Janina Wildfeuer

Discourse Semantics and Textual Logic: Methodological Considerations for Multimodal Analysis

Abstract

This chapter discusses two relevant methodological concepts for the analysis of multimodal media and the resulting artefacts and performances with regard to their usability and appropriateness for comprehensive multimodal analyses: the concept of discourse semantics and the notion of the textual logic of meaning-making. The chapter critically compares and triangulates both notions as essential methodological frameworks for the analysis of multimodal media and particularly the now needed empirical work with them. On the basis of a theoretically and methodologically oriented discussion, a short example analysis completes the chapter, illustrating the suitability of these concepts for larger corpus work.

Keywords: discourse semantics, text, logic, materiality, intersemiosis

1   Introduction

The immense variety of new developments and technological evolutions for communication today bring with them new affordances and challenges for multimodality research, particularly with regard to their ways of manipulating materiality and at the same time transgressing the borders between various sorts of media texts. It is therefore essential to have well-developed and robust frameworks for the analysis of these multimodal orchestrations that allow the precise and critical description of the actual material being used[1], the semiotic modes built from this material and, most importantly, the resulting interpretations of these modes and their intersemiotic interplay.

As one methodological specification for these analyses, the notion of discourse semantics has been introduced to the field as particularly suitable for the combination of both qualitative and quantitative approaches (see Bateman 2016, 2020). Within this context, it has most often been used as a particular configuration of semiotic modes which provides the necessary interpretative

---

[1]   The concept of material and materiality are here understood as "part of a general re-appraisal of the importance of embodiment and engagement with physical objects and their environments for almost all aspects of meaning-making" (Bateman 2021: 35).

mechanisms for determining the intended range of interpretations of these modes (Bateman et al. 2017). As Bateman (2020: 39, 40) highlights, discourse semantics "plays a crucial role in characterizing how […] regularities in material relate to their contextualized interpretations" and "this view of discourse semantics needs to impact directly on all multimodal analyses".

In this chapter, I will argue very similarly to Bateman (2020)'s recent contribution on discourse semantics and emphasize its suitability for an understanding of semiotic modes as well as for the comprehensive analysis of these modes and their interplay. In addition to Bateman's discussion, I will reason for the specific treatment of discourse semantics with a further notion intrinsically combined with this concept: that of the textual logic of meaning-making. The textual logic puts the analytical focus on inferences and logical conclusions to be made when interpreting multimodal artefacts (see Wildfeuer 2014; Wildfeuer 2018). With this focus, I particularly address some of the concerns and critical questions that have been raised in response to the suggestion of discourse semantics as an essential part of semiotic modes. Stöckl, for example, asks:

> "First, is not the semantics of discourse something that would need to be placed at the level of text or semiotic artefact (or better even genre) and concerns both cohesive structures in it and assumed cognitive processes in the recipient? And, secondly, in the light of multimodality, is it perhaps not more adequate to say that discourse semantics operates across and between modes rather than separately within single modes?" (Stöckl 2019: 64)

I will address exactly these questions with a critical discussion of some of the details of logically and textually analyzing multimodal texts and discourses with particular regard to the semiotic modes and their interplay. The aim is to meet Stöckl's concerns of the

> "pertinent task […] to outline what exactly a discourse semantics might be in relation to individual modes […] and where and in which ways a discourse semantics operates across or between the modes" (Stöckl 2019: 64).

I will also confirm and further add to Stöckl's immediate response to the second task in his quote above, when highlighting that "the latter [is] a question adequately addressed already in accounts of multimodal coherence" (Stöckl 2019: 64). With a detailed discussion of how coherence as only one textual quality among several others is an important means of logically analyzing

multimodal meaning, I will zoom in on the mechanisms[2] involved in multimodal meaning construction and concentrate on the logical operations of the multimodal artefact or performance on the level of the text.

For this, an essential understanding of text and discourse is as important as a detailed categorization of the discourse semantics mechanisms that operate in these texts and discourses. For this categorization, I will first look at the origins of discourse semantics in the broader context of semantics and (the mainly anglophone context of) discourse analysis evolving from the 1970s onwards.[3] I will then give a summary of the more contemporary understanding of discourse semantics as a stratum of the semiotic mode. Finally, I will focus on the understanding of the textuality of multimodal artefacts as being intrinsically connected on a logical level, i.e. on the basis of inferences and abductive hypotheses.

2 Discourse Semantics Then and Now

The concept of discourse semantics has taken up crucial roles in several areas of linguistic discourse analysis, both in more formally oriented semantic/discourse theories in the 1970s and 1980s as well as in functionally oriented contexts of systemic-functional linguistics (most explicitly in Martin (1992)'s work) and some evolving applications of these to multimodal artefacts. Both these context as well as their close interrelationships and mutual interferences are described in further detail and with regard to their usefulness for multimodal analyses in Bateman (2020)'s overview. In the following, I will therefore only give a very short and summarizing description of the main aspects of the concept in both traditions and focus on specific aspects that are relevant for my argumentation for the textual logic later.

---

2   As will hopefully become clear, the notion of 'mechanisms' used in this paper is a very broad one including all sorts of dynamic and asymmetric processes. It is by no means to be understood as a pre-defined or pre-destined sequence of actions. Since the term has been used in many other works related to discourse interpretation, I will further apply it to the current discussion.

3   Given the strong tradition in multimodality research of applying definitions and frameworks from the Anglo-American and Australian contexts of discourse analysis that all work with a strong notion of text as a concrete analytical unit (see more details below), discussions from the German-speaking context of discourse analysis that orientate towards Foucault's notion of discourse and often bypass or transgress the analysis of individual texts are mostly excluded from the current discussion.

2.1   Formal and Dynamic Discourse Semantics

The discipline of formal semantics is mainly concerned with the application of logic to semantic analysis in order to formally describe meanings in text and discourses. Meanings are thus depicted by means of formulae which are generated by a set of symbols and through different rules. With the help of the economic and formal tools of logic, the aim is to state the semantics of text in the first place.

