Perspective blending in graphic media

Bimpikou, Sofia

Perspective blending in graphic media*

Sofia Bimpikou

University of Groningen, Netherlands
s.bimpikou@rug.nl

Abstract. This paper discusses the representation of perceptual events in comics. I present "blended" pictures in which the experiencing character and her non-veridical perception are both represented from an external perspective in a single image. Inspired by Abusch & Rooth's (2017) analysis of free perception sequences and their modelling of veridical and non-veridical perception, I develop some proposals to model interpretation of non-veridical perception in blended pictures. I also discuss to what extent blended-perspective pictures are parallel to free indirect discourse in literature.

Keywords: picture semantics · pictorial narrative · perception · mental states · perspective shifting

1 Introduction

Is it possible for someone to have full access to another individual's perceptual experience? This may sound as fiction in real life, but in fiction, it is a common phenomenon. The pictures below, taken from Grant Morrison's comic book *Joe the Barbarian* and the cartoon *BoJack Horseman* illustrate this:

Fig. 1. (a) Image from graphic novel *Joe the Barbarian*, by Grant Morrison. (b) Snapshot from *Bojack Horseman* animated series, season 1 episode 11 *Downer Ending*.

In figure (1a), we see a character surrounded by his own hallucinations: here, Joe's toys have come to life. But the character is also depicted, which implicates

* This research is part of Emar Maier's research project *The Language of Fiction and Imagination* supported by NWO Vidi Grant 276-80-004.

that a narrator can 'see' both the character and his hallucinations from some external position. Similarly, in (1b), the protagonist, Bojack, depicted on the left of the picture, is under the influence of drugs and sees himself in the mirror as a real horse. We can simultaneously see Bojack and the image of himself in the mirror as he perceives it. I will call such instances "blended-perspective" or simply "blended" pictures, as they allude to the narrator's geometrical perspective from which the scene is projected, but also to the character's internal reality or epistemic perspective.[1]

In this paper, I present some proposals to deal with blended pictures in visual narratives by using tools from semantic theories applied to language. In section 2, I give an overview of previous work on semantics of pictures; in section 3, I discuss three ways to analyse blended pictures; finally, in section 4, I conclude and point out some issues for further research.

2 Background

Pictures convey their own content. This has led to the development of semantic analyses of pictures in recent work (see Greenberg 2011, Abusch 2012, 2015, Abusch & Rooth 2017). These approaches extend the possible worlds semantic framework to pictures, based on the idea that, like sentences in language, pictures express propositions, i.e. sets of worlds. Greenberg (2011) suggests that the content of a picture is the set of scenes (i.e. worlds at a time and a location) the picture accurately depicts. Accordingly, a picture is an accurate depiction of a scene if it can be derived from the scene by specific rules of geometrical projection.

In general, geometric projection is defined as $\pi\{w, v, l, M\} = p$, meaning that a world w (at time t)[2] is projected to picture p from viewpoint v given a marking rule M and a line projection rule l.[3] The main idea is that the content of a picture p is a set of worlds relative to a geometrical viewpoint. In other words, the content of a picture p is a set of pairs consisting of a world and a viewpoint:

$$[\![p]\!]^{M,l} = \{\langle w, v\rangle | \pi(w, v, l, M) = p\} \quad (1)$$

Before moving to their analysis of looking events, I will say a few words about how discourse referents are identified in pictures. I follow Abusch's (2012) analy-

[1] Note that the term "blending" as used here is different from the notion of "blending" as used in cognitive semantics literature.

[2] The time parameter t is neglected in the formalisations, but "world" will refer to a "world at a time".

[3] M and l are parameters of geometrical projection. More specifically, l determines the projection lines from a viewpoint v towards a scene and M determines how points in the picture plane are to be marked with respect to the projection lines and the scene. For more details, see Greenberg (2011), Abusch (2015). Except for formula (1), I will not include these parameters in the following formalisations in order to keep things simple.

sis of co-reference across panels. Abusch uses a Discourse Representation Theory framework to model co-reference of individuals in picture sequences. The reader distinguishes certain areas in the picture that correspond to the story characters. Discourse referents are made out of these areas. Identity relations between discourse referents across pictures are then formalised as identity predications between these areas. Abusch suggests that co-reference is done at a post-semantic level, that is, identity between areas in pictures is determined pragmatically and is not part of the literal content of a picture.

