Proceedings of SALT 27: 282–301, 2017

Two kinds of perspective taking in narrative texts·

Stefan Hinterwimmer
University of Cologne

Abstract In this paper, I argue for the existence of two distinct kinds of protagonists' perspective taking in narrative texts. The first, *Free Indirect Discourse*, represents conscious thoughts or utterances of protagonists and involves context shifting: All context-sensitive expressions with the exception of pronouns and tenses are interpreted with respect to the fictional context of some salient protagonist (Schlenker 2004; Sharvit 2008; Eckardt 2014, Maier 2015). The second, which I dub *viewpoint shifting*, does not necessarily represent conscious thoughts or utterances and it does not involve context shifting. Rather, a situation is described as it is perceived by a salient protagonist or in a way that reflects the doxastic state of such a protagonist, not with respect to the Common Ground (CG) of narrator and reader. While FID is only available at the root level, i.e. at the speech act level, viewpoint shifting is available at the level of finite matrix clauses.

Keywords: perspective, free indirect discourse, context shifting, speech acts

1 Introduction

In narrative texts sentences that are neither direct quotations nor embedded under a propositional attitude verb such as *say* or *think* can nevertheless often be understood as expressing the utterances or thoughts of a protagonist. This particular form of speech or thought representation, which is formally clearly distinguishable from both direct discourse (henceforth: DD) and indirect discourse (henceforth: ID), has long been known to literary scholars by the name

· I would like to thank Klaus von Heusinger, Sara Meuser, Petra Schumacher and the participants of *SALT* 27, the workshop *Agentivity and Event Structure* at the 38th Annual Meeting of the German Linguistic Association, the workshop *Linguistics and Cognition* at the University of Osnabrück, the Linguistics Colloquium at the University of Göttingen, and the workshop *Speech Acts* at the Leibniz-Zentrum für Allgemeine Sprachwissenschaft for questions and comments. The research reported in this paper was funded by the Deutsche Forschungsgemeinschaft (DFG) as part of the Collaborative Research Center 1252 *Prominence in Language* at the University of Cologne.

©2017 Hinterwimmer

free indirect discourse (henceforth: FID) (see Hamburger 1968 and Stanzel 1979 and the references therein), and has also been discussed by many linguists in descriptive terms (Harweg 1972; Plank 1986; von Roncador 1988; see Dirscherl and Pafel 2015 for a recent overview). More recently, FID has become a research topic for linguists working in the tradition of truth conditional semantics, and, building on ideas first formulated by Banfield (1982) and Doron (1991), various analyses have been proposed to capture the distinctive properties of FID in a formally precise way (Schlenker 2004; Sharvit 2008; Eckardt 2014; Maier 2015).[1] The most striking of these properties is that in FID all deictic expressions with the exception of pronouns and tenses are interpreted with respect to a fictional context (in the sense of Kaplan 1979) that has been set up by the preceding linguistic context. That fictional context consists of the protagonist whose thoughts or utterances are represented, the spatiotemporal location of that protagonist at the respective utterance or thinking time and the world containing that spatiotemporal location. Pronouns and tenses, in contrast, cannot be interpreted this way since both the respective protagonist and her (in case of speech representation) addressee can only be referred to by third person pronouns, not by first and second person ones, and verbal tense marking always contains a layer of past, not present. This mixed behavior of deictic elements has lead Schlenker (2004), Sharvit (2008) and Eckardt (2014) (whose analyses otherwise differ in many important respects) to postulate that sentences in FID mode are interpreted with respect to two different contexts: the narrator's context, which is only relevant for the interpretation of pronouns and tenses, and the fictional context of the respective protagonist, which is relevant with respect to all other context-sensitive expressions. Maier (2015), in contrast, analyzes FID as a special, highly conventionalized form of mixed quotation: The reader has to accommodate a prominent protagonist's speech or thought act to be partially quoted, with pronouns and tenses being unquoted.

In this paper I argue that there is a second kind of protagonists' perspective taking, which I dub *viewpoint shifting* (henceforth: VS) and which is likewise not marked as such via the presence of propositional attitude verbs, but which differs from FID in three important respects. First, it does not involve partial context shifting, i.e. all context sensitive expressions are interpreted with respect to the narrator's context. Second, what is rendered in VS is not necessarily the content of a conscious thought or utterance. Rather, a situation is described as it is perceived by a salient protagonist or in a way that reflects the doxastic state of such a protagonist at the relevant time, not with respect to the Common Ground (CG) of narrator and reader. Finally, while FID is only available at the level of

[1] See Vandelanotte 2009 and Dancygier 2012 and the references therein for analyses of FID from the point of view of cognitive linguistics.

entire sentences corresponding to speech or thought acts, VS is available at the level of finite matrix clauses that are contained in complex sentences.

The paper is structured as follows. In Section 2 the relevant characteristics of FID and VS are introduced and discussed. In Section 3 I present my analysis of VS in the first subsection. In the second subsection I discuss the question of which of the existing analyses of FID is best suited to account for the restrictions observed in Section 2. Section 4 discusses some directions for future research.

2 The formal characteristics of FID and VS

2.1 Free Indirect Discourse

Consider the three sentences in (1a-c):

(1) a. On her way home, Mary heard a song by Kendrick Lamar that she liked on the radio. She would buy his new album tomorrow.
b. On her way home, Mary heard a song by Kendrick Lamar that she liked on the radio. She thought that she would buy his new album on the following day.
c. On her way home, Mary heard a song by Kendrick Lamar that she liked on the radio. She thought: "I will buy his new album tomorrow".

The final sentence in (1a) exemplifies the characteristics of FID mentioned in the introduction: On the one hand, the deictic temporal adverb *tomorrow* is interpreted with respect to the fictional context provided by the immediately preceding sentence, i.e. as referring to the day following the day where the event introduced by the opening sentence took place. On the other hand, the layer of past tense marking on the modal auxiliary (giving rise to a future in the past interpretation[2]) in combination with the fact that a third person pronoun is used to refer to Mary clearly indicates that no complete shift to Mary's fictional context has occurred. If that were the case, the present tense form of the modal auxiliary, *will*, would have to be used, and Mary would be referred to by the first person pronoun *I*, as in the quoted sentence in (1c).

