation concerning just what they are being shown: namely, what the Photographer is looking at, or seeing, when he looked up. This interpretation is now so standard and entrenched in our way of seeing film that viewers will make it without conscious deliberation – it is, in fact, for good perceptual reasons difficult *not* to make this interpretation. Shot 3 then continues with a pan downwards and to the right away from the branches towards the Photographer. When we reach the Photographer (frame 482), however, he is not looking up at the branches after all: he is looking slightly forwards and out of the frame towards the left. He then looks down and then back to the right, finally turning right (frame 721) and walking off in this direction, the camera tracking with him.

This sequence commonly evokes an impression of surprise in attentive audiences – although they may not be able immediately to articulate where this response originated. This

■ **Table 1** Shot-by-shot breakdown of an example from Antonioni's *Blow Up* (1968). Shots are described in terms of the traditional distance continuum from long shot (LS), through medium long and medium close shots (MLS, MCS), down to head-and-shoulders shots (H&S).

Shot	Frame	Focus	Type	Description	Image
1 (6.66s)	0	P (David Hemmings)	MLS	crouching down, looking under a tree	
2 (6.08s)	159	P	MCS	looking around on the ground	
		P	MCS	looks up	
3 (3.54s)	306	tree branches	MLS	seen from below	
(3.79s)	391	branches	MLS	panning down-right ↘	
		branches background and P	MLS	panning down-right ↘	
(9.95s)	482	P	MCS	panning stops on Photographer	
	721	P	H&S	photographer turns and moves right, camera tracking with him	

demonstrates that more is occurring during the interpretation of the film than simply following what is happening. When the Photographer is seen looking to the left of the frame rather than up towards the tree in the middle of shot 3, the previous interpretation of the beginning of shot 3 as showing what he is looking at is thrown into doubt. Conventionally, a shot of branches from below following a shot of someone looking up will be interpreted as a point-of-view shot [12] even though the film never explicitly stated that this is what shot 3 was doing. What has happened is that the film's *structure* makes this claim. In other words, how the film has been structured leads viewers to the defeasible discourse hypothesis that the way that this shot is to be connected into the unfolding discourse is by means of a 'projection' [5], or mental perception, relationship.

This discourse interpretation then needs to be retracted when the pan down towards the right comes to an end showing that the Photographer is not looking at the tree at all. The situation here is even more complex, however, since we have additional cinematic conventions that come to bear. Actually there would be nothing 'in reality' preventing the beginning of shot 3 being related as a point-of-view shot and having the character look somewhere else during the pan downwards. *But this is not how point-of-view shots work.* An additional part of the convention is that during a point-of-view shot the character whose point-of-view is adopted cannot engage in independent action: they are constructed filmically as 'looking' and that is all. Antonioni plays with this convention in several of his films.

Slightly more formally, therefore, we have the following situation of interpretation. In the first two shots we have views of a single character that by virtue of filmic cohesion [33] of location, figure, posture, continuous sound and image properties (such as colour balance), contributes to a structure realising a narrative sequence with the shared topic 'the Photographer'. With the shift in shot 3, we hypothesise a perceptual relationship and so introduce an embedded topic structure subordinate to the point-of-view of the topic of the overall structure. All runs well with this hypothesised discourse structure until the character returns into shot: now we no longer have the option of maintaining an *embedded* topic structure because we appear to have rejoined the discourse structure that we started with. This raises a difficult task of reconciliation for the interpreter, which may well leave observable behavioural effects that can be subjected to empirical investigation. The only fully coherent single reading that combines all the possibilities is then that somehow the embedded topic structure came to an end without being cued explicitly in the filmic material – that is, the continuous pan does not provide any material support for this interpretation. To what extent viewers are willing to maintain interpretations for which there is no support will depend on several factors, including their attentiveness, familiarity with the medium, and willingness to overlook narrative gaps.

Taken together this demonstrates that the organisation of film is expectation generating, conventionalised, motivated (in that it commonly draws on perceptual routines) and violable. And all of these properties are, first, standard properties attributed to verbal discourse and the contributions made to such discourses and, second, well describable in terms of progressively applying defined discourse relations which constitute interpretative hypotheses concerning how the discourse is to be made sense of. In Bateman and Wildfeuer [10] we offer several examples of how discourse relations for various media can be defined; in Bateman [6] there is further discussion of how filmic discourse structures such as the example used here may be conventionalised into filmic idioms in similar ways to constructions in language.