Within a geometrical projection framework, Abusch & Rooth (2017) analyse free perception picture sequences. Their analysis is directly relevant for the data presented here. A free perception sequence (p, q) is a sequence in which one picture (p) depicts a character looking at a scene and the other one (q) - the free perception panel - depicts the scene looked at, as if directly through the character's eyes. Figure (2) is an illustration of a free perception sequence:

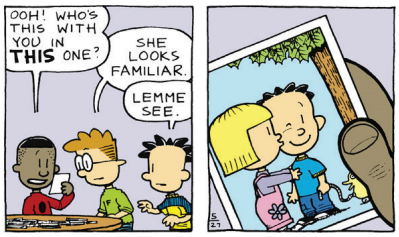

Fig. 2. A free perception sequence. (taken from *Big Nate*, by Lincoln Peirce)

A crucial distinction is that between veridical and non-veridical looking events. In the case of veridical looking events, the free perception panel has an extensional interpretation: it shows what the "base world" looks like from the character's geometric perspective, as in the picture above.[4] On the other hand, in the case of non-veridical perceptual events, the free perception panel depicts what the character sees but that may not correspond to how the base worlds looks like. This happens in cases of misperception, for example, when the protagonist hallucinates. This is shown in figure (3). Bart is looking at a jar which is actually empty, but what he sees instead is a dead fairy.

[4] I assume that the "base world" corresponds to the objective representation of the fictional world by the reliable narrator. In pictorial narratives, as opposed to language, there is not always a verbal narrator. In the absence of a verbal narrator, I assume an impersonal narrator or a camera eye. How the notion of the narrator in visual media should be construed is a matter of debate in narratological studies. I will not take a stance on this debate here, but no matter how a narrator is construed, I take the default visual observer to be a reliable one. Therefore, the base world is the representation of the fictional world according to that reliable observer.

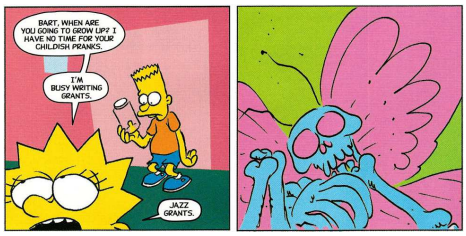

Fig. 3. A non-veridical free perception sequence. (taken from *Bart Simpson's Treehouse of Horror*, by Kyle Baker)

The authors propose covert syntactic embedding for free perception panels, inspired by natural language embedding structures.[5] More specifically, they propose syntactic embedding of the free perception panel under a covert operator P, a *see*-predicate. This is illustrated in figure (4). The authors propose the following logical form for embedding cases (here simplified):

$$w, v, O \models p\, d\, [1[P\, q]] \tag{2}$$

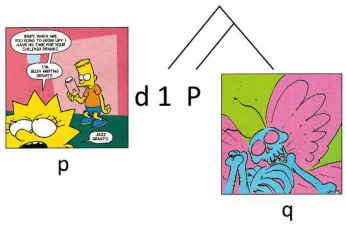

Fig. 4. An embedding structure for a free perception sequence (p, q).

Abusch & Rooth use a dynamic semantics framework. The above formula expresses a satisfaction clause where the tuple on the left of the turnstile satisfies the sequence (p, q) on the right. w is a world, viewpoint v is the viewpoint for the last picture of the sequence and O is a sequence of individuals onto which the discourse referents are mapped. On the right side of the turnstile, p and q are the two pictures and d is a discourse referent with index 1 introduced in picture p. According to the intensional LF (2), q is syntactically embedded under the covert operator P that takes the index introduced by the discourse referent as its subject. This means that w looks like q from agent $d1$'s perspective, but the base world may or may not look like q.