Whereas representational approaches within this tradition search for the level of representation of expressions, model-theoretic accounts try to relate elements of models to representations (see the overview in Portner & Partee 2002). As Bateman (2020) emphasizes, it is particularly the development of more dynamic approaches to meaning that drew more attention to units beyond the sentence level and pushed forward the notion of discourse within formal semantics. More focus was for example put on the efforts of a speaker to adapt their utterances to the recipient's knowledge and the respective context, and the notion of a dynamically unfolding, context-dependent interpretation came to the fore (see, e.g., Groenendijk & Stokhof 1991). These and further developments in dynamic semantics initiated a shift from a more traditional logical perspective to more complex operations that include not only the concrete linguistic forms and their semantics, but also the respective context and knowledge sources involved in the dynamic communicative situation (see also Bateman 2020: 43-47). As we have both highlighted several times in many of our (joint) works (e.g. Wildfeuer 2014; Bateman & Wildfeuer 2014; Wildfeuer 2018; Wildfeuer 2019, Bateman 2020), among the most important accounts within this development are the Discourse Representation Theory (DRT; Kamp 1981; Kamp & Reyle 1993) and the Segmented Discourse Representation Theory (SDRT; Asher & Lascarides 2003). I will not describe the details of these theories further, but concentrate on their advantages and a comprehensive application to multimodal artefacts below (see section 3 and 4).

With the aim of uncovering concrete rules and conditions for the interpretation of a discourse (see section 3 for a detailed definition), the notion of inference and especially the operation of abduction as a defeasible consequence relation play an important role within these accounts and the accompanying developments. Abduction goes back to the semiotic theory of Charles Sanders Peirce who distinguishes three classes of arguments which serve for ascertaining truth: deduction, induction, and abduction, whereby the latter "is the process

of forming an explanatory hypothesis. It is the only logical operation which introduces any new idea" (Peirce et al. 1979: §5.171).

These mechanisms of abduction have been suggested to play an important and necessary role in various kinds of media analysis, especially those that deal with dynamic artefacts (see, e.g., Moriarty 1996; Jappy 2013). In several of our own works, we demonstrate in further detail how these mechanisms similarly apply to multimodal artefacts.

Consequently, the abductive, dynamic construction of meaning (in the form of coherence and discourse structure, for example) is to be seen as a basic mechanism of discourse semantics that has been exemplified and tested at its best in the context of formal discourse semantics theories. As I will show in my further argumentation below in sections 3 and 4, it is moreover the formal description and detailed examination of all resources and elements involved in this process of abduction and inferential meaning-making which is a crucial aspect for the application of (formal) discourse semantics to multimodal artefacts.

2.2   Functional Discourse Semantics

The origin of discourse semantics in more functionally oriented approaches to discourse goes back to a need in the description of texts that is very similar to the attempts in formal discourse semantics to incorporate dynamic changes and contextual influences: the need to take into consideration the temporal construction of a text in terms of its logogenesis, the (dynamic) unfolding of meaning (Halliday & Matthiessen 2013: 63).

Martin's work from 1992 therefore proposes "establishing a discourse semantics stratum to complement the metafunctionally organized grammatical descriptions" (Martin 1992: 14) originally identified in the works of Halliday and Hasan (1976) and Halliday (1978). Within their systemic-functional linguistics paradigm, language is a stratified social system with a stratum of the expression plane and a stratum of the content plane or lexicogrammar, the latter then being enriched by the stratum of discourse semantics, as introduced by Martin (see Figure 1).
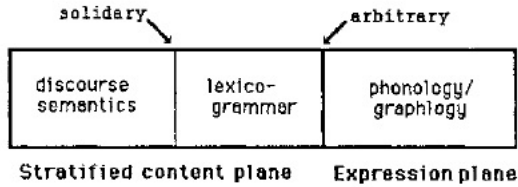
Figure 1: Stratification of the content plane with the addition of discourse semantics according to Martin (1992: 20)

Martin emphasizes:

> "the discourse semantics both generalizes across grammatical resources and accounts for relations between as well as within clause complexes. The discourse semantics is thus more abstract than, and deals with larger units than, lexicogrammar." (Martin 1992: 19)

According to this model, meanings in texts and discourses are realized across these strata, actualized in a text on the level of discourse semantics which again is realized by sentences. Or, as Bateman (2020: 41) summarizes: "Applying this mode of description to texts allowed Martin to claim that a text holds together precisely because of the discourse semantic options that have been taken up."

It becomes clear that with this strong focus on the text as a meaningful and semantic unit, Stöckl's questioning of the localization of discourse semantics (see above) is quite understandable: Discourse semantics in the functional context is mainly emphasized as a particularly *textual* quality, while a concrete distinction between text and discourse remains rather tacit. I will discuss in further detail below how this textual quality relates, on the one hand, to the notion of semiotic mode and, on the other hand, to the intersemiosis of these modes in a medium or genre.

Martin's further classification provides the description of four central discourse systems that constitute the stratum of discourse semantics. Each of these discourse systems is attributed to one of the three metafunctions described within the SFL context, so that "NEGOTIATION considers the discourse semantics of interpersonal meaning, IDENTIFICATION the discourse semantics of textual meaning, CONJUNCTION the discourse semantics of logical meaning, and IDEATION the discourse semantics of experiential meaning" (Martin 1992: 26). In later works, the discourse systems of APPRAISAL (see Martin & Rose 2003) and PERIODICITY (Martin 2019) are introduced.

Martin's concept of discourse semantics and the classification of several discourse systems thus offers a comprehensive approach to meaning with a variety of descriptive parameters that have proven useful both for an extensive language description as well as for the application to multimodal artefacts. While Bateman (2020: 47-48) gives only a short overview, I will elaborate on these applications and the potential for further and particularly empirical analyses in the following section.

2.3   Dimensions of Discourse Semantics Today

While the individual practicality and applicability of both formal and functional accounts to discourse semantics have proven useful and successful, a combination of both perspectives is even more fruitful and offers an integrative approach to multimodal meaning-making. With regard to the details outlined above for each tradition, this integrative approach combines several dimensions on which multimodal analyses have been and are undertaken extensively. I will briefly characterize these dimensions in the following and give a more detailed description and evaluation in terms of the textual logic involved in these dimensions in section 3.

(1) **Multimodal meaning-making can be analyzed in terms of discourse systems/regions or formal discourse representations.** These can be motivated both formally and functionally, as for example in a formal analysis of the semantic representations of events or actions (most developed in DRT) or in the functional description of regions such as identification (cohesion), negotiation, ideation, or conjunction.

(2) **Multimodal meaning-making can be analyzed in terms of discourse relations and discourse coherence.** Both formal as well as functional frameworks make available sets or system networks of these relations and their specific realizations. In formal accounts, coherence is one of the most important principles and the basis for any kind of abductive discourse interpretation. In Martin's functional account, the discourse system of conjunction provides logical relations that hold between discourse units and construct a coherent structure.

(3) **Multimodal meaning-making can therefore always also be analyzed in terms of discourse structures.** The aim of formal discourse analysis is to construct and formally represent these structures as a result of abductive conclusions about discourse relations. The discourse systems introduced

by Martin have also been developed in order to represent the structure of a discourse, even though this structure is mainly described on the basis of individual relations between units and not as an overall overview of the respective structure.