3   From multimodal discourse to transmedial narrative

The discussion so far has opened up many points of contact with standard proposals within narratology. According to Ryan [30, p. 4], for example, narrative necessarily involves:

- construction of a storyworld, individuating agents, objects and their spatial arrangements,
- contingency, including accidents and the deliberate actions of agents,
- linkages between physical states and goals, emotions, intentions so as to produce coherence, motivation, closure and intelligibility.

Moreover, narrative is seen as a means of 'sense-making', of providing explanations. All of these facets were already identified in the brief description of filmic interpretation from a discourse semantic perspective in the previous section. And, indeed, many of the properties commonly associated with narrative can also be seen to be at work in other communicative situations or text types. Introducing topics, maintaining these and providing additional information to produce coherence are common features of most discourse types, not only narrative.

This opens up the main line of inquiry being suggested in this paper: to what extent might it be the case that we can offer more formalised views of some basic narratological constructs by modelling them in terms of discourse semantics? Thus, for example, we might see classic (although still hotly debated) categories such as 'focalisation' (cf. [14]) as *discourse achievements* brought about by hypothesising particular sequences of discourse relations.

This may bring several benefits for advancing the formal and computational modelling of narrative constructs. Consider again the example of the previous section and ask how it would be described in terms of focalisation. Although the relation between point-of-view and focalisation is itself subject to detailed theoretical discussions, we can for current purposes simplify somewhat and propose that we have at the beginning of shot 3 of the sequence a move towards 'internal focalisation', where our access to the storyworld is made relative to that of the Photographer character. This narratological status is then immediately deconstructed as the pan reveals the violation of the convention. *This lets us consider labels such as 'focalisation' in terms of the discourse hypotheses that are taken to hold moment-by-moment during the discourse interpretation of some sequence.*

Several authors in narratology and related studies have emphasised that it is possible to interpret almost anything as a narrative. The more relevant consideration is therefore how much evidence does an artefact or performance *itself* give that such an interpretation is relevant or intended. Wolf describes such features in terms of narrative cues [35]: the concern is then to what extent an artefact or performance provides *narrativisation cues* – the greater the number of cues, the more likely it is that that artefact/performance can be profitably interpreted as narrative. And, as Wolf emphasises, such cues can be found in many media, including paintings and other static depictive representations. By relating such narrative cues back to discourse semantics, we begin to have a link from rather abstract notions of focalisation, on the one hand, and more established mechanisms of discourse planning as commonly applied in the computational modelling of narrative in any case, on the other (e.g., [36, 3, 15]).

In short, we may conceive a formalised account of discourse semantics as an additional and complementary level of description that mediates between manipulable technical features of a medium and more abstract, narratively-centred interpretations and descriptions that would themselves be too far more removed from the details of expression to be drawn directly into planning or interpretation processes. Under this view, an appropriate planning of such narrative devices may be more readily achieved by positioning a level of discourse semantics

between purpose and form – this would mean that increasing identification and empathy, increasing tension, achieving point-of-view, etc. all become *textual* goals to be achieved via discourse.

The planning process would not then need to relate high-level goals such as increasing tension or achieving a point-of-view shot to particular filmic expressive resources directly for planning purposes, but could work instead by means of posting intermediate discourse goals that have been established – primarily by empirical study – to have such desired effects as consequences. Such discourse goals may then range freely over all of the expressive resources that a medium provides (e.g., camera movements, lighting, dialogue, gesture and gaze, facial expression, temporal sequencing and so on), offering a far greater freedom of selection that is nevertheless related back to communicative goals.

4   Notes for the future

As a closing set of comments I will draw out some final considerations that could draw support from the line of inquiry suggested in the previous sections. These relate primarily to the drive to push accounts of narrative and its computational modelling beyond traditional narrative media such as verbal texts and narrative film and to take in the challenges of interactive narrative and computer-supported games. I will note two points of interest here related directly to the proposed role of discourse semantics.