[5] They also suggest that extensional free perception panels can be analysed as top-level conjuncts but this option will not be discussed here.

Note that this LF also allows for an extensional interpretation, as is the case in natural language embedding structures with verbs like *see* or *believe*. To capture ambiguity, the authors distinguish between veridical and non-veridical looking events: $l(x, q)$ and $m(x, q)$ respectively, that both translate into "x looks at a scene that projects to picture q from x's perspective", but the difference between them is that m has a precondition that the base world does not actually look like q, only x sees it as q. Roughly, the idea is that an embedding structure entails that the agent has looked and that w looks like q from his perspective, but the base world might either look like q (therefore the world ends up with a veridical looking event l) or it may not look like q (hence the world ends up with a non-veridical looking event m).

Overall, the authors offer a neat proposal in order to allow for both extensional and intensional interpretations in free perception sequences. In the following section, I will use the main idea of their proposal to account for blended-perspective pictures.

3 Perspective blending: exploring solutions

3.1 First proposal: Splitting & viewpoint-shifting

Although Abusch & Rooth's (2017) analysis can account for free perception sequences, it is not clear how it could work for pictures like (1a): these seem like free perception sequences that are merged or blended into a single image. In this section I explore the idea of "unblending" such pictures by turning them into free perception sequences, and following a similar embedding as proposed in Abusch & Rooth (2017).

When seeing pictures like (1a), and of course based on the previous narrative, we infer that the scene surrounding the figure of the protagonist, call him j, reflects not the 'objective' world of the fiction, but the subjective world of the protagonist, that is, the world as perceived by the protagonist. In that case, I assume that the reader re-analyses the picture as something similar to a free perception sequence in the following way. First, a picture p is covertly split in two parts resulting in a sequence of two pictures: the first picture, call it p_1, contains the figure of the character, and the second picture, p_2, includes the whole scene. I postulate a splitting function f whose definition is given below and results in the sequence shown in figure (5).

Definition 1. *A splitting function f applied to picture p yields a sequence of two pictures p_1 and p_2: $f(p) = (p_1, p_2)$, where p_1 includes the figure of a salient discourse referent and p_2 includes the whole scene.*

How does this splitting take place, that is, how is each picture determined to contain what it contains? Since Joe is the salient protagonist in the preceding narrative, I assume that the covert splitting is the result of some pragmatic mechanism that can isolate a figure corresponding to a salient discourse referent and separate this from the rest of the picture (see the discussion on discourse

Fig. 5. Output of splitting function f: p_1(left), p_2 (right)

referents in section 2). Joe is the salient protagonist of the story, the most prominent individual around whom the story revolves, thus, it is his epistemic state that is relevant for the interpretation of the non-veridical content of picture p.

The sequence created by the splitting function is not yet a free perception sequence: the second panel represents an external, third-person perspective, not an internal, first-person one. In order to use an intensional operator like Abusch & Rooth's, we have to adjust the perspective of the second picture, p_2, because we want to capture the fact that the world in p_2 is the world as perceived by the character. In order to accommodate this, we can hypothesise that a second function operates which adjusts the current, external viewpoint to the character's viewpoint. The result is a set of hypothetical pictures with a first-person, internal perspective (i.e. a set of free perception panels), that 'match' the content of p_2.

The reason why the viewpoint-shifting function results in a *set* of pictures and not in a unique picture is because we do not really know what the character precisely sees: we just come to imagine what the world would look like from his point of view, therefore many pictures could be compatible with what he saw. Suppose picture (6a) below is the initial, external-perspective picture. Pictures (6b) and (6c) depict two possible worlds viewed from the character's eye location. Picture (6d) does not depict a possible world: in (6a), there is no big blue rabbit between the cube and the triangle. On the other hand, from picture (6a) we cannot see what the character sees precisely, therefore there could be nothing between the cube and the triangle (picture 6b) or there could be some small rabbit that is not visible from above (picture 6c). Hence, the viewpoint-shifting function gives us all and only the scenarios that match with picture (6a).

