Computing Perspective Shift in Narratives*

Márta Abrusán
Institut Jean Nicod, Département d'études cognitives
ENS, EHESS, PSL Research University, CNRS, Paris France

October 5, 2020
to appear in
E. Maier & A. Stokke (eds.), *The Language of Fiction*, OUP

1 Introduction

Natural language allows changing the point of view in narrative texts without overt perspective shifting operators. A well-known example of such a perspective shift is *free indirect discourse* (FID). In this style, the hearer (reader) has to interpret certain elements of the sentence from the perspective of the narrator, and other elements from the perspective of a protagonist. But how do hearers (readers) know that they need to change the point of view in the first place? And when there are reasons to believe that the point of view is not that of the narrator, how do they know whose perspective is being developed?

Importantly, sentences in isolation often do not carry any indication of whether they should be understood under a shifted perspective or not (cf. Wiebe 1994). This is the case, for example, in (1a). When placed in a larger context, however, an interpretation of the same sentence with respect to some character's perspective might become plausible, as in (1b), or unlikely, as in (1c):

(1) a. It was raining.

*Thanks to the editors for their helpful suggestions as well as the audience at the Workshop on the Language of Fiction in Uppsala for their questions, comments and ideas. The research reported here was supported by the institutional grants ANR-10-LABX-0087IEC and ANR-10-IDEX-0001-02PSL.

b. Mary looked out the window anxiously. It was raining.
c. The barbecue was canceled. It was raining.

Thus elements in the broader discourse context can license the shift in perspective, even in the absence of any indication of perspective shift inside the target sentence. Nevertheless, there are also sentence-internal markers that strongly signal that a subjective perspective is being expressed. In the example below, the exclamation mark, the evaluative adverb and the progressive all suggest that the sentence conveys some protagonist's opinion:

(2) Mary was being absolutely impossible!

But note that even in these cases the identity of the perspectival center cannot be identified from the sentence alone. The best we can say about (2) is that the perspectival center is not Mary. To find out the identity of perspectival centres, the larger context is indispensable.

The aim of this paper is to get a better understanding of the principles that tell us (a) Whether a perspective shift occurs and (b) When there is a perspective shift, who is the perspectival center? These questions are rarely asked in the literature, but there are two notable exceptions I am aware of: The pioneering work of Wiebe (1990, 1994) that addresses both questions from a computational perspective and Hinterwimmer's (2019) analysis examining the second question with the tools of pragmatics. After reviewing these two proposals, I add a few observations of my own that show the importance of discourse structure. Finally, I sketch a framework based on SDRT (Asher and Lascarides 2003) extended with MSDRT/ADT (Kamp 2015; Kamp and Bende-Farkas 2018, Maier 2015a, 2016) that allows us to express the insights made in Wiebe (1994) and Hinterwimmer (2019) as well as the observations pertaining to discourse structure in a unified way.

A note on terminology: In this paper I understand *perspective shift* primarily as referring to FID. I assume that the mechanisms proposed could be extended to cases of what has been recently called *protagonist projection* (PP) or *viewpoint shift*, see Holton (1997), Stokke (2013, this volume), Hinterwimmer (2017), Abrusán (2020). However, I will not discuss such examples in this paper.[1] Perspective shift is also often used in connection with the interpretation of certain lexical items that have a subjective aspect to their meaning, e.g. predicates of personal taste, speaker-oriented adverbs, etc. When these elements appear outside of

[1] I also do not look at cases of the historical present, but see Anand and Toosarvandani (to appear).

direct or indirect discourse they are often markers of FID or PP. In such cases, at least, finding their perspectival center is the same question as finding the perspectival center of FID or PP.[2]

2 Wiebe's (1990, 1994) algorithm

Wiebe's (1990, 1994) groundbreaking work developed a computational algorithm that could recognise and track point of view in naturally occurring narrative discourse.[3] Point of view is important for understanding the meaning of what Wiebe (1994) calls *subjective sentences*, namely attitude reports and sentences in FID. While in the case of attitude reports the identity of the individual whose perspective is expressed is often given by the subject, finding the perspectival center of examples of FID is highly non-trivial.

The algorithm is based on regularities in the ways in which authors initiate, continue and resume a character's point of view. Based on a combination of the features of the target sentence (e.g. its tense, aspect, the presence of elements that express subjective meaning, etc.) and properties of the context (e.g. whether the previous sentence was subjective, the identity of a likely perspectival center, etc.), many point of view operations can be predicted by Wiebe's algorithm. Below I provide a summary of this algorithm, but note that this brief discussion cannot do justice to the richness of Wiebe's proposal.

The overall idea is captured by the Point of View (POV) function: This function maps the set of sentence features (*Feature Set*) and the current context (*Context*) into an *Interpretation*. The latter tells us whether or not the sentence is subjective and if yes, the identity of its perspectival center.

(3) Point of view (POV) function:
Feature Set $\times$ *Context* $\rightarrow$ *Interpretation*

In what follows I spell out the contents of the *Feature Set*, the *Context* and the *Interpretation*, respectively. I then illustrate the algorithm with an example.

[2]This does not mean that perspective shift is always PP or FID, see also Stokke (this volume) and Abrusán (2020) for discussion.

[3]Wiebe's early work on FID and other subjective sentences is unfortunately little known among linguists. One of the aims of this paper is to make this work more widely known among linguists and philosophers. Later, Janyce Wiebe (1959-2018) became one of the pioneers of the fields of sentiment analysis and opinion mining.

2.1 The *FeatureSet*

The *FeatureSet* contains many diverse pieces of information. The most important are the *potential subjective elements* in the sentence (more about these below). Another type of information contained in the *FeatureSet* is the type of state of affairs the sentence is about, for example whether it describes a private state action such as *looking and sighing* or some other types of action; or whether it is about a private state (i.e. an attitude, e.g. wondering, being afraid) or non-private state such as being *being six feet tall*. The *FeatureSet* also registers whether the head noun of the subject of the main clause is a private state noun (e.g. *pain, astonishment, etc.*) as well as the experiencers and actors of private states and private state actions. If the sentence contains a narrative parenthetical, the *FeatureSet* also registers the identity of the subject of the parenthetical. It further contains some other types of syntactic information as well.