As already said in the introduction, two different ways to capture the distinctive interpretive properties of FID have been proposed in the formal semantics literature. According to the first line of analysis, sentences are (optionally) interpreted with respect to two different contexts: the speaker's (in oral discourse) or narrator's (in narrative texts) context, and a protagonist's context (see Schlenker 2004; Sharvit 2008 and Eckardt 2014 for related, but technically different implementations of the same basic idea). Let us for the

[2] See Eckardt 2017 for detailed discussion of future in the past and its relevance for FID.

moment follow Eckardt (2014) for concreteness and assume that in the final sentence of (1a) a context *c*, whose parameters are indirectly provided by the immediately preceding sentence, is introduced in addition to the narrator's context *C*: The author of *c* is the experiencer of the event introduced by the preceding sentence, Mary, the time of *c* is the time immediately following the run time of that event, and the location of *c* is the location of Mary at the time of *c*. Crucially, all lexical items with the exception of pronouns and verbal tenses are lexically specified as being interpreted with respect to *C* by default and with respect to *c* whenever the latter is introduced. Pronouns and verbal tenses, in contrast, always have to be interpreted with respect to *C*. Consequently, the pronoun *she* and the past tense marking on the modal auxiliary are interpreted with respect to the narrator's context *C*, while the rest of the sentence is interpreted with respect to Mary's context *c*. Since the sentence does not contain any other context-sensitive elements, this is of course only relevant for the interpretation of *tomorrow*. Additionally, the proposition denoted by the sentence is not added to the Common Ground of narrator and reader directly. Rather, what is added is the proposition that the author of *c*, Mary, believes that proposition. Simplifying considerably, the two sentences in (1a) are accordingly interpreted as paraphrased in (2):

(2) In all worlds of the Common Ground there is an event *e* of Mary hearing a new song by Kendrick Lamar that she likes on the radio on her way home and *e* precedes the time of *C*, and in all worlds that are compatible with what Mary thinks at the time immediately following the run time of *e* there is an event *e´* of Mary buying the new album by Kendrick Lamar whose run time precedes the time of *C* and follows the run time of *e*, and the run time of *e´* is contained in the day following the day that contains the time of *c*.

The final sentence in (1b) denotes exactly the same proposition as the one that is added to the Common Ground as the denotation of the final sentence in (1a), but the crucial difference is that no additional context *c* is introduced. Rather, all deictic elements are interpreted with respect to the narrator's context *C*. Additionally, the information that the event under consideration is not (necessarily) true in the worlds of the Common Ground, but rather in Mary's belief worlds, is directly provided by the propositional attitude verb contained in the matrix clause. Finally, in the case of (1c), where the final sentence is quoted, all deictic elements are interpreted with respect to the sentence's original context, i.e. the context in which Mary had the quoted thought. Additionally, the thinking event itself is introduced by the propositional attitude verb preceding the colon.

Maier (2015, to appear) proposes a different analysis of FID: He assumes FID to be a special form of mixed quotation, i.e. all expressions contained in a sentence in FID mode, with the exception of pronouns and verbal tenses, are quoted, while pronouns and tenses are unquoted. According to Maier (2015, to

appear), the final sentence in (1a) is thus rather similar to the final sentence in (1c), with two important differences: First, whereas the thinking event whose content is quoted is introduced explicitly in the case of (1c), it has to be accommodated in the case of (1a). Second, while in (1c) the thought is quoted in its entirety, the pronoun *she* and the past tense of the modal auxiliary are unquoted in (1a). Concerning the question of why pronouns and tenses, as opposed to other items, are unquoted in FID, Maier (2015) assumes this to be a pragmatically driven convention that is also in effect in other forms of mixed quotation (see Maier to appear for discussion). The most important argument that Maier gives in favor of his analysis is the following one: FID involves more than a shifted interpretation of context-sensitive items. There are also cases such as the text segment given in italics in (3) where a protagonist's thoughts or utterances are rendered in the non-standard dialect spoken by that protagonist while the surrounding text is written in standard English. Such cases are unproblematic for an analysis of FID in terms of mixed quotation. It is unclear, however, how they are to be captured by an account which analyzes FID via partial context shift.

(3) He [Big Boy] remembered the day when Buck, jealous of his winning had tried to smash his kiln. *Yeah, that ol sonofabitch*! *Naw, Lawd! [. . .] Cussin the dead! Yeah, po ol Buck wuz dead now. N Lester too. Yeah it wuz awright fer Buck t smash his kiln. Sho. N he wished he hadnt socked ol Buck so hard tha day*[3].

Consider now the three sentences in (4a-c), which differ from those in (1a-c) as follows: The content of the first sentence is no longer given as an independent sentence, but rather as a *when*-clause modifying the main clause which corresponds to the respective second sentence in (1a-c).

(4) a. * When Mary heard a song by Kendrick Lamar that she liked on the radio on her way home, she would buy his new album tomorrow.
b. When Mary heard a song by Kendrick Lamar that she liked on the radio on her way home, she thought that she would buy his new album on the following day.
c. When Mary heard a song by Kendrick Lamar that she liked on the radio on her way home, she thought: "I will buy his new album tomorrow".

The main clause in (4a), in contrast to the second sentence in (1a), cannot be interpreted as expressing a thought of Mary, with *tomorrow* being interpreted with respect to the fictional context (implicitly) introduced by the *when*-clause. At the same time, no other coherent interpretation is available: On the one hand, the

[3] Wright, Richard (1979). Big Boy leaves home. In *The Literary South*. New York: JohnWiley & Sons. Cited by Maier (2015).

main clause event of Mary buying the new album by Kendrick Lamar is said to temporally overlap with the *when*-clause event, which is located in the past with respect to the narrator's context. On the other hand, the main clause event is located in the future with respect to the narrator's context by *tomorrow* as well as by the modal auxiliary verb. Since no event can meet both temporal specifications, the sentence is necessarily contradictory and thus sounds extremely awkward. Note that the sentence remains awkward, even if *tomorrow* is replaced by the non-deictic adverbial *on the following day,* due to those incoherent temporal specifications. The variants in (4b) and (4c), in contrast, where a thinking event with Mary as experiencer is explicitly introduced, are perfectly fine. Note that since in (4b) it is the thinking event which is required to temporally overlap with the *when*-clause event, the double temporal specification of the buying event is as unproblematic as in (1a) and (1b): The buying event is said to precede the time of the narrator's context and to follow the run time of the thinking event.