The main idea is that a world can be represented from different viewpoints. In our case, this means that p_2 and the output pictures are different representations of the reality as perceived by the character. To make this precise, we abstract away from viewpoints and use the definition of the uncentered content of a picture p (see Rooth & Abusch 2017):

$$[\![p]\!]^* = \{w \mid \exists v.\pi(w, v) = p\} \tag{3}$$

In words, the uncentered content of a picture p is the set of worlds w such that from some viewpoint v they are projected onto p.

Now we can give the definition of the viewpoint-shifting function:

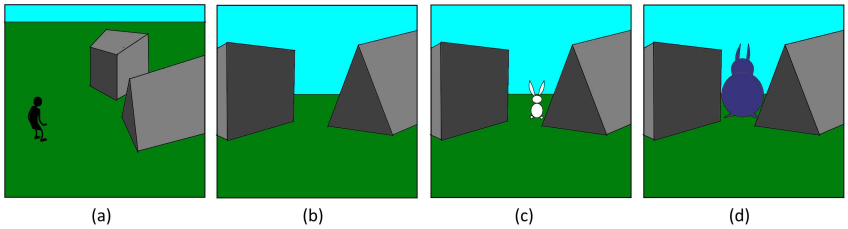

Fig. 6. (a) External (third-person) perspective picture. (b), (c), (d) Internal (first-person) perspective pictures.

$$g(v_j, p_2) = \{\pi(w, v_j) \mid w \in [\![p_2]\!]^*\} \tag{4}$$

Function g applies to the viewpoint of the discourse referent j introduced in p_1, i.e. v_j, and to picture p_2, and shifts the viewpoint of p_2 to v_j, yielding a set of first-person perspective pictures representing worlds that belong to the propositional, uncentered content of p_2. More simply, the output of g is a set of pictures $\{q_1, q_2 ... q_n\}$.

What we have now is actually a set of free perception sequences, whose first panel is p_1 and the second panel is a member of the output set of g, i.e. $\{q_1, q_2 ... q_n\}$. The remaining process is as proposed in Abusch & Rooth (2017), namely embedding each member of the set $\{q_1, q_2 ... q_n\}$ under an intensional operator P.

The proposal presented in this section is somewhat complex as it involves different steps. Is there a more simple analysis to model interpretation of blended pictures? In the following two sections, I consider two more ways.

3.2 Second proposal: Perspective blending as free indirect discourse?

A different approach would be to regard blended pictures as instances of free indirect discourse. In this section, I explore this idea.

Free indirect discourse is a literary style through which a character's thoughts and perceptual experience are represented with the mediation of the narrator's voice. This "intermediate mode" (McHale 2011) creates ambiguity as it is not clear if the point of view expressed is the narrator's or the character's. The usual pattern in free indirect discourse is that pronouns and tenses behave as in indirect discourse and reflect the narrator's perspective: third-person pronouns refer to the character(s) and past tenses refer to the character's present (at least in English). The rest of the expressions, such as temporal and locative adverbials (*now*, *here*), reflect the character's perspective (for a thorough overview of the expressions used in free indirect discourse, see Banfield 1982). For an illustration, see the example below:

(1) Tomorrow was Monday, Monday, the beginning of another schoolweek! (Lawrence, *Women in Love*, p. 185, as cited in Banfield 1982: 98)

In this example, *tomorrow* refers to the day following the day where the protagonist is temporally located, while the past tense *was* is anchored to the narrator's context.

Certain studies in the field of narratology discuss possible parallels of free indirect discourse in graphic novels and films (for example, Forceville 2002, Mikkonen 2008, Ghaffary & Nojoumian 2013). According to these studies, analogues of free indirect discourse in comics and/or in films are to be found in instances where the reader/viewer cannot determine whether what is represented visually (but also aurally in the case of films and animation) corresponds to the narrator's or the protagonist's perspective. For instance, free perception panels in comics and 'point-of-view' shots in film are taken in certain cases as possible parallels to free indirect discourse whenever they create ambiguity. But can this be supported semantically?