First, the long discussion contrasting narrative-based interpretations of games and game-based interpretations of games (cf. [19, 27, 31, 1]) has shown that it appears untenable to 'reduce' games and game playing to narrative, although there do appear to be narrative elements at work. This demonstrates that it cannot be the case that a field such as 'narratology' provides a suitable foundation for both narrative and games. In contrast, we can well see both narrative and many game activities as involving the construction of coherent discourses, particularly when we adopt the position common to treatments of multimodal discourse that *actions* play a central role (e.g, [26, 13, 16]). Here, moreover, the ready relation of discourse back to interaction in language makes it even more relevant for interactive media such as gaming. This allows us to pursue accounts whereby narrative aspects and gaming actions do not stand in any conflict. Discourse modelling may show both being supported at the same time: some of the discourse relations and structures brought about may be contributing to sequences of actions or turn-taking and so on; others may be acting as narrativisation cues, supporting narrative interpretations of what is occurring.

Second, and related to the first, the idea of interactive narrative has itself run into a range of theoretical concerns that raise doubts whether the two poles – interaction, involving freedom of choice and action on the part of participants, and narrative, involving a generally authored narrative arc with plot points and resolutions – are compatible. If participants can 'do what they want' then there can be little guarantee that a satisfying narrative results (cf. [23, 28]). An anchoring in an approach enriched with more explicit treatments of discourse semantics can also be applied to this state of affairs in order to provide some conceptual clarifications with potential implications for implementation strategies.

Briefly sketched, models of dynamic discourse semantics show that it is not possible to contrast interactive narratives (and games) with more classical narratives on the basis that the former involves an active participant and the latter does not. Since discourse interpretation is *always* seen as the active formation of discourse hypotheses in context in order to provide interpretations, the interpreter is always highly active. It is then possible to consider both the source of interpretative cues and the various kinds of interpretations
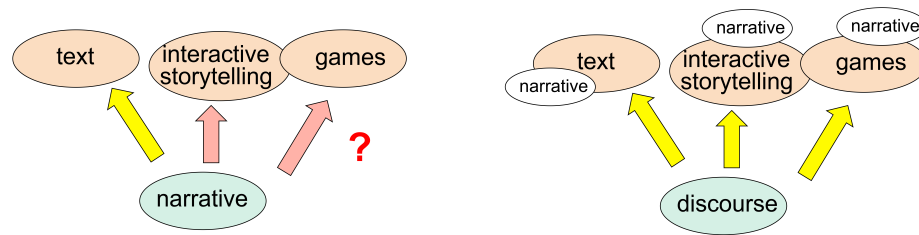
■ **Figure 1** Using discourse semantics to offer a bridge across media.

that may be pursued more carefully. In particular, any contributions that are made on the part of a computational system supporting the interactive narrative can be seen in the light of the narrative cues that they offer. These cues need to be planned and reconciled with the events and states of knowledge that are unfolding within the narrative. They naturally provide a locus for inserting system-derived narrative arcs into the interaction without forcing participants either to notice them or to follow them: this then reflects a discourse that is making a narrativisation 'offer' to its participants – with respect to the management of these cues, it is the computational system that best fits the role of 'narrator'.

However, and *simultaneously*, each and every participant must also be seen as an active participant and a necessary component of that participation is the construction of discourse coherence for what is occurring. This discourse coherence can just as well include narrative interpretations that differ from those that may have been planned by the computational narrator. Cues for such narrative interpretations can (but do not have to) be included in the actions (construed broadly) of those participants. In this sense, there may be many narratives unfolding on the basis of the 'same' developing scenario – consolidation and synchronisation of these narratives can only then be an *interactive achievement* precisely as studied in multimodal extensions to conversation analysis (cf. [24, 25]). Again, what appears essential here is the recognition of a variety of discourse mechanisms and how these provide basic tools for 'getting the job of narration done' – even (or rather, especially) when there may be multiple simultaneous narratives being constructed interactively.

Taken together, the conception of discourse discussed here then might make it possible to bridge across different but arguably related domains, such as: narrative, interactive stories, games, since they can all (at least partially) be characterised as instances of discourse. This is suggested graphically in Figure 1. On the left-hand side we see depicted how many discussions in the field have proceeded, with narrative placed at a foundational level. In contrast, on the right-hand side we see the position advocated in this paper. In this view, it is discourse that provides a bridge across different communicative situations and media. Narrative is then just one of the kinds of the interpretations that the appropriate manipulation of discourse expectations and hypotheses can bring about.