As shown by the analogous contrasts between (5a) and (6a), on the one hand, and (5b-c) and (6b-c), on the other, the restriction under consideration does not have anything to do with the specific temporal semantics of *when*: Rather, FID only seems to be available at the root level, but not sentence-internally.

(5) a. * After Mary had heard a song by Kendrick Lamar that she liked on the radio on her way home, she would buy his new album tomorrow.
b. After Mary had heard a song by Kendrick Lamar that she liked on the radio on her way home, she thought that she would buy his new album on the following day.
c. After Mary had heard a song by Kendrick Lamar that she liked on the radio on her way home, she thought: "I will buy his new album tomorrow".

(6) a. * Because Mary really liked the song by Kendrick Lamar that she heard on the radio on her way home, she would buy his new album tomorrow.
b. Because Mary really liked the song by Kendrick Lamar that she heard on the radio on her way home, she thought that she would buy his new album on the following day.
c. Because Mary really liked the song by Kendrick Lamar that she heard on the radio on her way home, she thought: "I will buy his new album tomorrow".

If one adopts an analysis of FID as involving the introduction of an additional protagonist's context (as in Eckardt 2014; see Schlenker 2004 and Sharvit 2008 for closely related, but technically different analyses), the contrasts discussed in

this section can be taken to show that such an additional context can only be introduced at the root level, but not sentence-internally. If one adopts an analysis of FID as a special kind of mixed quotation (as in Maier 2015), in contrast, the restriction of FID to the root level can be accounted for as follows: The speech or thought act to be partially quoted can only be accommodated at the root level, but not sentence-internally. In Section 3 I will turn to the question of which of the existing analyses of FID is best suited to account for the contrasts observed in this section, after having proposed my analysis of a second kind of protagonists' perspective-taking, *viewpoint shifting* (VS), which will be introduced in the next subsection.

2.2 Viewpoint shifting

Consider the text segment in (7), an excerpt from David Eggers' *A Heartbreaking Work of Staggering Genius*, which is to be understood against the following background: After the death of their parents, the ego-narrator, Dave, has to take care of his little brother, Toph. Toph has slept at a friend's house, and Dave has spent the night away from home with a woman, but wants Toph to believe that he stayed home all night.

(7) I wanted to be home in case he came back early, made it in time … The house was empty, and I dove into bed, fell back asleep, and when he [Toph] came back home *his brother was there, of course had been there the whole time, of course had never left*. (D. Eggers, *A Heartbreaking Work of Staggering Genius*: 112, cited in B. Dancygier 2012: 62).

Obviously, the main clause of the final sentence (in italics) can only be understood as describing the situation from Toph's perspective, as Dave imagines it[4]: First, it would make no sense at all for Dave, the ego-narrator, to refer to himself via the definite description *his brother* instead of the first person pronoun *I*. Second, Dave knows the propositional content of the clause to be false, i.e. he knows that he has not been home the whole time, and that he has left. For Toph, in contrast, it is plausible to believe it to be true, especially when he sees his brother lying in bed and sleeping.

While the clause in italics is clearly understood as expressing Toph's perspective, there are good reasons to assume that it is not an instance of FID. That this is so is, first, confirmed by the acceptability of the variant in (8), where the pronoun in the *when*-clause has been replaced by a definite description that clearly indicates that ego-narration is retained in the *when*-clause, and that the

[4] See Dancygier 2012 for detailed discussion of the complex network of perspectives in David Eggers' *A Heartbreaking Work of Staggering Genius*.

shift to Toph's perspective occurs in the main clause. Since we have seen numerous examples in the previous section that FID is only available at the root level, this provides the first crucial piece of evidence that what we see in the main clause of the final sentence in (7) is a different kind of perspective taking, which I, as already mentioned in the introduction, dub *viewpoint shifting* (VS).

(8) I wanted to be home in case he came back early, made it in time … The house was empty, and I dove into bed, fell back asleep, and when my little brother came back home *his brother was there, of course had been there the whole time, of course had never left*.

The second piece of evidence is provided by the following observation: Replacing the local adverb *there* in (7) by the adverb *here* leads to infelicity, as shown in (9). If the clause in italics was an instance of FID, however, *here* would be the expected choice: It is one of the defining characteristics of FID that all context-sensitive expressions with the exception of pronouns and tenses are interpreted from the respective protagonist's perspective, and the adverb refers to the location of Toph at the relevant time. This is further confirmed by the fact that in the variant in (10), where the content of the clause in italics is given in DD-mode, *here* is perfectly fine.

(9) # …, and when he came back home *his brother was here, of course had been here the whole time, of course had never left*.

(10) …, and when he came back home he thought: "My brother is here, of course has been here the whole time, of course has never left".

Consider next the two text segments in (11) and (12), to be understood against the following background: Two little dinosaurs, one of them named Billy, have run away from some predator that they could not identify and are hiding in a cave.

(11) [The T-Rex] hesitated. Maybe the little dinosaurs had hidden themselves in the cave on his left. ??[A T-Rex]/[The T-Rex] bent down to the entrance of the cave and squinted into the dark.

(12) [The T-Rex] hesitated. Maybe the little dinosaurs had hidden themselves in the cave on his left. When Billy looked up in his hiding place a few seconds later, [a T-Rex] bent down to the entrance of the cave and squinted into the dark.