Besides these three main dimensions, there are many more aspects of multi-modal meaning-making that have been discussed with regard to the contribution of a formally or functionally motivated discourse semantics. For instance, multimodal meaning-making can also be described in terms of discourse genres and registers. In particular, functional accounts situate multimodal artefacts in the socio-cultural context and with regard to the social practices in a culture producing specific registers and text/discourse genres. Meaning therefore has always to be interpreted within this specific socio-cultural context.

This interpretation in context not only marks an important feature of discourse semantics in comparison to other types (or levels) of semantics, but is also key to the questions raised by Stöckl and mentioned at the beginning of this contribution. According to the understanding of discourse semantics drafted above, discourse semantics is context-based with context being defined in terms of variables such as age, gender, or ethnicity or as social, political, cultural, or geographical backgrounds. Furthermore and most importantly, discourse semantics is context-based with context being defined as the 'context of an individual text' in which certain conditions influence the semantics and pragmatics, i.e. the basic meaning construction of a specific text, by delivering constraints and consequences for the interpretation of this text.

This contextualization level and the interpretation of discursive elements in context built the strongest motivation for the inclusion of discourse semantics in the definition of a semiotic mode, as Bateman explains:

> "We therefore consider the presence of a discourse semantics stratum to be the hallmark of the semiotic modes 'proper'. Without a discourse semantics, a semiotic mode can only be effective within very particular contexts of use with little possibility of extension […]; the additional stratification provided by a discourse semantics allows configurations to generalize across different contexts by providing guidance schemes for contextual interpretations." (Bateman 2016: 44).

As a consequence, discourse semantics as part of a semiotic mode is located at the interface between the level of discourse on the one hand and the level of text on the other – "precisely because discourse semantics incorporates the crucial operations of textuality" (Bateman 2016: 46). Stöckl (2019: 64) is

thus right in suggesting that discourse semantics needs to be, or is, placed – first and foremost – at the level of the text (see above), showing operations of textuality at work in a particular context. As Bateman furthermore explains:

> "We consider textuality to occur when artefacts and performances provide more or less explicit cues for guiding their own interpretation, both at the more general level of text types and genres [i.e. on the level above a single text, JW] and at the very specific level of how the text unfolds from clause to clause" (Bateman 2016: 46).

Secondly, discourse semantics thus also provides the means to generalize textual operations across more than one particular context – and with this more than one occurrence of a semiotic mode. Stöckl (2019: 64; see above) is therefore also right in saying "that discourse semantics operates across and between modes".

As outlined above, operations of textuality were indeed the basic motivation for developing the specific traits of discourse semantics – and they are the basic motivation for my argumentation in the following section. As I will explain in further detail below, all four dimensions of multimodal meaning-making mentioned above can be located as particularly textual operations and qualities. By giving an overview of these textual operations in terms of what I call the textual logic of discourse semantics, I will explain how discourse semantics operates both across and between modes as well as separately within single modes.

3   Textuality, Text, Discourse(s) and the Textual Logic of Discourse Semantics

According to the definition given above in the quote from Bateman (2016), textuality is a quality that is not simply inherent to a text, but also combines several individual texts to a text type or genre. The principle for this is the specific property of being "structured in order to be interpreted. Distinctions in materials are deployed in order to trigger or support particular lines of discourse interpretation" (Bateman et al. 2017: 132). This means that all operations in a text that give the recipient indications about how to construct meaning out of this text and/or its context can be taken into consideration.

Problems here arise, however, in the combination or collocation of 'text' with the term 'discourse'. Both are very often used interchangeably, especially in the context of multimodality, but also in much broader fields and research

areas; and both terms have in fact been generalized substantially. It is therefore essential to build on well-defined notions of both 'text' and 'discourse' for the context of multimodality. We have provided these in previous works and I will briefly summarize them in the following in order to be able to explain the interrelationships of text, discourse, and discourse semantics properly.

3.1   Definitions of 'Text' and 'Discourse'

As we outline in Bateman et al. (2017: chap. 4.3) and Wildfeuer et al. (2020: chap. 4.3) for a broad view of multimodality as well as in Wildfeuer (2017) for film and Wildfeuer (2018b) for discourse analysis in general, we understand the notions of 'text' and 'discourse' as individual units of analysis that are nevertheless strongly intertwined. Figure 2 visualizes our understanding of the two units and their connections and interrelationships on several levels, following a stratified view of meaning construction in texts and discourses (cf. Halliday & Hasan 1976). The specific idea of stratification as processes of realization and instantiation is shown by the various circles in Figure 2 that are intertwined or embedded within each other.
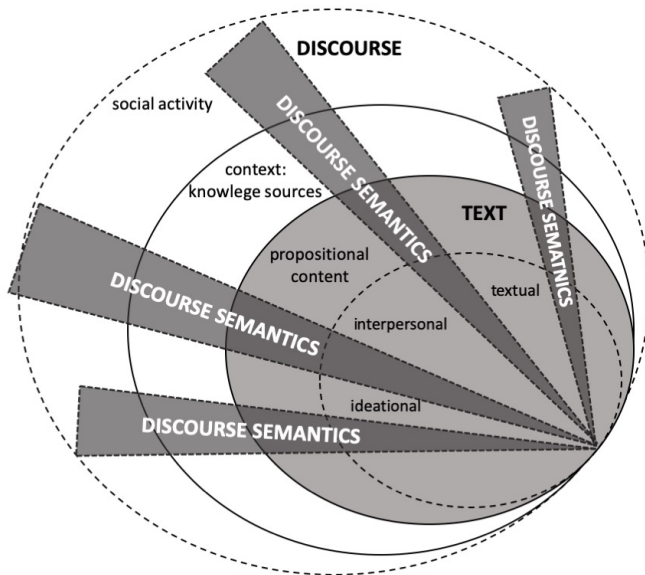

Figure 2: The stratified view of text and discourse including discourse semantics

Following this view, the multimodal text is a single artefact and concrete material object that shares with the traditional linguistic concept of the verbal text the characteristics of being a semantic unit, i.e. carrying meaning, which may be constructed out of diverse verbal and non-verbal semiotic resources. "A text can be defined as a unit that is produced as a result of deploying any semiotic modes that a medium might provide in order to produce a particular and intended structuring of the material" (Bateman et al. 2017: 132). The structuring and arrangement of the material is often described with regard to the respective (ideational, interpersonal or textual) metafunction in the specific context (cf. Halliday 1978). These functions are influenced by the general context as well as different knowledge sources that form the third stratum in Figure 2 and the social activity of the recipients in which the text is embedded (fourth stratum). The view of text and discourse includes a very broad understanding of 'context' that can on the hand be seen as the particular knowledge sources that play a role in the meaning-making process as well as the overall environment in which social communication takes place.[4]