Narrative is often considered from its own perspective – and this can no doubt be useful. But linguistically (and now multimodally), it is 'just one' among many forms that can be constructed via discourse. What remains to be done, therefore, is undertaking the considerable body of empirical work necessary in order to ascertain just how various media manage their expressive resources for the creation of such discourse.

—— **References** ———————————————————————————————————

1   Jonne Arjoranta. Narrative tools for games: focalization, granularity, and the mode of narration games. *Games and Culture*, pages 1–22, 2015.

2   Nicholas Asher and Alex Lascarides. *Logics of conversation.* Cambridge University Press, Cambridge, 2003.

3   Byung-Chull Bae, Yun-Gyung Cheong, and R. Michael Young. Toward a computational model of focalization in narrative. In *FDG'11: Proceedings of the 6th International Conference on Foundations of Digital Games*, pages 313–315, New York, NY, 2011. ACM.

4   Roland Barthes. Introduction to the structural analysis of narratives. In Stephen Heath, editor, *Image–Music–Text*, pages 79–124. Fontana, London, 1977 [1966]. originally published in French in *Communications*, 8.

5   John A. Bateman. Towards a *grande paradigmatique* of film: Christian Metz reloaded. *Semiotica*, 167(1/4):13–64, 2007.

6   John A. Bateman. Film and representation: making filmic meaning. In Wolfgang Wildgen and Barend van Heusden, editors, *Metarepresentation, Self-Organization and Art*, European Semiotics, pages 137–162. Lang, Bern, 2009.

7   John A. Bateman. The Decomposability of Semiotic Modes. In Kay L. O'Halloran and Bradley A. Smith, editors, *Multimodal Studies: Multiple Approaches and Domains*, Routledge Studies in Multimodality, pages 17–38. Routledge, London, 2011.

8   John A. Bateman. Methodological and theoretical issues for the empirical investigation of multimodality. In Nina-Maria Klug and Hartmut Stöckl, editors, *Sprache im multimodalen Kontext / Language and Multimodality*, number 7 in Handbooks of Linguistics and Communication Science (HSK), pages 36–74. de Gruyter Mouton, Berlin, 2016.

9   John A. Bateman and Karl-Heinrich Schmidt. *Multimodal Film Analysis: How Films Mean.* Routledge Studies in Multimodality. Routledge, London, 2012.

10   John A. Bateman and Janina Wildfeuer. A multimodal discourse theory of visual narrative. *Journal of Pragmatics*, 74:180–218, 2014. `doi:10.1016/j.pragma.2014.10.001`

11   John A. Bateman and Janina Wildfeuer. Defining units of analysis for the systematic analysis of comics: A discourse-based approach. *Studies in Comics*, 5(2):371–401, 2014. `doi:10.1386/stic.5.2.371_1`

12   Edward Branigan. Point of view in the fiction film. *Wide Angle*, 8(3):4–7, 1986.

13   Hans-Jürgen Bucher. Multimodales Verstehen oder Rezeption als Interaktion. Theoretische und empirische Grundlagen einer systematischen Analyse der Multimodalität. In Hans-Joachim Diekmannshenke, Michael Klemm, and Hartmut Stöckl, editors, *Bildlinguistik. Theorien – Methoden – Fallbeispiele*, pages 123–156. Erich Schmidt, Berlin, 2011.

14   Seymour Chatman. Characters and narrators: filter, center, slant and interest-focus. *Poetics Today*, 7(2):189–204, 1986.

15   Markus Eger, Cámille Barot, and R. Micheal Young. Impulse: a formal characterization of story. In Mark A. Finlayson, Ben Miller, Antonio Lieto, and Remi Ronfard, editors, *Proceedings of 6th Workshop on Computational Models of Narrative (CMN 2015)*, pages 45–53, 2015.

16   James Paul Gee. *Unified discourse analysis: language, reality, virtual worlds, and video games.* Routledge, London and New York, 2015.

17   Irene Heim. File change semantics and the familiarity theory of definiteness. In C. Schwarze and A. von Stechow, editors, *Meaning, Use and Interpretation of Language*, pages 164–178. De Gruyter, Berlin, 1983.