In the case of (11), the indefinite *a T-Rex* can only be understood as introducing a novel T-Rex that is distinct from the one referred to by the definite description in the first sentence. In fact, the sentence sounds slightly marked in the context of the first sentence and requires *another* instead of *a* to be fully acceptable (see Grønn & Sæbø 2012). In the case of (12), in contrast, the indefinite is not

understood as introducing a novel T-Rex, but rather as picking up the T-Rex referred to by the definite description in the first sentence. Intuitively, the contrast comes about as follows: In (11), using an indefinite to pick up the already familiar T-Rex would violate the *familiarity condition* (Heim 1982) or, alternatively, *Maximize Presupposition*! (Heim 1991, Singh 2011). After all, the final sentence is automatically understood as being told from the narrator's perspective, just like the opening sentence, while the second sentence reports a thought of the T-Rex in FID-mode (as indicated by the presence of *maybe*, which would make little sense from the narrator's perspective). Consequently, the Common Ground of narrator and reader is relevant for the interpretation of the indefinite, and with respect to that Common Ground the T-Rex chasing the two little dinosaurs is already familiar/contextually unique. It can thus only be picked up by a definite description or a pronoun, and the indefinite *a T-Rex* can only be understood as introducing a novel T-Rex. In the case of (12), in contrast, the left-adjoined *when*-clause licenses a shift from the T-Rex's to Billy's perspective. Consequently, the main clause describes what Billy sees when he looks up in a way that is compatible with Billy's doxastic state, not with respect to the Common Ground of reader and narrator. Since with respect to Billy's doxastic state the T-Rex chasing him is novel, using the definite article is not licensed, and the indefinite can easily be understood as picking up the T-Rex that the narrator has referred to via a definite description in the opening sentence without violating the familiarity condition or *Maximize Presupposition*! (see Passonneau 1994 and Abbott & Horn 2011 for discussion of similar uses of indefinites[5]).

What is most important for our current purposes is that the shift to Billy's perspective again occurs sentence-internally, not at the root level: It is the *when*-clause that makes Billy available as a perspective-taker in virtue of giving a neutral description of an event whose experiencer is Billy from the narrator's perspective. This is obvious by comparing (12) with (11), which does not contain a left-adjoined *when*-clause, and where a shift to Billy's perspective is not available. The short text segment in (12) thus involves three different perspectives: The first sentence and the *when*-clause of the third sentence are interpreted with respect to the narrator's perspective, while the second sentence is an instance of FID with the T-Rex as perspective-taker. Concerning the interpretation of the main clause of the third sentence, the fact that the shift to Billy's perspective occurs sentence-internally provides evidence that it is not an instance of FID, but rather of VS. A second reason to assume this is that, intuitively, the clause does not report (or partially quote) the content of a conscious thought of Billy, as it would if it was FID. Rather, what Billy sees is described in a way that is compatible with his doxastic state, i.e. as he presumably

[5] I am grateful to Alexander Williams for making me aware of this research.

would describe it if he were asked to. To see this, consider the variant of (12) given in (13), where the content of the main clause of the third sentence is given in DD-mode, i.e. as a full quote of a conscious thought of Billy that is explicitly introduced as such. The text segment in (13) is not equivalent to (12), in contrast to cases like (1a) and (1c) from above. Rather, the final sentence in (13) sounds rather unnatural, or at least far less natural than the one in (12).

(13) [The T-Rex] hesitated. Maybe the little dinosaurs had hidden themselves in the cave on his left. When Billy looked up in his hiding place a few seconds later, he thought: "A T-Rex is bending down to the entrance of the cave and squinting into the dark".

As a final instance of VS consider the main clause of the sentence in (14): It has a prominent reading on which the ground did not really shake, but Mary just had that feeling as a result of her sense of balance being disturbed by the boat trip. As in the previous examples, the shift from the narrator's to the protagonist's perspective occurs sentence-internally, i.e. the *when*-clause gives a neutral description of an event whose agent is Mary from the narrator's perspective and thereby establishes Mary as a (potential) perspective-taker with regard to the content of the main clause. Again, the main clause is not (or at least not necessarily) understood as reporting a conscious thought of Mary, but rather as describing a sensation of her in a way that is compatible with her doxastic state at the time at which she has that sensation.

(14) When Mary stepped out of the boat, the ground was shaking beneath her feet for a couple of seconds.

In this section I have provided evidence for the existence of VS as a kind of protagonist's perspective taking that is clearly different from FID in at least three respects: First, it is not only available at the root level, but sentence-internally. Second, it does not (necessarily) report a protagonist's conscious thoughts, but rather describes an eventuality perceived by a protagonist or a sensation of a protagonist in a way that is compatible with that protagonist's doxastic state at the time at which she perceives the respective eventuality/has the respective sensation. Third, it does not involve a shifted interpretation of context-sensitive items. In Section 3 I will first propose a formal analysis of VS and then return to the question of which of the existing formal analyses of FID is best suited to account for the restriction of FID to the root level in a natural way.

3 Analysis

3.1 Viewpoint shifting

In the previous section, we have seen that VS, in contrast to FID, is not restricted to the root level, but can occur sentence-internally. I assume that this is because VS does not involve the accommodation of a speech or thought act. Rather, an eventuality (where the term *eventuality* applies to dynamic events as well as to states, as in Landman 2000) is described in a way that is compatible with the doxastic state of a prominent protagonist at the time of him or her perceiving it. Consequently, a perceiving eventuality needs to be accommodated on the basis of contextually prominent information. In the cases considered in this paper that information is provided by the left-adjoined *when*-clause whose run time determines the run time of the perceiving event and whose agent or experiencer provides the experiencer argument of the perceiving eventuality.

I assume that just like everything that all interlocuters perceive consciously during a conversation becomes part of their Common Ground (Stalnaker 1978, 2002), everything a protagonist in a narrative text perceives consciously becomes part of the set of propositions representing his/her beliefs at the relevant time. Crucially, under certain conditions (namely in those where VS applies), the content of a clause is not added to the Common Ground of reader and narrator directly, but rather to the set of propositions that represent the doxastic state of a prominent protagonist. For concreteness, I assume the following technical implementation of VS: A covert operator OP_{VS} is optionally inserted at the TP-node of finite matrix clauses. That operator takes an eventuality predicate *P* as its argument and returns a predicate of perceiving eventualities whose further properties are determined by the context and whose experiencer is some contextually prominent individual *x*. Crucially, in all worlds representing the doxastic state of *x* at the time immediately following (the run time of) the respective perceiving eventuality, there is an eventuality satisfying the original eventuality predicate *P*. The covert VS-operator is defined in (15). Note that it takes a covert pronoun introducing a free variable that ranges over individuals (given as *1*) and a covert pronoun introducing a free variable that ranges over eventuality predicates (given as C_2) as arguments.