The list of *potential subjective elements* in the sentence is largely based on Banfield's (1982) list of subjective elements, with the difference that for Wiebe (1994) these items are not unambiguous marks of subjectivity, they only mark subjectivity given certain contexts and with varying degrees of strength. Here is the list of *potential subjective elements* for English:[4] (NB: Similar lists of perspective sensitive items can be found in e.g. in Eckardt (2014, this volume) and Bylinina et al. (2014). See also Stojanovic (this volume) for a discussion of a subclass of evaluative nouns, viz. derogatory terms.)

(4) *Potential subjective elements* (PSEs)
- a. exclamations and direct questions
- b. elements that express evaluation or judgment
  - (i) adjectives (*awful, poor*)
  - (ii) nouns (*old bag*)
  - (iii) adverbs (*oddly, incredibly*)
  - (iv) modals that express judgment or obligation (*had better, ought to, should, be supposed to*)
  - (v) adverbs such as *scarcely* and *hardly* (e.g. "She could hardly be expected to live there.")
- c. elements that express lack of knowledge
  - (i) subordinators such as *whoever, whatever* (e.g. "Whatever it was, it has flown by quickly.")

[4] In languages other than English, tenses and aspect can be FID triggers, e.g. the imparfait in French, or—in the case of free indirect speech—the Konjunktiv in German.

(ii) adjectival phrases such as *some kind of* (e.g. "The object in her hand was some kind of weapon.")

d. sentence fragments
e. kinship terms
f. evidentials (that express certainty, or uncertainty, hedges, evidentials that address expectations, met or unmet)
g. certain adverbial discourse connectives such as *first, in addition, for instance, on the other hand, after all, anyway, yet,* etc. (e.g. "Yet, they were the pride of the family.")
h. conditional clauses
i. comparative *like* (e.g. "They followed her like acolytes behind a goddess.")
j. habitual sentences
k. the past perfective but only in the main verb phrase
l. the progressive, but only in the main verb phrase

The potential subjective elements have different strengths. The weakest ones, such as the past perfective and the progressive can typically only continue a character's POV and only within a paragraph. Stronger elements can continue or resume a POV after a paragraph break. Still stronger ones (e.g. evidentials and sentence fragments) can resume a last perspectival center's POV as long as they are expected perspectival centers. Finally, the strongest subjective elements, such as exclamations and questions, are always subjective, even when there is no expected character to whom we can attribute a sentence.

How do we know the target sentence's situation in the larger context and whether there is an expected perspectival center? This information is given by the *Context*, to which we turn below.

2.2 The *Context*

The *Context* consists of (a) the identity of the perspectival center (PC)[5] of the last subjective sentence that appeared in the text, if there was one, (b) the identity of the last active character, if there was one (c) the identity of any characters whose point of view was taken earlier in the text and (d) the current Text Situation:[6]

[5]Wiebe uses the terminology of *subjective character* for what I call perspectival center (PC). I have changed it to PC in this discussion in order to be consistent with the rest of the paper.

[6]This approach of tracking potential subjective centers is inspired by Grosz and Sidner's (1986) Centering Theory. As Wiebe points out however, although anaphora resolution is necessary for

(5) $Context = \langle LastPC,\ LastActiveCharacter,\ PreviousPCs,\ TextSituation \rangle$

The *Text Situation* describes whether the current sentence has been preceded by a subjective sentence or a sentence with an active character or a paragraph break has preceded the current sentence. Accordingly, Wiebe (1994) distinguishes four types of text situations:

(6) *Text Situations*
   a. continuing-subjective
   b. broken-subjective, interrupted subjective
   c. presubjective-active, postsubjective-nonactive, postsubjective-active
   d. presubjective-nonactive

The *Text Situation* interacts with the potential subjective elements in the following way. Expectations for a subjective sentence are strongest in continuing-subjective situations and weakest in the presubjective-nonactive situation. Even the weakest potential subjective elements are taken to be subjective in the first case, but only the strongest ones to be subjective in the last case. More precisely, each potential subjective element is associated with a set of text situations t such that the algorithm interprets the potential subjective elements to be subjective iff the current text situation is in t.

Finally, we can turn to the algorithm for identifying the perspectival center (PC): There are two cases: The first case is when the PC can be indentified on the basis of the target sentence alone, as in many attitude reports or examples of FID that contain parentheticals. The algorithm that summarises the procedure is given in (7) (adapted from Wiebe and Rappaport 1988):[7]

(7) *Identifying the PC from the target sentence:*
   If the sentence contains a narrative parenthetical then
   PC is the subject of the parenthetical
   else if the sentence is a private-state sentence then
   if it has a non-subordinated subjective element[8]
   or the text situation is continuing-subjective then PC is identified from the

tracking point of view, it is not sufficient.

[7]NB: Wiebe's (1994) primary aim is to identify FID in texts. The interpretation of unshiftable indexicals (e.g. first person indexicals) is (presumably) not influenced by the PC, but fixed by the global context.

[8]This would be the case for example in *Mary hardly wanted to leave*, in which the subjective adjective *hardly* indicates that the PC is not Mary.

previous context
else PC is the experiencer
end if
else PC is identified from the previous context
end if

When the PC cannot be identified from the target sentence, it needs to be identified based on the previous context.

(8) *Identifying the PC from the previous context:*
If there are two expected perspectival centers then
if the sentence is about the last active character then PC is the last perspectival center
else PC is the last active character
end if
else if there is an expected perspectival center then
PC is the expected perspectival center else PC is unidentified
end if

The expected perspectival centers (EPCs) are the Last PC and the last active character of the current context.

2.3 The *Interpretation*

The Interpretation functions tells us either that the sentence is a subjective sentence of a particular character, or that the sentence is objective[9] and has a particular active character (*ActiveCharacter* might also be the empty set).