18   David Herman. Toward a transmedial narratology. In Marie-Laure Ryan, editor, *Narrative across Media: The Languages of Storytelling*, pages 47–75. University of Nebraska Press, Lincoln, 2004.

19   Jesper Juul. A clash between game and narrative: a thesis on computer games and interactive fiction. Master's thesis, Institute of Nordic Language and Literature, University of Copenhagen, Copenhagen, February 2001[1999]. translated from Danish by the author. URL: `https://www.jesperjuul.net/thesis/`.

20 Hans Kamp. A Theory of Truth and Semantic Representation. In Jeroen A.G. Groenendijk, T.M.V. Janssen, and Martin B.J. Stokhof, editors, *Formal methods in the study of language. Part 1*, number 136 in Mathematical Centre Tracts, pages 277–322. Mathematisch Centrum Amsterdam, Amsterdam, 1981.

21 Alex Lascarides and N. Asher. Discourse relations and defeasible knowledge. In *Proceedings of the 29th. Annual Meeting of the Association for Computational Linguistics*, pages 55–63, Berkeley, California, 1991. Association for Computational Linguistics. URL: `http://acl.ldc.upenn.edu/P/P91/P91-1008.pdf`.

22 James R. Martin. *English text: systems and structure*. Benjamins, Amsterdam, 1992.

23 Michael Mateas. A preliminary poetics for interactive drama and games. In Noah Wardrip-Fruin and Pat Harrigan, editors, *FirstPerson: New Media as Story, Performance, and Game*, pages 19–34. MIT Press, Cambridge, MA, 2004.

24 Lorenza Mondada. The local constitution of multimodal resources for social interaction. *Journal of Pragmatics*, 65:137–156, 2014.

25 Lorenza Mondada. Multimodal resources and the organization of social interaction. In Andrea Rocci and Louis de Saussure, editors, *Verbal Communication*, pages 329–350. Mouton de Gruyter, 2016.

26 Sigrid Norris. Modal density and modal configurations: multimodal actions. In Carey Jewitt, editor, *The Routledge Handbook of multimodal analysis*, pages 78–90. Routledge, London, 2009.

27 Ken Perlin. Can there be a form between a game and a story? In Noah Wardrip-Fruin and Pat Harrigan, editors, *FirstPerson: New Media as Story, Performance, and Game*, pages 12–18. MIT Press, Cambridge, MA, 2004.

28 James Owen Ryan, Michael Mateas, and Noah Wardrip-Fruin. Open design challenges for interactive emergent narrative. In H. Schoena-Fog, L.E. Bruni, S. Louchart, and S. Baceviciute, editors, *Interactive Storytelling. Proceedings of the 8th International Conference on Interactive Digital Storytelling (ICIDS 2015)*, pages 14–26, Berlin, December 2015. Springer. URL: `https://games.soe.ucsc.edu/sites/default/files/ryanEtAl_OpenDesignChallengesForInteractiveEmergentNarrative.pdf`.

29 Marie-Laure Ryan, editor. *Narrative across Media: The Languages of Storytelling*. University of Nebraska Press, Lincoln, 2004.

30 Marie-Laure Ryan. On the Theoretical Foundations of Transmedial Narratology. In Jan Christoph Meister, Tom Kindt, Wilhelm Schermus, and Malte Stein, editors, *Narrative Beyond Literary Criticism*, pages 1–23. De Gruyter, Berlin, 2005.

31 Marie-Laure Ryan. From narrative games to playable stories towards a poetics of interactive narrative. *StoryWorlds: a journal of narrative studies*, 1:43–59, 2009.

32 Marie-Laure Ryan and Jan-Noël Thon, editors. *Storyworlds across Media. Toward a Media-Conscious Narratology*. University of Nebraska Press, Lincoln and London, 2014.

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

34 Janina Wildfeuer. *Film Discourse Interpretation. Towards a New Paradigm for Multimodal Film Analysis*. Routledge Studies in Multimodality. Routledge, London, New York, 2014.

35 Werner Wolf. 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, 2003.

36 R. Michael Young and Johanna D. Moore. DPOCL: a principled approach to discourse planning. In *Proceedings of the Seventh International Workshop on Natural Language Generation, Kennebunkport, Maine, USA, June 21-24, 1994*, Kennebunkport, Maine, USA, 1994.