(15) $[[\mathrm{OP_{VS}}\ 1\ C_2]]^{g,C} = \lambda P_{<ev,<s,t>>}\ .\ \lambda e\ .\ \lambda w\ .\ \mathrm{PERC}(e)(w) \wedge g(C_2)(e)(w) \wedge \mathrm{Experiencer}(e, g(1))(w) \wedge \forall w' \in \mathrm{DOX}_{(g(1))(\tau(e)+)(w)} [\exists e'\ [P(e')(w) \wedge \mathrm{overlap}(\tau(e), \tau(e'))]]$,

where *g* is the assignment function, *ev* is the type of eventualities, *s* is the type of possible worlds, *PERC(e)(w)* means that *e* is a perceiving eventuality in world *w*, *τ(e)* is the run time of *e*, *τ(e)+* is the time immediately following the run time of *e* and $DOX_{(g(1))(\tau(e)+)(w)}$ is the set of

possible worlds that are compatible with what the individual assigned by *g* to the index *1* believes at $\tau(e)+$ in *w*.

Now, various conditions have to be met in order for the insertion of OP_{VS} to be felicitous. First, there has to be a contextually prominent individual that can be assigned to the free individual variable. Second, since a value for C_2 needs to be determined, the context has to make available further information concerning the respective perceiving eventuality. Third, since I assume the insertion of a covert operator to be costly, there has to be a reason for inserting OP_{VS}. In the cases considered in this paper, the reason is that the respective clauses would otherwise receive an inadequate interpretation. Finally, I assume that (again for economy reasons) there has to be a connection between the perceiving eventualities *e* and the doxastic states of their respective experiencers *x* at the time *t* immediately following the run time of *e* in the following sense: The doxastic state of *x* at *t* may not be identical to the doxastic state of *x* at the time preceding *t*.

With these assumptions in place, let us now have detailed look at how the examples discussed in Section 2.2 are interpreted. I will start with the example in (7), repeated here as (16):

(16) I wanted to be home in case he came back early, made it in time … The house was empty, and I dove into bed, fell back asleep, and when he [Toph] came back home *his brother was there, of course had been there the whole time, of course had never left*. (D. Eggers, *A Heartbreaking Work of Staggering Genius*: 112, cited in B. Dancygier 2012: 62).

In the case of the clause in italics in (7)/(16), all conditions for the insertion of OP_{VS} are met. First, the free individual variable can be resolved to the ego-narrator's brother, Toph: Being the agent of the locally highly prominent event introduced by the *when*-clause makes him a very suitable candidate. Second, the required information concerning the further specification of the perceiving eventuality is indirectly made available by the events introduced by the preceding sentences in combination with the one introduced by the *when*-clause: C_2 can thus presumably be resolved to the property of being an eventuality of seeing whose experiencer is Toph and whose theme is an event of Toph's brother lying in bed and sleeping. Third, without the insertion of OP_{VS} the clause would receive an inadequate interpretation since (a) it would be extremely awkward for the ego narrator to refer to himself via the definite description *his brother* instead of the first person pronoun *I* and (b) the ego narrator knows the proposition denoted by that clause to be false. Finally, there is a plausible connection between the eventualities *e* of Toph seeing his brother lying in bed and sleeping and Toph's doxastic state immediately after the run time of *e*: I assume that prior to *e*, Toph already assumed that his brother would be home and would have been home all

night, as indicated by the two occurrences of *of course*. Nevertheless, he could not be entirely sure, i.e. it was not true in *all* of his belief worlds, but only in all of his belief worlds that are compatible with his assumptions concerning the stereotypical course of events. After seeing his brother lying in bed and sleeping, it is true in all of his belief worlds that his brother is home and has been home all night.

In (17) the (slightly simplified, since I have set aside the meaning of *of course* and abbreviated the value of C_2 as *see*) denotation of the clause in italics is derived step by step, under the assumption that OP_{VS} is inserted directly at the TP-node (together with the covert individual and eventuality predicate pronouns) and takes the eventuality predicate denoted by the TP as its argument. Note that T_C is the contextually determined time interval providing the instants quantified over by *the whole time* and *never* – namely the previous night.

(17) $\lambda P_{<ev,<s,t>>} . \lambda e . \lambda w . PERC(e)(w) \wedge g(C_2)(e)(w) \wedge Experiencer(e, g(1))(w)$
$\wedge \forall w' \in DOX_{(g(1))(\tau(e)+)(w)} [\exists e' [P(e')(w) \wedge overlap(\tau(e), \tau(e'))]]$
$(\lambda e . \lambda w . home(e, \iota x. brother(x, toph)(w))(w) \wedge \tau(e) < t_0$
$\wedge \forall t \in T_C [\exists e'[home(e', \iota x. brother(x, toph)(w))(w) \wedge \tau(e') \subseteq t \wedge \tau(e') < \tau(e)]] \wedge \neg\exists t \in T_C [\exists e''[leave(e'') \wedge Agent(e'', \iota x. brother(x, toph)(w))(w)$
$\wedge \tau(e'') \subseteq t \wedge \tau(e'') < \tau(e)]]) =$

$\lambda e. \lambda w . PERC(e)(w) \wedge see(e)(w) \wedge Experiencer(e, toph)(w)$
$\wedge \forall w' \in DOX_{(toph))(\tau(e)+)(w)} [\exists e' [home(e', \iota x. brother(x, toph)(w'))(w') \wedge \tau(e') < t_0 \wedge \forall t \in T_C [\exists e''[home(e'', \iota x. brother(x, toph)(w)')(w) \wedge \tau(e'') \subseteq t \wedge \tau(e'') < \tau(e')]] \wedge \neg\exists t \in T_C [\exists e'''[leave(e''') \wedge Agent(e''', \iota x. brother(x, toph)(w'))(w') \wedge \tau(e''') \subseteq t \wedge \tau(e''') < \tau(e')]] \wedge overlap(\tau(e), \tau(e'))]].$

Combining the eventuality predicate derived in (17) with the one denoted by the *when*-clause gives us the proposition in (18) as the denotation of the final sentence in (7)/(16).