(9) $Interpretation \in$
$\{\langle subjective,\ PC \rangle \mid PC \subseteq Characters\} \cup$
$\{\langle objective,\ ActiveCharacter \rangle \mid ActiveCharacter \subseteq Characters\}$

[9]Wiebe (1994) does not define what it means exactly for a sentence to be 'objective', but it is used in opposition with 'subjective', which is taken to represent the point of view of a character in the story. The narrator is not assumed to be character of the story (except in first person narration) and there is no PC associated with the narrator in objective sentences.

2.4 An example

(10) (a) The eyes were an incredibly light blue, like the sea with sunlight touching the waves. (b) Lemech greeted him respectfully. (c) "Adnarel, we thank you."(d) Then he said to Sandy, "The seraph will be able to help you. Seraphim know much about healing."
(e) So this was a seraph. (6) Tall, even taller than the twins. *(L'Engle, Many waters, cited in Wiebe (1994), p.28)*

Assume that the situation is *continuing-subjective* at the beginning, and Sandy is the last perspectival center. The sentence in (a) continues Sandy's subjective context, because it contains the subjective elements *incredibly* and the comparative *like*. The sentences in (b)-(d) are objective (note: direct discourse is ignored by the algorithm). Since there is a paragraph break before(e), the Text Situation is set to *postsubjective-nonactive*, with the *Last Perspectival Center* being Sandy). The sentence in (e) is determined to be Sandy's subjective sentence, because it contains *so* used as a conjunct, which is subjective as long as there is an expected perspectival center.

2.5 Tests of Wiebe's (1994) algorithm

The algorithm was tested on 450 sentential input items, which were complete sentences of every 5th page of two novels. The output of the POV function was judged against author's evaluations. Of the 450 input items, the algorithm made 27 primary errors, i.e. errors that were not dependent on other errors. Of these, 20 actual subjective sentences were interpreted as objective (13), or subjective but with wrong PC (7) and 7 actual objective sentences were interpreted as subjective. A very impressive result, overall.

2.6 Later work on probabilistic classifiers for POV

Wiebe and Bruce (2001) present a probabilistic approach for tracking point of view. This work, however, is much more limited in scope than Wiebe (1990, 1994). The underlying idea is that texts can be segmented into blocks such that all subjective sentences within the block express the point of view of the same character. The computational task that they define is segmenting the text into such blocks using a probabilistic classifier. The hypothesis for finding the subjective agent for each block is that these characters tend to enjoy a higher level of focus,

i.e. they will appear more frequently. The subject noun phrases of the main clauses of each blocks are therefore extracted and the perspectival center of each block is predicted from these.

The idea that blocks of texts tend to have the same perspectival center finds an echo in the work of Hinterwimmer (2019), to which I turn now.

3 Hinterwimmer (2019): Prominent protagonists

Hinterwimmer (2019) asks a more modest question than Wiebe's extremely ambitious proposal. Assuming that we already know whether a given sentence is to be interpreted as FID, he asks what conditions need to be met for a protagonist in order to be identified as the perspectival center.

He points out that not all protagonists prominent enough to be taken up by personal pronouns are available as anchors for FID. (This point was also noted in Wiebe 1994.) Observe the examples cited by Hinterwimmer (2019):

(11) *FID*

a. Susan looked at George hatefully. The dumb jerk had managed to make her look like an idiot at the meeting yesterday.

b. Susan looked at George hatefully. #The mean old hag had managed to make him look like an idiot at the meeting yesterday.

(12) *Pronoun Resolution*

a. Susan looked at George hatefully. She quickly turned away when he returned the look.

b. Susan looked at George hatefully. He returned the look, then turned his back on her and walked away.

The protagonist *George* is salient enough to be picked up by a personal pronoun in (12b), but not to be a perspectival center in (11b).[10]

Hinterwimmer (2019) proposes that licit perspective takers are the protagonists that either are sentient with respect to the eventuality introduced by the verb of the immediately preceding sentence or function as topics with respect to larger text segments. He thus proposes two strategies for identifying a discourse referent

[10]Intuitions about the availability of George as a perspectival center in (11b) are not completely clear, though. Andreas Stokke (pc.) finds that the example becomes acceptable if the context makes it probable that George hates Susan as well.

as a perspectival center: one based on *local prominence* and one based on *global prominence*.

Local prominence: Protagonists that have the highest number of agentivity features in the preceding sentence are prominent. Agentivity thus is more important than subjecthood. This can be seen in the case of object experiencer verbs: In (13) below the experiencer *Mary* is salient for being the prespectival center of the following sentence, is but the subject *John* is not.

(13) George bore Mary to death.
   a. Tomorrow she would definitely avoid sitting at a table with the bloated idiot again.
   b. #How sleepy she looked today!

Thus the character who is sentient with respect to the eventuality introduced by the verb of the immediately preceding sentence is the preferred perspectival center of the event to be accommodated as FID.

Global prominence: Hinterwimmer (2019) argues that in addition to local prominence, global prominence, i.e. topichood also matters for choosing the perspectival center. (Topichood is understood in the sense of QUD theories, cf. Roberts 2011, 2012.) To illustrate the importance of topichood, he contrasts the two following examples:

(14) George entered the room and looked around cautiously. Susan was sitting at the table in the corner with her best friend. Susan looked at George hatefully. The mean old hag had managed to make him look like an idiot at the meeting yesterday.

(15) Susan was sitting at a table in the corner with her best friend. George entered the room and looked around cautiously. Susan looked at George hatefully. #The mean old hag had managed to make him look like an idiot at the meeting yesterday.

The discourse in (14) is a story about *George* in the sense that *George* is felt to be its discourse topic. This is why *George* is available as the perspectival center for the last sentence of (14). The discourse in (15), on the other hand, seems to be about *Susan*, i.e. *Susan* is its discourse topic. This is why *George* is not available for being the perspectival center, despite of the fact that the sentence immediately

preceding the FID passage is identical in both cases.

In sum, Hinterwimmer (2019) complements Wiebe (1990, 1994) with two very interesting observations. The first is that characters with the highest number of agentivity features are the (locally) expected likely candidates for being perspectival centers.[11] The observation about the effect of topicality allows a more refined characterisation than a simple chunking approach of Wiebe and Bruce (2001) (though presumably it is much harder to track computationally).