At the same time, this level of 'context' forms the moment of transition from one unit, the text, to the other unit, the discourse, which brings with it notions of abstractness and context-dependency. With this, 'discourse' is represented as realized on all levels of description, as a somewhat superordinate entity which includes the text and cannot be realized without it. The circle representing 'discourse' in Figure 2 is therefore enclosing all other strata but at the same time open for extensions from outside. In contrast to text, 'discourse' is not a material unit, but an abstract concept that is realized by the text and its textuality: "The relationship between text and discourse is thus one of 'realization' and text and textuality refer to the realization of the entire set of semiotic modes in some act of communication as well as the semiotic relations involved" (Bateman et al. 2017: 133). Martin/Rose (2003: 5) describe this principle as "a kind of re-coding" that includes concepts such as 'symbolizing', 'encoding', 'expressing', and 'manifesting'. (Originally verbal) Meanings are therefore always realized across strata, across both levels of discourse and text.

In Bateman et al. (2017) we distinguish further between Discourse (with a 'big D') and discourse (with a 'small d') where the former is the broader

---

4   A complex understanding of contextual relations and with this a more detailed concept of 'context' has, for example, also been provided by Dietrich Busse in his concept of historical discourse semantics (see, e.g., Busse 2000). A combination with the works discussed here is surely a fruitful endeavor, but not the aim of this paper. Thanks go to the anonymous reviewer who pointed out this connection.

context, a 'way of thinking about' something, and the latter is a phenomenon on a local scale, on the level of text. In Wildfeuer (2017) and Wildfeuer et al. (2020), we basically distinguish between text as the material and discourse as the abstract unit.

Multimodal analysis always deals with both single text examples as well as broader discourses and corpora and asks, for example, for socially constructed knowledge and practices in these texts, their contexts, and their discourse(s, with a 'big D'). Meaning construction of these multimodal artefacts is then always an interpretation in context that takes place on the basis of knowledge sources and socio-cultural background information.

As said above and now made visible in Figure 2, discourse semantics forms an overarching principle of the two levels of text and discourse and combines both. It is, on the one hand, a quality that is located on the level of the text, i.e. within the concrete material unit – and thus also within a specific mode, for example. For the interpretation of this material unit, discourse semantics then, on the other hand, operates in connection to the context and the broader level of discourse, i.e. across several texts – or modes.

It is to be noted here that texts and semiotic modes are not to be equalized. While descriptions of discourse semantics in the context of multimodality research have mostly focused on its location *within* or as part of a semiotic mode, i.e. often on a smaller scale than the level of text, a text here is understood as "a result of deploying any semiotic modes that a medium might provide" (Bateman et al. 2017: 132). This can in rare cases be only one individual mode, but in the majority of cases several modes are used together in a text. The individual discourse semantics of each mode can then be manifested in its contribution to the text and its connection to the discourse.

This is exactly where the notion of the textual logic comes into play. Semiotic modes being used in a text construct meaning according to the specific structures of this text or, in other words: the combined discourse semantics of any semiotic modes deployed in a text build the central intrinsic property of a text (see Bateman et al. 2017: 133), i.e. the logic of this text. The specific part of discourse semantics in Figure 2 that is located on the level of the text is thus also the place where the textual logic can be situated.

3.2   The Textual Logic of Multimodal Artefacts and Performances

As indicated above, the textual logic refers to the principles of multimodal meaning-making on the level of the text and is strongly connected to the stratum of discourse semantics in each semiotic mode.

I understand this process, first and foremost, as more or less universal for all multimodal artefacts and performances. Their meaning arises out of the multiple interaction of various modes, their intersemiosis, and this interaction requires the active participation of the recipients in order to comprehend and interpret the artefact or performance. While actively participating, recipients draw inferences from the multimodal content and context due to their world and specific artefact/performance knowledge and further information sources. Understanding and interpreting a multimodal artefact is thus not a matter of simply decoding the semiotic resources, but a process of abductive reasoning and logically concluding the content.

Looking at the inferences drawn by the recipients, the aim is to analytically reconstruct their hypothesis-making by describing and modeling both possible and necessary inferences and the available information sources as well as further contextual circumstances (see Wildfeuer 2019: 219). This makes explicit which inferential steps are taken by the recipients to abductively arrive at the interpretation, and it traces the process of understanding and interpreting a multimodal artefact or performance on a detailed level.

As introduced above, this detailed level can be shown with several mechanisms that build the specific textual qualities that are needed to analyze multimodal meaning. I will now elaborate on these mechanisms, i.e. the textual qualities as listed in section 2.3, with regard to some of the approaches and frameworks that have been developed in the context of multimodality research. I will then explain more specifically how the meaning-making process for each of these mechanisms can be reconstructed for and within a specific example in section 4.

3.2.1   Discourse Representations and Discourse Systems/Regions

Multimodal meaning-making can be analyzed in terms of discourse systems/regions or formal discourse representations.

In order to analyze the semantic content of multimodal artefacts and to reconstruct participants/characters, settings, etc., multimodal artefacts have successfully been analyzed in terms of several of the discourse systems provided by systemic-functional discourse semantics. Chiao-I Tseng, for instance, has

developed and adjusted the functional discourse region of IDENTIFICATION to film as well as comics by looking at cohesion as "one significant dimension within the dynamic discourse semantic model developed in SFL" (Tseng 2013: 9). Tseng exactly highlights in her work what has been outlined above as the main principle of discourse semantics:

> "The analysis of discourse semantics encompasses viewers' dynamic comprehension activity and describes how meanings can be assigned to dynamically unfolding film without resorting to text grammar or syntax. The analysis is conducted with moment-by-moment contingency. This dynamism echoes precisely the dynamic view held by film theorists who argue for investigating active aspects of art and viewers' inferences." (Tseng 2013: 9)

By focusing on moment-by-moment contingency, constructing cohesive identity chains and building more abstract metaphorical structures of action patterns, the analysis of multimodal cohesion provides a method of co-patterning actions, characters, objects and settings on the textual level of a multimodal artefact, i.e. by looking at the concrete material units that are combined. With this, Tseng's work is one of the strongest applications of the discourse system of IDENTIFICATION to multimodality and has been further developed and adjusted for empirical analysis.