(18) $\lambda w . \exists e \exists e'[come\_home(e)(w) \wedge Agent(e, toph)(w) \wedge \tau(e) < t_0$
$\wedge PERC(e')(w) \wedge see(e')(w) \wedge Experiencer(e', toph)(w)$
$\wedge \forall w' \in DOX_{(toph))(\tau(e')+)(w)} [\exists e'' [home(e'', \iota x. brother(x, toph)(w'))(w')$
$\wedge \tau(e'') < t_0 \wedge \forall t \in T_C [\exists e'''[home(e''', \iota x. brother(x, toph)(w')(w') \wedge$
$\tau(e''') \subseteq t \wedge \tau(e''') < \tau(e'')]] \wedge \neg\exists t \in T_C [\exists e''''[leave(e'''')$
$\wedge Agent(e'''', \iota x. brother(x, toph)(w'))(w') \wedge \tau(e'''') \subseteq t \wedge \tau(e'''') < \tau(e'')]] \wedge overlap(\tau(e'), \tau(e''))]] \wedge overlap(\tau(e), \tau(e'))]$

Note that since the entire novel containing the text segment in (7)/(16) is told from the perspective of the ego-narrator, Dave, the proposition in (18) is intersected with the set of worlds compatible with Dave's beliefs.

Let us consider the text segment in (12) next, repeated here as (19):

(19) [The T-Rex] hesitated. Maybe the little dinosaurs had hidden themselves in the cave on his left. When Billy looked up in his hiding place a few seconds later, [a T-Rex] bent down to the entrance of the cave and squinted into the dark.

Concerning the main clause of the third sentence in (12)/(19), the conditions for the insertion of the covert VS-operator are likewise fulfilled. First, the free individual variable can be resolved to Billy since being the agent of the event introduced by the *when*-clause makes him locally highly prominent. Second, the required information concerning the further specification of the perceiving eventuality can be inferred on the basis of the information provided by the *when*-clause in combination with the content of the clause itself, since it is obviously an event of Billy seeing the event introduced by main clause. Third, there is an obvious connection between the perceiving eventuality and Billy's doxastic state at the time immediately following the run time of that eventuality, since prior to the perceiving eventuality he does not believe there to be a T-Rex that is bending down to the entrance of the cave and squinting into the dark. Finally, the main clause would receive an inadequate interpretation without the insertion of the covert VS-operator for the following reason: On the one hand, interpreting the indefinite as picking up the T-Rex referred to by a definite description in the opening sentence would violate the familiarity condition (Heim 1982) and/or *Maximize Presupposition*! (Heim 1991, Singh 2011). On the other hand, interpreting it as introducing a second, novel T-Rex would, first, make the text segment rather incoherent. Second, as shown by Grønn & Sæbø (2012), replacing *a* by *another* is almost obligatory or at least strongly preferred in such cases.

In (20) the (again slightly simplified) denotation of the main clause in the third sentence in (19) is given:

(20) $\lambda e.\ \lambda w\ .\ PERC(e)(w) \wedge see(e)(w) \wedge Experiencer(e, billy)(w)$
$\wedge\ \forall w' \in DOX_{(billy)(\cdot(e)+)(w)}[\exists e'\exists x[T\text{-}Rex(x)(w') \wedge bend\_down(e')(w')$
$\wedge\ Agent(e', x)(w') \wedge \tau(e') < t_0 \wedge overlap(\tau(e), \tau(e'))]]$

Combining the eventuality predicate in (20) with the one denoted by the *when*-clause and applying existential closure gives us the proposition in (21) as the denotation of the final sentence in (12)/(19).

(21) $\lambda w\ .\ \exists e\exists e'[look\_up(e)(w) \wedge Agent(e, billy)(w) \wedge \tau(e) < t_0\ \wedge PERC(e')(w) \wedge$
$see(e')(w) \wedge Experiencer(e', billy)(w) \wedge \forall w' \in DOX_{(billy)(\cdot(e')+)(w)}[\exists e''\exists x[T\text{-}$
$Rex(x)(w') \wedge bend\_down(e'')(w') \wedge Agent(e'', x)(w') \wedge \tau(e'') < t_0 \wedge$
$overlap(\tau(e'), \tau(e''))]] \wedge overlap(\tau(e), \tau(e'))]$

In the case of the text segment in (12)/(19), there is no ego-narrator. Rather, the story is told by a neutral, omniscient narrator. Consequently, I assume that the proposition in (21) is intersected with the set of worlds representing the Common Ground of narrator and reader.

Finally, consider the sentence in (14), repeated here as (22). Here, the conditions for the insertion of the covert VS-operator at the TP-node of the main clause are met again: First, being the agent of the event introduced by the *when*-clause, Mary is prominent enough for the free individual variable to be resolved to her. Second, a suitable value for the free variable ranging over eventuality predicates can be inferred on the basis of the content of the *when*-clause in combination with the one of the main clause itself. Third, the sentence (at least in the absence of an earthquake scenario) would receive an inadequate interpretation without the insertion of the VS-operator. Finally, there is an obvious connection between the perceiving eventuality *e* and Mary's doxastic state at the time immediately following the run time of *e*: Without having the internal sensation caused by a disturbance of her sense of balance resulting from the boat trip, Mary would not believe that the ground was shaking beneath her feet (a belief that will most likely not last longer than a few seconds). The sentence is thus interpreted as shown in (23).

(22) When Mary stepped out of the boat, the ground was shaking beneath her feet for a couple of seconds.

(23) $\lambda w . \exists e \exists e' [\text{step\_out\_of\_boat}(e)(w) \wedge \text{Agent}(e, \text{mary})(w) \wedge \tau(e) < t_0$
$\wedge \text{PERC}(e')(w) \wedge \text{internal\_sensation}(e')(w) \wedge \text{Experiencer}(e', \text{mary})(w)$
$\wedge \forall w' \in \text{DOX}_{(\text{mary})(\iota(e')+)(w)} [\exists e'' [\text{shake}(e'')(w') \wedge \text{Theme}(e'', \text{ground})(w')$
$\wedge \tau(e'') < t_0 \wedge \text{overlap}(\tau(e'), \tau(e''))]] \wedge \text{overlap}(\tau(e), \tau(e'))]$

In this section I have proposed a formal analysis of VS according to which it comes about via the insertion of a covert operator at the TP-node of the respective main clause. The proposed analysis is flexible enough to capture the two following kinds of cases: First, cases such as (7)/(16) and (14)/(22), in which the proposition *p* that is true in the respective protagonist's doxastic state is false with respect to the Common Ground of narrator and reader. Second, cases such as (12)/(19), in which *p* is true in the Common Ground of speaker and narrator as well, but would have to be replaced by a presuppositionally stronger one if it was added to that Common Ground directly. At the same time, the application of VS is not unconstrained: There has to be a reason for the insertion of the covert VS-operator, and the context has to make available suitable values for the free variables ranging over individuals and eventuality predicates, respectively. In the following Section, I will return to the question of which of the analyses of FID sketched in Section 2.1 is best suited to account for its being restricted to the root level.