In semantic literature, different analyses are proposed about the status of free indirect discourse (see Banfield 1982, Schlenker 2004, Sharvit 2008, Eckardt 2014, Maier 2015, among others). Here I will remain agnostic as to which analysis is the most tenable. What is of special relevance is the behaviour of indexical expressions in free indirect discourse, which seems to be one of its defining characteristics. In standard discourse, the context of utterance is responsible for the interpretation of all indexical expressions such as first- and second-person pronouns, and also temporal and locative adverbials like *here* and *now*. Direct discourse can be considered a context-shifting mechanism because all expressions in a direct discourse report refer to the context of utterance being reported. By contrast, as mentioned above, in free indirect discourse, indexicals do not behave uniformly. This could lead to the assumption that free indirect discourse is basically a context-shifting mechanism for certain expressions.

Now, let's move to pictures. Here, there is no visual parallel of context-dependent linguistic expressions like temporal adverbials. However, change in the geometrical perspective of a picture is change in the locational point of reference, so this could be considered parallel to context-shifting. Free perception sequences involve change of perspective from one picture to another, so they can be viewed as instantiations of context-shifting. On the other hand, single pictures like (1a) do not involve any shifting as for a single picture there is by default only one corresponding viewpoint from which the whole scene (the character and the rest) is depicted (unless there are other embedded pictures).[6] Hence, pictures like (1a) do not seem to be semantic parallels to free indirect discourse if this split of indexicals is its defining feature.

[6] One general remark should be made. Throughout a graphic narrative, there is continuous switching between multiple viewpoints. However, I reduce my discussion to *two* viewpoints in order to refer to two broader notions: a) an external viewpoint that corresponds to the possibly multiple locations in space-time that the narrator can take, and b) an internal viewpoint that corresponds to a protagonist's first-person perspective.

However, in line with the observations in narratological studies, as far as the effect on the reader is concerned, blended pictures seem to convey the same ambiguity as passages in free indirect discourse do. More specifically, free indirect discourse reports are usually 'free' in that they are not embedded under an attitude or saying verb. Of course they can include a parenthetical verb indicating whether the sentence is a speech or thought event (example 2 below), but only optionally. The absence of such a verb (example 3) may make the reader wonder if what is described is 'uttered' by the narrator and is therefore true in the story, or if the sentence only represents a character's thought or perceptual experience. See the examples below:

(2) It was seven o'clock, he thought.

(3) It was seven o'clock. (examples from Banfield 1982:205)

The same challenge is placed onto the reader of a graphic novel when seeing (1a): is this picture a representation of the actual world in the story or is the character hallucinating?[7]

To sum up, regarding the effect on the reader's interpretation, blended pictures have a similar impact as passages in free indirect discourse in novels and they too appear to be 'syntactically free'. Nonetheless, from a semantic point of view, such a view cannot be supported. In the following section, I will discuss a more plausible analysis.

3.3 Third proposal: Blended pictures as indirect discourse

In this section, I will explore an alternative analysis that makes use of an intensional belief-operator.

Instead of following the decompositional approach presented in section 3.1, we can suggest a simpler analysis. Pictures like (1a) can be regarded as parallel to indirect thought reports in language like "Joe thinks that he is surrounded by superheroes". In indirect discourse in written/spoken language, the narrator's perspective in the embedded clause is reflected through the use of the 3rd person pronoun. Something similar happens in pictures like (1a): the protagonist is also represented from a third-person perspective in the image. I assume that via inferential reasoning the reader will come to realise that what is going on in the picture is actually a hallucination. Inspired by Eckardt (2014), I assume a 'cautious update' for blended-perspective pictures (Maier & Bimpikou 2018).