Hinterwimmer's (2019) proposal was recently tested empirically by Bimpikou (2020). She found that locally prominent characters are more salient than globally prominent ones. At the same time, narrators are more salient than globally prominent characters but not more salient than locally prominent characters.

4 Rhetorical relations and FID

In this section I propose that rhetorical relations might be informative for inferring that perspective shift has taken place as well as for finding the perspectival center. In particular, certain types of discourse relations favour the interpretation of the target sentence from a shifted perspective, while others make it less likely or favour an unexpected perspective holder. The relevant division seems to be the distinction between subordinating and coordinating relations. A non-exhaustive list of each type of relation is given below, drawn from Asher and Vieu (2005).

(16) a. **Subordinating:** Elaboration, Instance, Topic, Explanation, Precondition, Commentary.
b. **Coordinating:** Narration, Background, Result, Continuation, Parallel, Contrast.

A good approximation of the difference between the two major types of relations is that coordinating relations tend to move the story forward, while subordinating relations do not; rather, sentences linked with the latter relations tend to expand on a given issue.

[11] The preference for previous subjects and active characters was already noted in Wiebe 1994, but Hinterwimmer's 2017 suggestion is more explicit.

4.1 Empirical observations

The first observation is that when an objective sentence is followed by a potentially perspective shifted (FID) sentence, subordinating relations make the interpretation as FID more likely, while coordinating relations favour a non-shifted interpretation. This is in accordance with Caenepeel's (1989) observation according to which perspectivally non-situated sentences typically move narrative time forward, while perspectivally situated sentences do not convey forward movement in time (see also Cumming this volume). This is because sentences connected with coordinating but not with subordinating relations tend to move the narrative time forward. Observe the example below:

(17) Mary started to cry.$_{\pi 1}$ She was fired again.$_{\pi 2}$

a. *explanation*(π_1,π_2): FID is salient
b. *result*(π_1,π_2): no perspective shift.

The second observation is that when a sentence linked with a coordinating relation is nevertheless understood as FID, the perspectival center tends to be a character who is not the expected perspective taker (in the sense of Wiebe or Hinterwimmer).[12]

(18) Sam threatened Justin with a knife.$_{\pi 1}$ He had to defend himself!$_{\pi 2}$

a. Explanation(π_1,π_2): PC=Sam, he=Sam[13]
b. Result(π_1,π_2): PC=Justin, he=Justin

(19) John handed a book to Bill.$_{\pi 1}$ He had to read this.$_{\pi 2}$

a. Explanation(π_1,π_2): PC=John, he=Bill
b. Result(π_1,π_2): PC=Bill, he=Bill.

In both examples, the subject of the first sentence is the expected perspective center: It is the last active character (Wiebe 1994) and also the character with the most agentivity features (Hinterwimmer 2019). When the target sentence is understood as being linked with the subordinating relation Explanation, it is the subject of the previous sentence that can function as the perspective holder. When the second sentence is understood as being linked with the co-ordinating relation Result, the preferred perspective center is the (indirect) object, which is not the expected perspective holder for either.

[12] Many of the examples below are inspired by Kehler et al. (2008).

[13] Possibly, PC: Justin, if he is reasoning about Sam(=he)

As for sentences in FID linked with a subordinating relation such as Explanation, their PC is usually the expected PC, but not always. Observe first two examples in which the character with the most agentive features is indeed the PC of the following sentence.

(20) Jane hit Mary.$_{\pi 1}$ She had stolen a tennis racket!$_{\pi 2}$

a. Explanation(π_1,π_2): PC=Jane, she=Mary

(21) Jane angered Mary.$_{\pi 1}$ She had stolen a tennis racket!$_{\pi 2}$

a. Explanation(π_1,π_2): PC=Mary, she=Jane

However, sometimes the rhetorical relation Explanation can co-occur with a shift of the PC as well: this is the situation if we interpret the second sentence of the next example as in (22b):

(22) Sam threatened Justin with a knife$_{\pi 1}$. He lost control.$_{\pi 2}$

a. Explanation(π_1,π_2): PC=Sam, he=Sam
b. Explanation(π_1,π_2): PC=Justin, he=Sam
c. Result/Narration(π_1,π_2): no perspective shift

An interesting question is what happens when FID sentences are followed or preceded by a parenthetical (*blabla..., he thought*). One possibility is that it is the parenthetical that is linked to the preceding discourse with a discourse relation (cf. Bimpikou et al. 2020, Cumming this volume). The embedded clause is linked to the parenthetical clause with the subordinating Attribution relation. In such cases, it is the relation between the parenthetical and the previous discourse that influences the choice of PC, since the attitude holder is the PC:

(23) Sam threatened Justin with a knife.$_{\pi 1}$ Next he would threaten Paul!$_{\pi 3}$(,he thought$_{\pi 2}$.)

a. Elaboration(?)(π_1,π_2): PC=Sam, he in π_2=Sam.
Attribution (π_2,π_3), he in π_3=Sam
b. Result (π_1,π_2): PC=Justin, he in π_2=Justin.
Attribution (π_2,π_3) he in π_3=Sam.

But in some cases, as discussed by Hunter (2016), it is the embedded clause that is linked with a (possibly modalised) rhetorical relation to the previous discourse. For example:

(24) a. [John didn't come to my party.]$_{\pi 1}$ [Jill said]$_{\pi 2}$ [he was out of town.]$_{\pi 3}$

b. $\diamond$Explanation (π_1,π_3), Attribution (π_2,π_3)

We can easily create similar examples with FID, see e.g. (25), and therefore we should keep open the possibility of rhetorically linking embedded clauses directly to the previous discourse:

(25) a. [John didn't look at Mary.]$_{\pi 1}$ [He was mad at her]$_{\pi 3}$, [she thought]$_{\pi 2}$
b. $\diamond$Explanation (π_1,π_3), Attribution (π_2,π_3)

Further, as discussed by Hunter (2016) in connection with speech reports, we probably also need to allow cases in which the parenthetical and the embedded clause are linked to the preceding discourse with separate discourse relations.