The identification of characters and settings also plays an important role in formally oriented approaches to the construction of discourse representations, mostly based on DRT (Kamp & Reyle 1993, see above), for example. While my own work includes the construction of so-called logical forms of discourse, i.e. discourse segments that list represented participants/characters, objects, specific settings as well as technical features in order to explain the process of inferring the semantic content (see Wildfeuer 2014 for film; Wildfeuer 2019 for comics), the work by Dorit Abusch (2012) similarly provides a (DRT) formalism for indexing (or co-referencing) identities and settings in pictures and comic panels. Likewise, Emar Maier (2019) is developing an extension of discourse semantics approaches that were originally designed for verbal discourse analysis to the analysis of pictorial meaning. Similar attempts can be found in Abusch & Rooth (2017), Maier & Bimpikou (2019) and Fernando (2020). All these works aim at the formal representation of specific aspects of multimodal meaning, mostly focusing on characters or specific technical features of a medium (such as a speech bubble in comics, for example). So far, most of them have been used in qualitative analyses and more empirical work has yet to come.

Further discourse systems that have been extensively discussed and applied to multimodal artefacts are the systems of IDEATION, CONJUNCTION, and APPRAISAL. In contrast to Tseng's comprehensive framework of cohesion/ IDENTIFICATION for the analysis of several narrative artefacts, works dealing with these other systems often focus on more individual aspects of these regions and analyze either individual aspects (such as, for example, taxonomies on the level of ideation) or a combination of several of these regions. Often, these approaches do not first and foremost aim at providing a basis for empirical analysis.

CONJUNCTION is perhaps the discourse system that has experienced most attention and development for multimodal artefacts. Early work has already started in the 1990s with Theo van Leeuwen's analysis of conjunction in film (Van Leeuwen 1991) and has until now resulted in many fundamental accounts of discourse relations holding between all sorts of semiotic modes (e.g., among many others, Martinec & Salway 2005; Liu & O'Halloran 2009; O'Halloran et al. 2016). The main aim of all these accounts is to make explicit the inferences between units in a multimodal text or performance by identifying relationships holding between these units.

3.2.2   Discourse Relations, Discourse Coherence and Discourse Structure

Multimodal meaning-making can be analyzed in terms of discourse relations and discourse coherence. Multimodal meaning-making can therefore always also be analyzed in terms of discourse structures.

Functional accounts to discourse relations mostly operate under the system of conjunction and provide extensive classification systems for these relations. Interestingly, they often do not describe the unfolding discourse structure of the overall artefact that is generally resulting from these relations, even though, as Bateman (2020: 42) outlines, "each of the regions addressed [in the discourse semantics as developed by Martin (1992), J.W.] provides a different perspective on 'discourse structure'. In formal accounts, in contrast, this discourse structure and the resulting overarching coherence is more in the foreground of description and often also the starting point of discussion and analysis, being the basis for any kind of abductive discourse interpretation.

In order to make this abductive interpretation more explicit and to demonstrate the necessary steps of inferring relationships between entities, we have pushed the formal reconstruction of multimodal discourse structures in our own works with a particular focus on the dynamic unfolding of these structures and the challenges they bring with them during the interpretation

(see Wildfeuer 2014; Bateman & Wildfeuer 2014). These works follow the basic notion of coherence in formal discourse semantics, formulated as the principle of Maximize Discourse Coherence (MDC): "The logical form for a discourse [the structure, J.W.] is always a logical form that's maximal in the partial order of the possible interpretations (Asher & Lascarides 2003: 21).

An overview of the broad context of multimodal coherence research resulting from the different perspectives mentioned is given in Bateman (2014) in which the author adds to the modeling of multimodal relations on the basis of discourse semantics also directions such as grammar, rhetorics, as well interaction and action theories. He summarizes: "In general, relationships can only be found by an active act of interpretation on the part of the reader/ viewer (who may even turn out to be wrong). This is a very general property of the *discourse* stratum of abstraction" (Bateman 2014: 170).

While putting the focus back on discourse semantics, Bateman then also explains:

> "Rather more detailed descriptions will need to be pursued of the fine-grained internal structure, layout and modal contributions of multimodal artefacts. This is necessary in order to support empirical research and practical application." (Bateman 2014: 171)

Since the time of Bateman's publication, some progress has been made in analyzing the structure of multimodal artefacts (see, e.g., Thomas 2014; Hiippala 2015). As Hiippala (2015: 6) outlines: "Describing the structure of a multimodal artefact therefore makes a valuable theoretical contribution by mapping the ground between the concepts of mode and genre." The analysis of discourse coherence and discourse structure thus situates itself at the threshold of the textual and the discourse level of interpretation.

While coherence itself has often been seen more as a text-external quality and definitely not as an objective property of the produced text (see, e.g., Givón 1993), it is nevertheless the meaning-making units that drive the construction of coherence and the unfolding discourse structures. The starting point of any coherence relation to be inferred by the recipient is thus still in the text and not fully based on context-related details. The logical processes originate in the text and can be seen as being textual; the textual logic of inferring coherence and structure therefore deals with the description and analysis of the semiotic units at play and their particular interaction. The example analysis in the following section will demonstrate these textual processes in further detail.

4   Bringing It All Together: Discourse Semantics at Work

Figure 3: Choking Poster by Lara Antal
(Source: https://lara-antal.com/Infographics)

As Bateman (2016: 44) puts it in a nutshell, discourse semantics "allows semiotic configurations to generalize across different contexts by providing guidance schemes for contextual interpretations". These guiding schemes can be manifested in the multimodal text by looking at the logical processes of meaning-construction on the basis of abductive inferences. I will now demonstrate this manifestation in terms of the textual logic with an example analysis of the multimodal artefact that is reproduced in Figure 3.

This artefact represents a highly complex multimodal document that combines typical patterns from both information posters as well as comic pages. On the one hand, this artefact uses a graphic novel or comic style with a typical 3x3 grid structure, captions, and speech bubbles that together form a very typical comic page. On the other hand, the document follows the style of multimodal instructions that are usually designed in the form of posters and that "include pictures and text and instruct users to perform procedural tasks" (van der Sluis et al. 2017), in this case the so-called Heimlich maneuver to save someone from choking.

This specific page, or 'instruction comic', was designed by Lara Antal in order to replace the official "Choking First Aid poster" provided by New York city's health department[5], following an initiative by several artists in the mid 2010s. Since it shows interesting features of different media, it represents an interesting hybrid form that combines formal and functional patterns from several other narrative and instructional genres.

Applying a discourse semantics approach to this multimodal artefact, both characters, the setting, and typical features in terms of the discourse representations available for this page as well as the unfolding discourse structure in terms of relations holding between the representations can be analyzed. This will make visible how patterns of the textual logic in this artefact guide the recipients' interpretation in understanding this page.