The idea behind cautious update is that, even in normal, non-fictional discourse, we do not always update the common ground directly with the propositional content p of a sentence. If, for example, we consider the speaker confused, we take p to be part of the speaker's belief state only and not to form part of the shared common ground. Thus, instead of updating the common ground with

[7] Ambiguity can be resolved through text, but here I just consider cases where no captions are included.

the set of worlds where proposition p is true, we update with the proposition that the speaker believes that p.

Extending this to fictional discourse, a cautious update may take place whenever the reader assumes that a certain proposition is not true in the world of the fiction but true according to a protagonist, i.e. true in the protagonist's belief or imagination worlds. When seeing (1a), the reader infers that the protagonist hallucinates and so she has to perform a cautious update, i.e. embed the propositional content of the picture under an intensional belief-operator BEL. The update will result in interpreting the picture as something like "Joe believes that he is surrounded by superheroes". This results in the picture being interpreted as depicting the character's subjective world and not the actual world of the fiction. More generally, for a picture p and a salient protagonist j, $\text{BEL}_j p$ is true iff for all worlds $w\prime \in Bel_j$ (where Bel_j is j's belief state), $w\prime \in [\![p]\!]^*$ (here we use the uncentered, classical propositional content as defined in formula 3).

Should we appeal to a different operator, e.g. an imagination operator IMG? It should be remarked that imagination is different from hallucination or, more generally, faulty perception. When imagining, for example, when engaging in role playing or when daydreaming, we do know that our imaginary worlds are different from the actual, real world. On the other hand, in the case of faulty perception, there is no such awareness on the part of the perceiver. Therefore, imagining is distinct from misperceiving: the first involves the (aware) construction of a mental representation on the part of the agent, whereas misperception involves no distinction on her part between the actual and the imaginary world. Hence, when misperceiving, the agent actually believes that what she perceives is true. This is why a belief operator seems more appropriate. What is common in both imagining and believing though is that there are two different 'layers', the external and the internal reality. As for how the perceiver's awareness of the distinction between actual and mental is conveyed in each case through pictures, there seems to be a difference in marking in comics, as illustrated contrastively in figure (7). The blended picture (1a) is repeated in (7a); figure (7b) is a made-up image where the character appears instead with a thought bubble. Thought bubbles are conventionally used for imaginings and thoughts, so the most natural interpretation for figure (7b) would be that Joe is consciously thinking or imagining something. By contrast, for hallucinations there is no overt marking enclosing the character's perception and we can either have blended pictures (fig. 7a) or free perception panels (fig. 3).[8] So figures (7a) and (7b) seem to prompt different interpretations. Any particular choice (overt embedding with a bubble or non-embedding) has a significant effect on the reading process and consequently on the reader's interpretation. This makes the prediction that a picture with a thought bubble should be unambiguous, whereas pictures like (7a) can be ambiguous: a reader might fail to understand that what is represented in the picture is true only in the character's mind.

[8] For a detailed discussion on speech and thought bubbles in comics and a somewhat different approach to their relation to "awareness", see Cohn (2013).

Fig. 7. (a) Blended picture. (b) Picture with a thought bubble.

Although I mainly discuss single pictures, it is very common in comics to have sequences of blended pictures spanning a large part of the narrative, as in Bill Watterson's comic series *Calvin and Hobbes*. What would be the most satisfactory proposal from the ones suggested so far, also from the point of view of cognitive processing? According to the first proposal, the reader has to re-imagine the scene from a first-person perspective. The indirect discourse approach involves the insertion of a belief operator without applying extra operations such as a viewpoint-shifting mechanism. For sequences of blended pictures, we may assume that these are grouped together as a constituent and that the intensional operator scopes over the whole constituent. This is also relevant for animation and film. Consider the movie *Fight Club* or the episode of *Bojack Horseman* (fig. 1b) where Bojack hallucinates. It is hard to imagine how splitting and reorienting suggested in 3.1 for pictures could be applied in continuous shots. For animation, we could suggest that the intensional operator could apply at the level of a whole scene (taking a scene to correspond to a series of successive shots that represent a certain spatio-temporal slice of the fictional world).