One might ask if there is an implicit parenthetical in the examples of FID discussed above in (17a), (18)-(22), as suggested by Bimpikou et al. (2020). Suppose that this is the case. In these particular examples, the relation Explanation/Result then holds between the first clause and the parenthetical. (Another possibility is that the relation holds between the first clause and the complex unit formed by the parenthetical and the embedded clause, cf. Cumming this volume for discussion.)

A final empirical observation concerns rhetorical relations between sentences in FID. In such cases coordinating relations such as Narration, Parallel or Result do not induce a change of PC:

(26) Sam threatened Justin with a knife.$_{\pi 1}$ Next he would threaten Paul!$_{\pi 3}$ Then he would threaten Jane as well!$_{\pi 4}$(,he thought$_{\pi 2}$.)
a. Elaboration(π_1,π_2): PC=Sam;
or Result(π_1,π_2): PC=Justin,
Parallel(π_3,π_4)

In general, whoever is the PC of the FID passages, and whatever is the rhetorical relation between them, the PC has to be the same for the two sentences.

4.2 Summary of the empirical observations

Rhetorical relations between two continuous sentences in FID do not matter: the two sentences have the same PC. However, rhetorical relations between a narrative and a potentially FID sentence do matter for establishing perspective shift: Subordinating relations favour perspective shift, and the PC is typically the Expected PC. However, a change of PC is not excluded with subordinating relations. On the other hand, coordinating relations usually suggest the absence of perspective

shift; or, if perspective shift still happens, the PC is most likely different from the expected PC.

Note that these observations are compatible with Hinterwimmer's topicality idea. This is because sentences linked with subordinating relations tend to elaborate on the same topic, while sentences liked with coordinating relations move the story forward, and introduce new topics.

5 Putting it all together: towards a framework

Is there a way to incorporate the insights of the works discussed above as well as the observations about discourse relations in one unified theoretical framework? I propose that using SDRT (Asher and Lascarides 2003) extended with a representation of mental states (MSDRT/ADT, Kamp 2015, this volume, Kamp and Bende-Farkas 2018, Maier 2015a, 2016) allows us to do just this. Below I first provide a very brief introduction to these theories and then sketch a way to formulate the answers to our two leading questions.

5.1 Background on SDRT and MSDRT/ADT

SDRT (Asher and Lascarides 2003) is an extension of DRT (Kamp 1981 and subsequent work) that aims to express the semantics and structure of rhetorical relations within the discourse representation. One reason why this is important is to allow us to capture the constraints that govern anaphoric relations. That classic DRT is not sufficiently restrictive is shown by the following textbook example:

(27) a. John had a great evening (π_1). He had a great meal (π_2). He ate salmon (π_3). He devoured lots of cheese (π_4). Then he won a dancing competition (π_5).
b. # It was a beautiful pink(π_6).

If we try to follow up the discourse in (27a) with the sentence in (27b), the pronoun *it* cannot refer back to the discourse referent introduced by *salmon* in the previous discourse. But this is not predicted by classic DRT, in which the structure of the above discourse is flat and the discourse referent for *salmon* should be available as an antecedent for the pronoun, as shown in Figure 1.[14]

[14]NB: The DRSs are highly simplified in order to focus on the questions discussed in the paper.

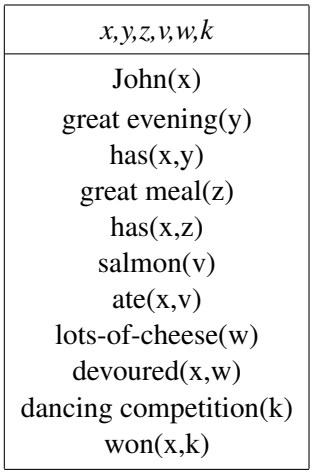


Figure 1: DRS of (27a)

Rhetorical structure thus imposes constraints on anaphora resolution (among other domains). To capture this effect, SDRT extends the language of DRSs by introducing two new types of expressions expressions: (a) Speech act discourse referents and (b) rhetorical relations that relate speech act discourse referents. The resulting structures are called segmented DRSs (SDRSs). The SDRS of (27a) is given in Figure 2: π are speech act discourse referents, Ks stand for the DRS's associated with the sentences that introduce speech act discourse referents, and rhetorical relations (e.g. Narration, Elaboration) are given as relations between speech act discourse referents.

The hierarchical relations captured by the embedding structure of the SDRS can be seen more clearly in a graph format, cf. Figure 3.

The main constraint on anaphora resolution that predicts why (27a) cannot be followed by (27b) is the Right Frontier Constraint (RFC). This constraint says that referents in the constituent to which the new sentence is attached to (the current constituent) are accessible as well as those that dominate the current constituent. To find antecedents, this constraint thus allows to look left on the graph one step only or to look up. Since (27b) attaches to π_5, the RFC limits the available antecedents to those that appear in π_5 or the node dominating it, π_1.[15]

Another extension of DRT that is relevant for the present discussion is MS-DRT, proposed recently by Kamp (2015); Kamp and Bende-Farkas (2018) and elaborated further in Maier (2015a, 2016) under the name ADT (see also Kamp

[15]The speech act discourse referent π_6 is constructed for the topic of the narration.

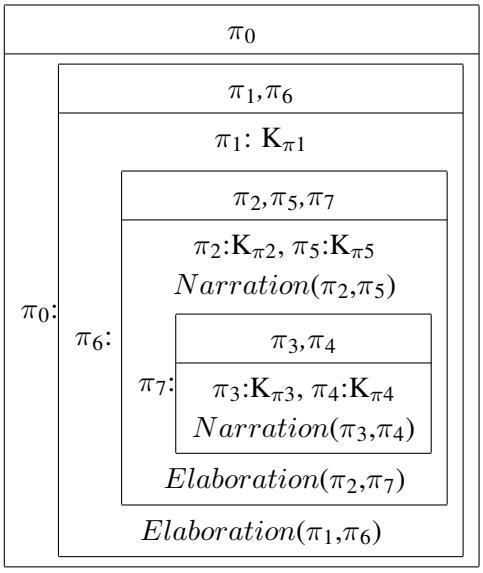


Figure 2: SDRS of (27a)