4.1   Discourse Representations and Identification in the Instruction Comic

One way to describe discourse representations of the various units and segments of this page is to apply the formal approach of the so-called logical of multimodal discourse interpretation, which I have been developing in several accounts for the analysis of artefacts such as films, comics, or video games (Wildfeuer 2014; Wildfeuer 2019; Wildfeuer & Stamenković forthcoming).

---

5   The official poster can be found here: https://www1.nyc.gov/site/doh/business/food-op-erators/required-signs.page

Using the formal notation style provided in this framework, the analysis produces so-called logical forms of the discourse for which a specific example is given in Figure 4.
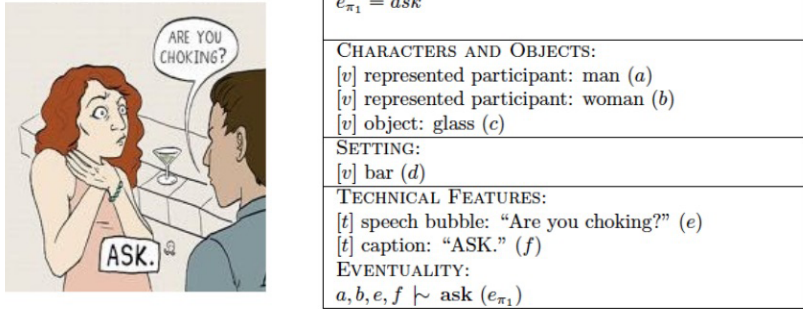

Figure 4: Logical form of the first panel in Lara Antal's choking poster/instruction comic

The logical form on the right in Figure 4 shows an abstract formulation of what is happening in the panel shown on the left. The first line gives the inferred 'verbalization' for the content identified in the panel: ask. This is a semantic description that represents a hypothesis about the recipients' inference process in making sense of the pictorial and verbal resources. The description is based on the list of elements identified in the second part of the box, listing characters and objects, the setting, and specific technical features. This shows exactly which material units in the text are identified and directly attributes specific functions to the semiotic elements involved in the meaning-making process. The final line of the semantic representation in the box makes explicit which of the referents and features listed in the upper part of the box are sources of evidence for the interpretation of the eventuality. The logical operator given in the last line indicates that this inference is a defeasible consequence relation drawn on the basis of the recipients' world knowledge (see for further details Wildfeuer 2019: 224-226). The final line therefore makes visible how recipients normally – that is, on the basis of abductive reasoning – interprets the combination of resources: as the event of asking "Are you choking?".

It would now be possible to construct such discourse representations, i.e. logical forms, for each panel of the page and to outline which particular semiotic elements in these panels guide the recipients in their interpretation. For this, it would for example be important to list the arrows that are used

in some of the subsequent panels in order to show the procedural actions of processing the Heimlich maneuver.

By constructing these discourse representations, it is also possible to identify the cohesive chains that are established for the most dominant elements of this page: for the two characters, but also for the setting and the invisible object in the character's throat that becomes visible later on. By following Tseng's framework for comics (Tseng et al. 2019), it then becomes visible that the identity of the victim is already introduced verbally as "person" in the caption text on top of the page. The same identity is presumed/tracked pictorially in the first panel, and the cohesive chain for this character starts to build across these two elements and these two distinct semiotic modes, the written language and the image. The identity can further be tracked throughout the page by the visual re-appearances in all other panels which creates further cross-modal references, sometimes also only to parts of the identity's body, as well as by verbal references as "victim". Similar cohesive chains can also be constructed for the second identity that is represented only pictorially as well as for the setting that is shown in the first and last panel.

Both the construction of discourse representations as well as the use of the discourse system of identification as part of the meaning-making processes for this poster give particular details of the textual manifestations of these constructions. They thus show the textual logic at work by identifying the material units that guide the recipients to the identification of characters as well as the reconstruction of the basic narrative and instructional events/actions.

Processes of discourse semantics beyond the textual level mainly operate with regard to the context and genre of this artefact. The identification of speech bubbles and captions and the interpretation of their specific functions in the meaning-making process is for example based on the recipients' knowledge of the typical design and formal features of comics pages, while the function of the page as an instruction for the particular first-aid task will perhaps only be realized and interpreted in the respective environment in which the page will most likely be perceived, namely in a bar or restaurant in New York.

By looking at these textual details with a discourse semantics perspective, it is then also an interesting question for multimodality research to find out more about the specific functions of these speech bubbles and captions. From a corpus analysis of the layout of several of these posters (see Wildfeuer et al. forthcoming), it became clear that the specific task of asking the victim whether he or she is choking is almost always given in a rather extraposed caption in one of the first panels of the page. This is also the case for the specific task of

calling 911, which is here given in the title-like caption on top of the page. A more systematic corpus analysis with a particular view of the functions of these design features could then perhaps identify these features as particular semiotic modes of this specific genre of instructions comics according to their own discourse semantics.

For this, it is also necessary to think about further interpretative mechanisms that guide the recipients to a particular understanding and interpretation. One of these mechanisms is the analysis of discourse relations holding between these features and other features of the page. It could for example be hypothesized that these captions that instruct the reader to ask the victim are always embedded into a larger structure via an Elaboration relation that gives further details of the processes shown in the images and additionally verbalized by a speech bubble.

4.2   Coherence and Discourse Structure of the Instruction Comic

This embedding of the caption can best be analyzed with regard to the discourse relations holding between the various discourse segments/representations analyzed in the first step in section 4.1 and by describing the unfolding discourse structure of the page. Without going into further details of the set of discourse relations that we provide for comics and graphic novels (see Bateman & Wildfeuer 2014), Figure 5 presents the overall structure of the page and shows that the two-part structure of the instruction itself is in fact embedded in a narrative structure that frames the whole comic. The first and the last panel show a narrative setting in a bar in which the two characters obviously have a drink together and continue to do so after the life-saving procedure.

The instruction itself is unfolding within two substructures to this narrative structure that develop in parallel, according to the two conditions introduced on the page by two further extraposed captions: "If the person is standing" and "If the person has collapsed". This complex structure has to be recognized by the recipients by inferring a Contrast relation between the two conditions, thus by identifying that these panels do not follow each other in a temporal sequence. The typical table grid structure is indeed not giving any further indications for this besides the two captions; instead, the contrast has to be analyzed between the semantic content of the verbal captions and the semantic content of the respective panels. The two structures then resolve in a panel that is accompanied by the extraposed caption "Thereafter" and then followed again by a further panel that is depicted on another layer. The use

of this layering effect is indeed a typical feature of recent comics pages (see Bateman et al. 2019) and thus also a specific discourse knowledge that comes into play in the interpretation process.