Overall, blended pictures can be paralleled to indirect discourse reports in language. That makes the indirect discourse analysis more appealing because it can apply to both pictorial and linguistic data.

4 Discussion

In this paper, I discussed perception representation in graphic narratives, mainly comics, and I focused on depictions of characters that are surrounded by their hallucinations. These data are similar to free perception sequences in that they also depict de se experience and therefore can also represent non-veridical perception. Unlike free perception sequences though, our data are single pictures. Our goal was to build on Abusch & Rooth's (2017) account in order to include these data as well. I proposed two ways to analyse blended pictures (sections 3.1 and 3.3) and suggested that a third option, namely comparing blended-perspective pictures to free indirect discourse (section 3.2), is not a tenable approach.

An interesting case is the representation of different kinds of perception in visual narratives. For instance, dreaming is a kind of perceptual experience that, on the one hand, is not exactly like thinking or conscious imagining and, on the

other hand, it is not exactly like hallucination or misperception. The following questions arise: first, how are different kinds of perceptual experience conveyed in pictures and how are representations of perceptual events different or similar across different media? secondly, how are distinct perceptual phenomena in pictures to be modelled semantically? For instance, to the extent that hallucinating and misperceiving are different from dreaming, should we appeal to different kinds of modal operators in the mental representations of blended-perspective pictures?

A different question is whether the above observations can be tested experimentally. I already suggested that blended pictures and pictures with bubbles represent different kinds of perceptual experience especially with respect to how they encode the agent's awareness of the imagined content. Do different ways of representation cause significant differences in readers' interpretations as was suggested in section 3.3? These issues are left for future work. I hope to have pointed out some interesting directions for further research.

Acknowledgements

I would like to thank my supervisor Emar Maier for his help and feedback, as well as the anonymous ESSLLI reviewers for their valuable comments.

References

Abusch, D.: Applying discourse semantics and pragmatics to co-reference in picture sequences. Proceedings of Sinn und Bedeutung 17 (2012)

Abusch, D.: Possible worlds semantics for pictures. Handbook article. Draft (2015)

Abusch, D., Rooth, M.: The formal semantics of free perception in pictorial narratives. Proceedings of the 21st Amsterdam Colloquium (2017)

Banfield, A.: Unspeakable sentences. Routledge and Kegan Paul, Boston (1982)

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

Eckardt, R.: The semantics of free indirect discourse: How texts allow us to mind-read and eavesdrop. Brill (2014)

Forceville, C.: The conspiracy in the comfort of strangers: Narration in the novel and the film. Language and Literature **11**(2), 119–135 (2002)

Ghaffary, M., Nojoumian, A.A.: A poetics of free indirect discourse in narrative film. Performance Studies: Rupkatha Journal on Interdisciplinary Studies in Humanities, Volume V, Number 2, 2013 p. 269 (2013)

Greenberg, G.J.: The semiotic spectrum. Rutgers The State University of New Jersey-New Brunswick (2011)

Maier, E.: Quotation and unquotation in free indirect discourse. Mind & Language **30**(3), 345–373 (2015)

Maier, E., Bimpikou, S.: Blending perspectives in pictorial narratives. Abstract accepted at Sinn & Bedeutung 23 (2018)

McHale, B.: Speech representation. `http://www.lhn.uni-hamburg.de/article/speech-representation` (2011), revised 2014 and accessed June 30, 2018

Mikkonen, K.: Presenting minds in graphic narratives. Partial Answers: Journal of Literature and the History of Ideas **6**(2), 301–321 (2008)
Rooth, M., Abusch, D.: Picture descriptions and centered content. Proceedings of Sinn und Bedeutung 21 (2017)
Schlenker, P.: Context of thought and context of utterance: A note on free indirect discourse and the historical present. Mind & Language **19**(3), 279–304 (2004)
Sharvit, Y.: The puzzle of free indirect discourse. Linguistics and Philosophy **31**(3), 353–395 (2008)