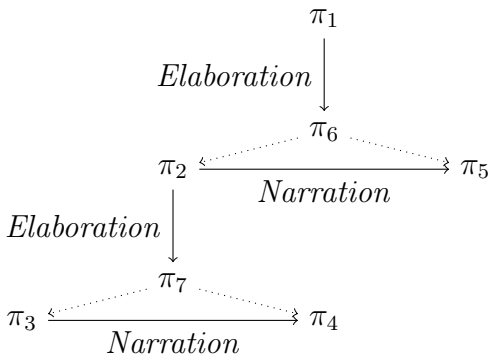


Figure 3: Graph of the SDRS for (27a)

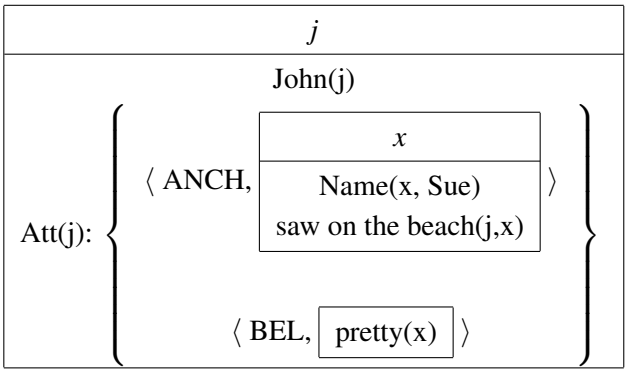
this volume). Its aim is to provide an accurate description of the semantics of propositional attitudes and the referring expressions embedded in their scope. The theory gives simple and elegant solutions to various classic puzzles of reference, e.g. cases of mistaken identity, etc.[16] For the purposes of the present paper, I will only introduce some of its most basic aspects that are relevant for the present discussion.

[16]MSDRT/ADT is not the first approach to these issues within DRT, see for example Asher (1986, 1987).

In MSDRT/ADT the semantic representation of attitude reports are representations formed by the interpreter about the mental state of the attitude holder. Observe the example (28a) and its ADT representation in (28b). The mental state represented in the DRS describes John being in the mental state consisting of a belief and an anchored entity representation of the individual *Sue*. The latter consists of a discourse referent for the entity represented and a description of how the attitude holder is related to the entity.[17]

(28) a. John thought that Sue was pretty.

b.



As it was discussed in Wiebe (1994), when we are trying to identify the PC of FID sentence, we can distinguish two types of cases. Either the PC can be determined from the sentence (as in the cases with examples with parenthetical), or the PC cannot be determined from the sentence alone. In the first case, the ADT representation of the sentence can be determined exactly as in (28b) above.

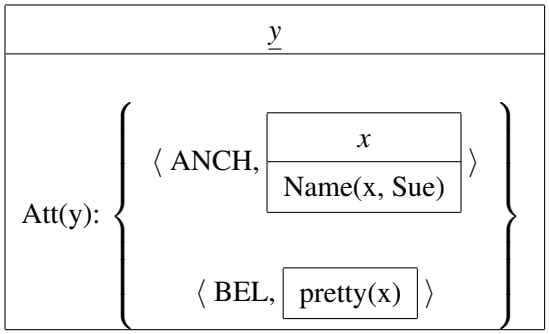
(29) Sue was pretty, John thought.

In the second case, namely when perspective shift is signalled by the sentence features and the context but the PC cannot be determined from the sentence alone, an attitude representation needs to be projected with a variable standing for the attitude holder. The task of finding out the PC is then the task of finding the value for this projected variable.

(30) a. Sue was pretty!

[17]The entity representation above is an internally anchored representation. In order to capture that there is indeed something that the attitude is about, an external anchor needs to be added to the representation. I omit this here for simplicity.

b.



NB: In many cases, the identity of the projected attitude has to be figured out as well. However, we might assume that *believe* is a good default candidate.[18]

5.2 A sketch of a proposal

The advantage of building on an SDRT+MSDRT/ADT representation of the text is that a lot of what we need to implement Wiebe's ideas (as well as the observations made by Hinterwimmer and myself) is already contained in the discourse representation. This includes: (a) The available discourse referents on the Right Frontier, which can take over the function of the list of Last Active Character in Wiebe's system. (b) The attitude holders (perspectival centers) for subjective sentences whose PC can be recovered from the sentence alone (attitude reports, parentheticals), or whose PC has already been determined (past cases of FID or other PS). These are the subjects of attitude predicates in the MSDRT/ADT representation. (c) the type of predicate (private-state, active, etc), and the roles of their arguments (experiencers, e.g.) as these are part of the semantic analysis of the DRS of the sentence. (d) information about discourse structure, e.g. whether a scene break has occurred and some other information that is captured by Wiebe's *TextSituation*. (e) If needed, information about topics can be represented in SDRT as well, along the lines proposed in Hunter and Abrusán (2016) (I come back to this issue below).

Thus the question of tracking subjective perspective in the text boils down to two issues: First, we need to decide whether a sentence is subjective or objective (=*Question 1*). If it is determined likely subjective, an MSDRT/ADT structure needs to be projected and we need to find the identity of the attitude holder

[18]For a theory of FID that assumes mixed quotation (Maier 2015b), inferring an implicit verb of saying might be more natural. But even then the attitude *believe* might have to be projected in a 'parasitic' fashion (unless one has reasons to assume that the character is lying).

(=*Question 2*). (Note: unlike for Wiebe, for objective sentences nothing in particular needs to be done as its the discourse referents are encoded anyway in the representation of the sentence.) There is thus a two-step procedure: First, when adding a new sentence to the previous discourse, decide about *Question 1*. If the answer to this question is that the sentence is subjective, then we try to determine the answer to *Question 2*.[19]

Question 1: How do we know we have perspective shift? Literary scholars have identified various triggers for FID (e.g. Banfield 1982, Fludernik 1993). Banfield (1982) called these *subjective elements*. This list was the main inspiration for Wiebe's (1994) list in her *FeatureSet* which she called *potential subjective elements*, realising that many linguistic features can contribute to signaling perspecitve shift but they rarely do so in an unambiguous and determinate way. Recent experimental results confirm the effect on perspective shifting for certain triggers, cf. Kaiser (2015) concerning epithets and certain adverbs. As it is well known, tenses and aspect can be FID triggers in many languages, e.g. the imparfait in French, or—in the case of free indirect speech—the Konjunktiv in German.