A similar analysis in terms of discourse relations holding between the discourse segments and with regard to dependencies in the overall structure would be possible with an application of other structural accounts such as, for example, Rhetorical Structure Theory (Mann & Thompson 1987, Taboada & Mann 2006) or by analyzing several layers of the GeM model (Bateman 2008; Hiippala 2015).

Similar to the first step of the analysis as shown in section 4.1, the examination of the discourse relations and with this the unfolding discourse structure of the page, which at the same time shows its coherence, has identified textual details and logical processes that are necessary for the understanding and interpretation of the page. Mechanisms of discourse semantics that are situated on the broader level of the discourse context and in combination with specific knowledge sources have also been identified. This provides the basis for further empirical analyses of all examples in the corpus of instruction comics. It will then be interesting to find out whether all pages use such a complex discourse structure in which the instruction is embedded in a narrative or whether other structures come to the fore. Further details to be addressed could be the arrangement and design of the two conditions for this specific instruction or the relation between the content of the instruction and additional information as provided in title captions or other parts of the page.
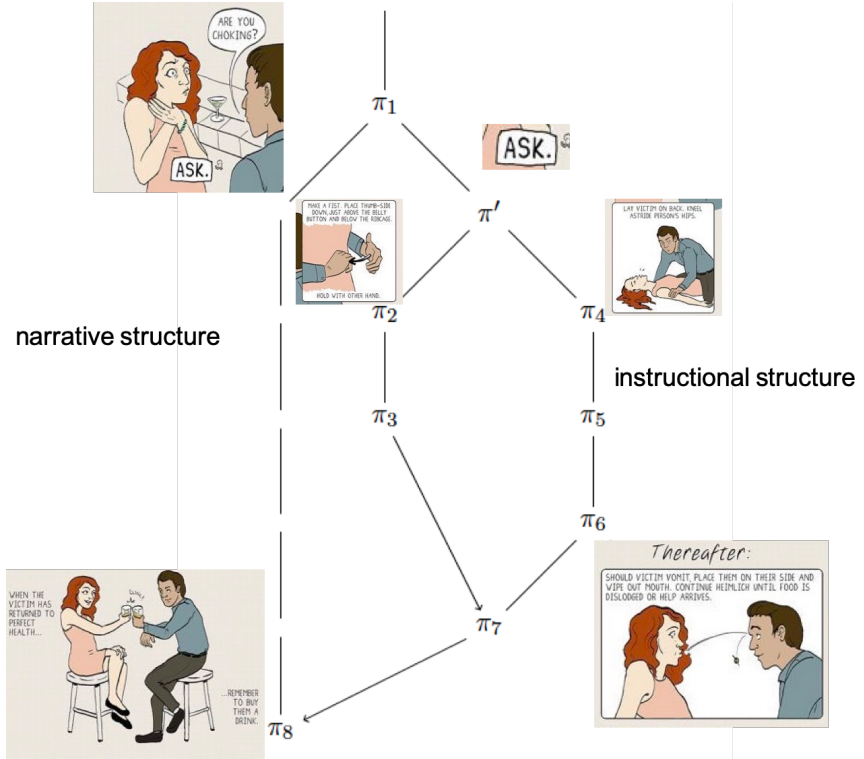
Figure 5: Discourse structure of the choking poster/instruction comic by Lara Antal

5   Conclusion

In this chapter, I have shown how the methodological concepts of discourse semantics and textual logic can be triangulated effectively in order to analyze meaning-making processes in multimodal artefacts. By focusing on the inference processes of the recipients and describing the textual logic that guides these inference processes, I have tried to demonstrate the interwovenness of the two concepts with their own individual specificities and their suitability for several levels of multimodal analysis. I have also addressed critical questions raised with regard to the concept of discourse semantics in multimodality research and I have shown that a better understanding of the notions of 'text'

and 'discourse' help manifesting this concept as an essential methodological framework for the analysis of multimodal media.

On the basis of the short qualitative analysis and the adaptability of the frameworks used for this analysis, it will now be possible to pursue larger empirical analysis with corpora of diverse multimodal artefacts. While the corpus of instruction comics only includes around 40 examples, the broad applicability of the discourse semantics framework will allow the analysis of discourse representations and discourse structures of various static as well as dynamic artefacts. This will allow not only a precise description of the materiality at play in these artefacts, but also a comprehensive analysis of the semiotic modes built from this material and each including a specific discourse semantics.

References

Abusch, Dorit (2012): Applying Discourse Semantics and Pragmatics to Co-Reference in Picture Sequences. Proceedings of Sinn und Bedeutung 17. URL: https://hdl.handle.net/1813/30598.

Abusch, Dorit; Roth, Mats (2017): The Formal Semantics of Free Perception in Pictorial Narratives. Proceedings of the 21st Amsterdam Colloquium.

Asher, Nicolas; Lascarides, Alex (2003): Logics of Conversation. Cambridge: Cambridge University Press.

Bateman, John A. (2008): Multimodality and Genre. A Foundation for the Systematic Analysis of Multimodal Documents. Hampshire: Macmillan Palgrave.

Bateman, John A. (2014): Multimodal Coherence Research and its Applications. In: Gruber, Helmut; Redeker, Gisela (eds.): The Pragmatics of Discourse Coherence. Amsterdam: John Benjamins, 145-177.

Bateman, John A. (2016): Methodological and Theoretical Issues for the Empirical Investigation of Multimodality. In: Klug, Nina-Maria; Stöckl, Hartmut (eds.): *Handbuch Sprache im multimodalen Kontext*. Berlin: De Gruyter, 36-74.

Bateman, John A. (2020): The Foundational Role of Discourse Semantics beyond Language. In: Zappavigna, Michele; Dreyfus, Shoshana (eds.): Discourses of Hope and Reconciliation. On J. R. Martin's Contribution to Systemic Functional Linguistics. London: Bloomsbury, 39-55.

Bateman, John A. (2021): Dimensions of Materiality. In: Pflaeging, Jana; Wildfeuer, Janina; Bateman, John (eds.): Empirical Multimodality Research: Methods, Evaluations, Implications. Berlin: de Gruyter, 35-63.

Bateman, John A.; Wildfeuer, Janina (2014): A Multimodal Discourse Theory of Visual Narrative. Journal of Pragmatics 74: 180-218.

Bateman, John A.; Wildfeuer, Janina; Hiippala, Tuomo (2017): Multimodality. Foundations, Research, Analysis. A Problem-Oriented Introduction. Berlin: de Gruyter.