3.2 Free Indirect Discourse

In Section 2.1 we have seen that FID in contrast to DD and ID as well as VS is only available at the root level, but not sentence-internally, as shown by contrasts like the one between (1a) and (4b-c) (repeated here as (24a) and (24c-d), respectively), on the one hand, and (4a) (repeated here as (24b)), on the other.

(24) a. On her way home, Mary heard a song by Kendrick Lamar that she liked on the radio. She would buy his new album tomorrow.
b. * When Mary heard a song by Kendrick Lamar that she liked on the radio on her way home, she would buy his new album tomorrow.
c. When Mary heard a song by Kendrick Lamar that she liked on the radio on her way home, she thought that she would buy his new album on the following day.
d. When Mary heard a song by Kendrick Lamar that she liked on the radio on her way home, she thought: "I will buy his new album tomorrow".

As already mentioned in Section 2.1, there are two different lines of analysis of FID: According to the first line, FID results from the presence of a second context in addition to the narrator's context, with all context-sensitive expressions except pronouns and tenses being interpreted with respect to that second context (Schlenker 2004; Sharvit 2008; Eckardt 2014, building on Banfiled 1982 and Doron 1991). According to the second line, FID is a special form of mixed quotation, with pronouns and tenses being unquoted (Maier 2015, to appear).

Let us first have a look at how the restriction of FID to the root level can be implemented in analyses of the first kind. Sharvit (2008) assumes that in FID a covert operator is inserted that is similar to a propositional attitude verb such as *think* or *say*, but does not quantify over worlds that are compatible with the respective attitude holder's beliefs, but rather over contexts *c* such that the respective protagonist does not rule out the possibility that she is in *c*. Consequently, FID is in effect analyzed just like ID, the only difference being that all context-sensitive expressions with the exception of pronouns and tenses receive a shifted interpretation. Contrasts like the one between (24a) and (24b) are thus unexpected, i.e. one would not expect the covert FID-operator to behave differently from propositional attitude verbs such as *say* or *think* with respect to the syntactic environments in which it can be inserted. Therefore, the restriction of FID to the root level would simply have to be stipulated in the form of a restriction on the insertion options of the covert FID-operator.

Schlenker (2004) assumes all sentences to be interpreted with respect to two contexts – the Context of Utterance (ν) and the Context of Thought (θ), where the former is the context where the respective sentence is uttered and the latter the context where the corresponding speech or thought act originates. Usually, the

two contexts are identical, but in principle they can diverge, with FID being a case in point. In FID, ν is the context of the omniscient narrator, and θ the context where the respective protagonist speaks or thinks. Since all context-sensitive expressions with the exception of pronouns and tenses (which are always interpreted with respect to ν) are lexically specified as being interpreted with respect to θ, they all receive a shifted interpretation in FID.

While the restriction of FID to the root level does not automatically follow from the analysis just sketched, it can be derived as follows: Since (with some well-known exceptions) only entire sentences correspond to speech acts, a speech or thought act distinct from the one performed by the narrator in telling the entire story can only be accommodated at the root level. Consequently, it is only with respect to entire sentences that ν and θ can come apart, not at the level of matrix clauses contained in complex sentences.

Eckardt (2014) assumes that a protagonist's context *c* can optionally be added to the speaker's or narrator's context *C*, with all context-sensitive expressions except pronouns and tenses being interpreted with respect to *c* whenever it is present. The restriction of FID to the root level could be derived from Eckardt's (2014) analysis along lines similar to those sketched for Schlenker's (2004) account: In order for the addition of *c* to be feasible, a speech or thought act of some contextually prominent protagonist has to be accommodated by the reader, and since (generally) only entire sentences correspond to speech or thought acts, the accommodation of speech or thought acts is restricted to the root level.

Let us finally turn to Maier's (2015, to appear) analysis of FID as a special form of mixed quotation, according to which FID involves the accommodation of a speech or thought act that is partially quoted by the narrator, with pronouns and tenses being unquoted. On this account, the restriction of FID to the root level could also be derived from the fact that (generally) only entire sentences correspond to speech or thought acts.

Given the discussion so far, there does not seem to be much to choose between the analyses of Schlenker (2004) and Eckardt (2014), on the one hand, and the one of Maier (2015, to appear), on the other, as far as their ability to account for the restriction of FID to the root level is concerned. In each case, the restriction can be derived from the combination of the two following factors: First, the reader needs to accommodate a protagonist's speech or thought act. Second, speech or thought acts (generally) correspond to entire sentences, not to finite matrix clauses that are part of complex sentences. Nevertheless, there is a difference: In Schlenker's (2004) and Eckardt's (2014) analyses, the connection between the speech or thought act to be accommodated and the additional protagonist's context remains rather indirect and obscure, i.e. it is not really part of the analysis. Rather, what is spelled out in formally precise terms is only the interpretive effect of the addition of a second context to the narrator's context. On

Maier's (2015, to appear) account, in contrast, the accommodated speech or thought act is an integral part of the analysis, since in its absence there simply is nothing to be quoted. For this reason and because an analysis in terms of mixed quotation captures the speech-like character of FID as opposed to VS most clearly, I tentatively adopt Maier's (2015, to appear) analysis of FID for the purposes of this paper and propose that its restriction to the root level results from the need to accommodate a speech or thought act to be partially quoted.

4 Directions for future research

As already observed by Montague (1969), perception verbs such as *see* and *hear* in addition to their standard veridical uses have non-veridical uses as well: A sentence such as (25), for example, has a reading on which it does not entail the existence of unicorns, i.e. it can be true in a situation in which Angela in fact saw a white horse that she mistook for a unicorn.

(25) Angela saw a unicorn.