In the framework that I am using in this paper, the question 'How do we know that we have perspective shift?' translates as 'How do we know that we need to project an attitude layer in the MSDRT/ADT representation?'. I propose to follow Wiebe (1994) by looking at her *FeatureSet*, together with the *Text Situation*, the latter now understood as the state of the SDRT representation. In addition to textual transitions, *Text Situation* now also represents rhetorical relations. This allows us to encode that perspective shift is more likely with subordinating relations.

The question we are posing can be thought of as a classification problem: Given various features of the target sentence S and the preceding text, is S objective or subjective (i.e. understood under a perspective shift)? Given the nature of the contextual cues and the interaction of the various features, the answer ideally should be given by probabilistic classifier rather than by a rule-based system (cf. Wiebe and Bruce 2001). However, developing such a system would go beyond the scope of this paper. We might note, still, that one crucial issue would be deciding the right level of interaction between the various sentential and textual features.

[19]NB: sometimes the answer to these questions cannot be determined. For example, here is an example of a sentence in FID from Wiebe (1994) that appears as the first sentence of a novel:

(i) *Captain Scalawag's treasure!* This was the first thing Pete thought of when he got up.

The approach needs to be capable of handling such cases as well.

Question 2: How do we find the identity of the projected PC? If the sentence was found to be subjective, the second question is how to find its perspectival center. As discussed by Wiebe (1994), there are two basic cases. In the first case, the target sentence is an explicit attitude report (potentially in the parenthetical form). In this case the perspectival center can often be identified from the target sentence alone. In the second case, the target sentence itself does not allow us to identify the perspectival enter, it needs to be identified from the previous context.

Note that unlike the previous question, this question is not very natural to think of in terms of a simple probabilistic classification problem. This is because for a probabilistic classifier the possible values of the classification variable need to be set in advance and then each candidate is evaluated separately. However, we do not have the list of all potential perspectival centers before parsing the text (cf. Wiebe and Bruce 2001) and it is not natural to think of evaluating each candidate separately. It might make more sense to think in terms of a ranking algorithm that compares the most likely candidates for being the PC, similarly to some approaches for anaphora resolution (cf. Denis and Baldridge 2007). Nevertheless, in this paper I stick to assuming a rule-based model à la Wiebe (1994), but adding some information about discourse structure.

Let me review the first, easy case, the case of attitude reports. These are understood as providing information about the perspective of the attitude holder, except in cases where there are reasons to believe that the attitude report is itself the content of some other character's mental state. This might be the case e.g. in *Mary hardly wanted to leave*, in which the subjective adverb in the main clause might indicate that we need to infer someone else's perspective. Another exception is if the sentence continues a previously established subjective perspective. Here is then a slightly updated version of Wiebe's rule for identifying the PC from attitude reports:

(31) *Identifying the PC from the target sentence:*
If the sentence contains a narrative parenthetical then
PC is the subject of the parenthetical
else if the sentence is an attitude report then
if it has a non-subordinated subjective element,
or the text situation is rhetorically linked to a subjective sentence then PC is identified from the previous context
else PC is the attitude holder
end if
else PC is identified from the previous context

end if

The second, harder case is finding the PC if it cannot be identified from the target sentence alone. I propose an updated version of Wiebe's rule that makes reference to discourse structure constraints, rhetorical relations and also incorporates Hinterwimmer's observation relating to experiencers.

Let us first assume a set of expected PCs consisting of the Last PC and the most recent discourse referent on the Right Frontier (RF) that is an experiencer. Here is then a (highly tentative!) rule for finding the PC from the context:

(32) *Updated rule for finding the PC from the previous context*

a. the sentence connects with a subordinating relation:
If there are two expected PCs then
if the sentence is about the most recent experiencer DR on the RF, PC is the Last PC
else PC is the last experiencer DR on the RF
else if there is one expected perspectival center then
PC is the expected perspectival center
else PC is unidentified

b. the sentence connects with a coordinating relation to an objective sentence: PC is a DR on the RF that is not an expected PC.

Note that the above rule is not fully deterministic in the sense that it does not always give a unique output, e.g. in the case of sentences connecting with a co-ordinating relation there might be more suitable outputs.[20] Note also that the observation that coordinating relations indicate a shift away from the expected PC incorporates also the finding of Wiebe and Bruce (2001) according to which natural 'blocks' of text tend to have the same perspectival center. This is because coordinating relations often introduce new 'blocks' of discourse. Potentially, the difference between co-ordinating and subordinating discourse relations can also explain Hinterwimmer's observation about topicality. I turn to this next.

5.3 Topicality vs. rhetorical structure

Hinterwimmer (2019) argued that topicality itself is a factor for determining the perspectival center. One way of incorporating this insight into the current system is to introduce a representation of topicality into SDRT. This could be done along

[20]The rule above can only predict (22b) if Justin was the LastPC.

the lines of Hunter and Abrusán (2016), who associate questions under discussions with subgraphs (CDUs) of the SDRT graph. Another possibility though is to explore whether the phenomenon can be reduced to something else. This is what I tentatively suggest below.

Recall the examples that motivated Hinterwimmer (2019) to argue for the importance of topicality:

(14) George entered the room and looked around cautiously$_1$. Susan was sitting at the table in the corner with her best friend.$_2$ Susan looked at George hatefully.$_3$ The mean old hag had managed to make him look like an idiot at the meeting yesterday.$_4$

(15) Susan was sitting at a table in the corner with her best friend.$_1$ George entered the room and looked around cautiously.$_2$ Susan looked at George hatefully.$_3$ #The mean old hag had managed to make him look like an idiot at the meeting yesterday.$_4$

If we look closer, the Right Frontier constraint predicts the difference between the two examples as well. Notice that the rhetorical structure of the two examples is quite different: In (14) all the sentences following (1) elaborate on it and for this reason the Right Frontier is (1-3). George is the experiencer of (1), to which (4) connects. In (15), however, the right frontier is only (3), and George is not an experiencer of (3). This predicts that George will not be an available PC for (4).