Bateman, John A.; Veloso, Francisco O.D.; Lau, Yan Ling (2019): On the Track of Visual Style: A Diachronic Study of Page Composition in Comics and its Functional Motivation. Visual Communication 20, 2. DOI: https://doi.org/10.1177/1470357219839101.

Busse, Dietrich (2000): Historische Diskurssemantik. Ein linguistischer Beitrag zur Analyse gesellschaftlichen Wissens. In: Stuckenbrock, Anja; Scharloth, Joachim (eds.): Linguistische Diskursgeschichte. Sprache und Literatur in Wissenschaft und Unterricht 31, Volume 86, 39-53.

Fernando, Tim (2020): Pictorial Narratives and Temporal Refinement. Semantics and Linguistic Theory (SALT) 29.

Givón, Tomas (1993): Coherence in Text, Coherence in Mind. Pragmatics & Cognition 1,2, 171-227.

Groenendijk, Jeroen; Stokhof, Martin (1991): Dynamic Predicate Logic. Linguistics and Philosophy 14, 39-100.

Halliday, Michael A.K. (1978): Language as Social Semiotic: The Social Interpretation of Language and Meaning. London: Arnold.

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

Hiippala, Tuomo (2015): The Structure of Multimodal Documents: An Empirical Approach. London: Routledge.

Jappy, Tony (2013): Introduction to Peircean Visual Semiotics. London, New York: Bloomsbury.

Kamp, Hans (1981): A Theory of Truth and Semantic Representation. In: Groenendijk, Jeroen; Janssen, T.M.V.; Stokhof, Martin (eds.): Formal Methods in the Study of Language Part 1. Amsterdam: Mathematisch Centrum, 277-322.

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

Maier, Emar (2019): Picturing Words: The Semantics of Speech Balloons. Proceedings of the 22nd Amsterdam Colloquium. URL: https://ling.auf.net/lingbuzz/004938.

Maier, Emar; Bimpikou, Sophia (2019): Shifting Perspectives in Pictorial Narratives. Sinn und Bedeutung 23.

Mann, William C.; Thompson, Sandra (1987): Rhetorical Structure Theory: A Theory of Text Organization. Text 8,3, 243-281.

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

Martin, James R. (2019): Discourse Semantics. In: Thompson, G.; Bowcher, W.L.; Fontaine, L.; Schöntal, D. (eds.): The Cambridge Handbook of Systemic Functional Linguistics. Cambridge: Cambridge University Press, 358-381.

Martin, James R.; Rose, David (2003): Working with Discourse: Meaning beyond the Clause. London: Continuum.

Martinec, Radan; Salway, Anthony (2005): A System for Image-Text Relations in New and Old Media. Visual Communication 4,3, 337-371.

Moriarty, Sandra E. (1996): Abduction: A Theory of Visual Interpretation. Communication Theory 6, 2, 167-187.

O'Halloran, Kay; Tan, Sabine; Wignell, Peter; Bateman, John A.; Pham, Duc-Son; Grossman, Michele; Vande Moere, Andrew (2016): Interpreting Text and Image Relations in Violent Extremist Discourse: A Mixed Methods Approach for Big Data Analytics. Terrorism and Political Violence 31,3, 454-474.

Peirce, Charles S.; Hartshorne, Charles; Weiss, Paul (1979): Collected Papers of Charles Sanders Peirce. Cambridge: Belknap Press of Harvard University Press.

Portner, Paul; Partee, Barbara (2002): Formal Semantics. The Essential Readings. Oxford: Blackwell Publishing.

Stöckl, Hartmut (2019): Linguistic Multimodality - Multimodal Linguistics: A State-of-the-Art Sketch. In: Wildfeuer, Janina; Pflaeging, Jana; Bateman, John A.; Seizov, Ognyan; Tseng, Chiao-I (eds.): Multimodality. Disciplinary Thoughts and the Challenge of Diversity. Berlin: de Gruyter, 41-68.

Taboada, Maite; Mann, William C. (2006): Applications of Rhetorical Structure Theory. Discourse Studies 8,4, 567-588.

Thomas, Martin (2014): Evidence and Circularity in Multimodal Discourse Analysis. Visual Communication 13,2, 163-189.

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

Tseng, Chiao-I; Laubrock, Jochen; Pflaeging, Jana (2018): Character Developments in Comics and Graphic Novels: A Systematic Analytical Scheme. In: Dunst, Alexander; Laubrock, Jochen; Wildfeuer, Janina (eds.): Empirical Comics Research. Digital, Multimodal, and Cognitive Methods. London: Routledge, 154-175.

van der Sluis, Ielka; Eppinga; Anne N.; Redeker, Gisela (2017): Text-Picture Relations in Multimodal Instructions. In: Asher, Nicholas (ed.): Proceedings of the

Workshop on Foundations of Situated or Multimodal Communication (FMSC) at IWCS 2017.

van Leeuwen, Theo (1991): Conjunctive Structure in Documentary Film and Television. Continuum: Journal of Media and Cultural Studies 5,1, 76-114.

Wildfeuer, Janina (2014): Film Discourse Interpretation. Towards a New Paradigm for Multimodal Analysis. London: Routledge.

Wildfeuer, Janina (2017): From Text to Recipient: Pragmatic Insights for Filmic Meaning Construction. In: Wildfeuer, Janina; Bateman, John A. (eds.): Film Text Analysis. New Perspectives on the Analysis of Filmic Meaning. London: Routledge, 118-140.

Wildfeuer, Janina (2018a): It's all about Logics?! Analyzing the Rhetorical Structure of Multimodal Filmic Text. Semiotica 220, 95-121.

Wildfeuer, Janina (2018b): Diskurslinguistik und Text. In: Warnke, Ingo H. (ed.) Handbuch Diskurs. Berlin, de Gruyter 134-151.

Wildfeuer, Janina (2019): The Inferential Semantics of Comics: Panels and their Meaning. Poetics Today 40,2: 215-234.

Wildfeuer, Janina; Bateman, John A.; Hiippala, Tuomo (2020): Multimodalität. Grundlagen, Forschung, Analyse. Eine problemorientierte Einführung. Berlin: de Gruyter.

Wildfeuer, Janina; Stamenković, Dušan (forthcoming): A Multimodal Discourse Theory of Video Games: Analyzing the Discourse Structure of Tutorials. Under Review.

Wildfeuer, Janina; van der Sluis, Ielka; Redeker, Gisela (forthcoming): No Laughing Matter!? Analyzing the Page Layout of Instruction Comics.