The analysis argued for in this paper now opens up the possibility that such readings do not come about because of lexical ambiguity or because of the presence of an elided occurrence of *seems* (Montague 1969), but rather because of an insertion of the covert VS-operator[6]. Adopting this view, the sentence in (25) would be interpreted as shown in (26), with the free individual variable being resolved to Angela and the free variable ranging over eventualities being resolved to the predicate *λe. see(e)*. On this reading, the sentence can be paraphrased as follows: Mary saw something and directly afterwards she believed that she saw a unicorn, which seems to be adequate.

(26) $\lambda w . \exists e[\mathrm{PERC}(e)(w) \wedge \mathrm{see}(e)(w) \wedge \mathrm{Experiencer}(e, \mathrm{angela})(w) \wedge \forall w' \in \mathrm{DOX}_{(\mathrm{angela})(\iota(e)+)(w)} [\exists e' \exists x[\mathrm{unicorn}(x)(w') \wedge \mathrm{see}(e')(w') \wedge \mathrm{Experiencer}(e', \mathrm{angela})(w') \wedge \mathrm{Theme}(e', x)(w') \wedge \tau(e') < t_0 \wedge \mathrm{overlap}(\tau(e), \tau(e'))]]]$

Such an analysis is highly attractive insofar as it reduces what at first sight appears to be a special property of a restricted class of verbs to a more general phenomenon. It remains to be seen, however, whether the conditions under which non-veridical uses of perception verbs are available are identical to those under which VS is available in narrative texts. Since a detailed discussion of this matter is beyond the scope of the present paper, I leave it as a topic for future research.

[6] I am grateful to Ede Zimmermann for pointing out to me this potential connection between my research on VS and non-veridical uses of perception verbs.

References

Abbott, Barbara & Laurence R. Horn. 2011. Nonfamiliarity and indefinite descriptions. Paper presented at the *Annual Meeting of the Linguistic Society of America*, Pittsburgh.

Banfield, Ann. 1982. *Unspeakable Sentences*: *Narration and Representation in the Language of Fiction*. Boston: Routledge.

Dancygier, Barbara. 2012. *The Language of Stories: A Cognitive Approach*. Cambridge: Cambridge University Press.

Dirscherl, Fabian & Jürgen Pafel. 2015. Die vier Arten der Rede- und Gedankendarstellung: Zwischen Zitieren und Referieren. *Linguistische Berichte* 241, 3–47.

Doron, Edit. 1991. Point of view as a factor of content. In S. Moore & A. Z. Wyner (eds.), *Semantics and Linguistic Theory* (SALT) I, 51–64. Ithaca, NY: Cornell University.

Eckardt, Regine. 2014. *The Semantics of Free Indirect Discourse*. *How Texts Allow to Mind-read and Eavesdrop*. Leiden: Brill.

Eckardt, Regine. 2017. Perspective and the future-in-the-past. *Glossa*: *A Journal of General Linguistics*. 2(1), 71. DOI: http://doi.org/10.5334/gjgl.199

Grønn, Atle & Kjell Johan Sæbø. 2012. A, the, another: A game of same and different. *Journal of Logic, Language and Information* 21(1), 75–95.

Hamburger, Käte. 1968. *Die Logik der Dichtung*. Stuttgart: Ernst Klett.

Harweg, Roland. 1972. Die reduzierte Rede. *Linguistics* 80, 44–55.

Heim, Irene. 1982. The Semantics of definite and indefinite noun phrases. Amherst, MA: University of Massachusetts dissertation.

Heim, Irene. 1991. Artikel und Definitheit. In A. von Stechow & D. Wunderlich (eds.), *Semantik*: *Ein internationales Handbuch der zeitgenössischen Forschung*. Berlin: Mouton de Gruyter, 487–535.

Kaplan, David. 1989. Demonstratives. In J. Almog, J. Perry & H. Wettstein (eds.), *Themes from Kaplan*, 565–614. Oxford: Oxford University Press.

Landman, Fred. 2000. *Events and Plurality*. Dordrecht: Kluwer.

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

Maier, Emar. To appear. The pragmatics of attraction: Explaining unquotation in direct and free indirect discourse. In P. Saka & M. Johnson (eds.), *The Semantics and Pragmatics of Quotation*.

Montague, Richard. 1969. On the nature of certain philosophical entities. *The Monist* 53(2), 159-194.

Passonneau, Rebecca J. 1994. Frame shifts: Indefinite reference to familiar referents. In *the Conference on Focus and Natural Language Processing. Working Paper 8 of The Institute for Logic and Linguistics*, volume 3: Discourse.

1986. Über den Personenwechsel und den anderer deiktischer Kategorien in der wiedergegebenen Rede. *Zeitschrift für Germanistische Linguistik* 13(3), 284–308.

von Roncador, Manfred. 1988. *Zwischen Direkter und Indirekter Rede*. *Nicht Wörtliche, Direkte Rede, Erlebte Rede, Logophorische Konstruktionen und Verwandtes*. Tübingen: Niemeyer.

Sharvit, Yael. 2008. The puzzle of free indirect discourse. *Linguistics and Philosophy* 31, 353 - 395.

Singh, Raj. 2011. Maximize Presuppositions! and local contexts. *Natural Language Semantics* 19, 149–68.

Schlenker, Philippe. 2004. Context of thought and context of utterance. A note on free indirect discourse and the historical present. *Mind and Language* 19(3), 279–304.

Stalnaker, Robert C. 1978. Assertion. In C. Peter (Ed.), *Syntax and Semantics 9: Pragmatics*, 315–332. New York: Academic Press.

Stalnaker, Robert C. 2002. Common ground. *Linguistics and Philosophy* 25, 701–721.

Stanzel, Franz. 1979. *Theorie des Erzählens*. Göttingen: Vandenhoeck & Ruprecht.

Vandelanotte, Lieven 2009. *Speech and Thought Representation in English: A Cognitive-functional Approach*. Berlin: de Gruyter.

Literary works cited:

Eggers, David. 2000. *A Heartbreaking Work of Staggering Genius*. New York: Simon & Schuster.

Wright, Richard. 1979. Big Boy leaves home. In L. D. Rubin, Jr. (ed.), *The Literary South*. New York: JohnWiley & Sons.