If we change (15) in such a way that (3) is itself interpretable from the perspective of George, an interpretation of (4) with George as PC becomes marginally possible:

(33) Susan was sitting at a table in the corner with her best friend.$_1$ George entered the room and looked around cautiously.$_2$ Susan looked at *him* hatefully.$_3$ *?*The mean old hag had managed to make him look like an idiot at the meeting yesterday.$_4$

This is because in (33) now we have a continuing-subjective situation, and George is an expected PC when we get to sentence (4).

6 Conclusion

Readers (hearers) are able to figure out effortlessly if a perspective shift occurs in a discourse. They can also identify, in most cases, the likely perspective holder of the relevant text. How we are able to achieve this is a question that has been

rarely addressed in the linguistic literature, with the notable exceptions of Wiebe (1990, 1994) and Hinterwimmer (2019). This paper reviewed these proposals, added a few observations of my own about the importance of rhetorical structure and proposed to incorporate all the previous insights into one unified framework.

References

Abrusán, Márta (2020). The spectrum of perspective shift: protagonist projection versus free indirect discourse. *Linguistics and Philosophy*, https://doi.org/10.1007/s10988–020–09300–z.

Anand, Pranav and Maziar Toosarvandani (to appear). Narrative and point of view. In D. Altshuler (Ed.), *Linguistics meets Philosophy*. Cambridge University Press.

Asher, Nicholas (1986). Belief in discourse representation theory. *Journal of Philosophical Logic 15*(2), 127–189.

Asher, Nicholas (1987). A typology for attitude verbs and their anaphoric properties. *Linguistics and Philosophy 10*(2), 125–197.

Asher, Nicholas and Alex Lascarides (2003). *Logics of Conversation*. Cambridge University Press.

Asher, Nicholas and Laure Vieu (2005). Subordinating and Coordinating Discourse Relations. *Lingua 115*(4), 591–610.

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

Bimpikou, Sofia (2020). Who perceives? who thinks? anchoring free reports of perception and thought in narratives. *Open Library of Humanities 6*(2).

Bimpikou, Sofia, Emar Maier, and Petra Hendriks (2020). The discourse structure of free indirect discourse reports. ms. University of Groningen.

Bylinina, Lisa, Eric McCready, and Yasutada Sudo (2014). The landscape of perspective shifting. In *Talk given at the Workshop "Pronouns in Embedded Contexts at the Syntax-Semantics Interface". Universität Tübingen*.

Caenepeel, Mimo (1989). *Aspect, temporal ordering and perspective in narrative fiction*. Ph. D. thesis, The University of Edinburgh.

Cumming, Sam (this volume). Narative and point of view. In E. Maier and A. Stokke (Eds.), *The Language of Fiction*. Oxford: Oxford University Press.

Denis, Pascal and Jason Baldridge (2007). A ranking approach to pronoun resolution. In *IJCAI*, Volume 158821593.

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

Eckardt, Regine (this volume). In search of the narrator. In E. Maier and A. Stokke (Eds.), *The Language of Fiction*. Oxford: Oxford University Press.

Fludernik, Monika (1993). The fictions of language and the language of fiction. *The Linguistic Representation of Speech and Consciousness. London and New York: Routledge*.

Grosz, Barbara J. and Candace L. Sidner (1986). Attention, intentions, and the structure of discourse. *Computational linguistics 12*(3), 175–204.

Hinterwimmer, Stefan (2017). Two kinds of perspective taking in narrative texts. In *Semantics and Linguistic Theory*, Volume 27, pp. 282–301.

Hinterwimmer, Stefan (2019). Prominent protagonists. *Journal of Pragmatics 154*, 79–91.

Holton, Richard (1997). Some telling examples: A reply to Tsohatzidis. *Journal of Pragmatics 28*(5), 625–628.

Hunter, Julie (2016). Reports in discourse. *Dialogue and Discourse 7*(4), 1–35.

Hunter, Julie and Márta Abrusán (2016). Rhetorical relations and QUDs. In M. Otake, S. Kurahashi, Y. Ota, K. Aatoh, and D. Bekki (Eds.), *New Frontiers in Artificial Intelligence: JSAI-isAI Workshops LENLS, JURISIN, KCSD, LLLL Revised Selected Papers.*, pp. 41–57. Springer.

Kaiser, Elsi (2015). Perspective-shifting and free indirect discourse: Experimental investigations. In *Semantics and Linguistic Theory*, Volume 25, pp. 346–372.

Kamp, Hans (1981). A theory of truth and semantic representation. *Formal semantics-the essential readings*, 189–222.

Kamp, Hans (2015). Using proper names as intermediaries between labelled entity representations. *Erkenntnis 80*(2), 263–312.

Kamp, Hans (this volume). Sharing real and fictional reference. In E. Maier and A. Stokke (Eds.), *The Language of Fiction*. Oxford: Oxford University Press.

Kamp, Hans and Ágnes Bende-Farkas (2018). Epistemic specificity from a communication-theoretic perspective. *Journal of Semantics 36*(1), 1–51.

Kehler, Andrew, Laura Kertz, Hannah Rohde, and Jeffrey L Elman (2008). Coherence and coreference revisited. *Journal of semantics 25*(1), 1–44.

Maier, Emar (2015a). Parasitic attitudes. *Linguistics and Philosophy 38*(3), 205–236.

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

Maier, Emar (2016). Attitudes and mental files in discourse representation theory. *Review of philosophy and psychology 7*(2), 473–490.

Roberts, Craige (2011). Topics. In C. Maienborn, K. von Heusinger, and P. Portner (Eds.), *Semantics: An International Handbook of Natural Language Meaning*, Volume 2, pp. 1908–1934. Berlin/Boston.

Roberts, Craige (2012, December). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics 5*(6), 1–69.

Stojanovic, Isidora (this volume). Derogatory terms in free indirect discourse. In E. Maier and A. Stokke (Eds.), *The Language of Fiction*. Oxford: Oxford University Press.

Stokke, Andreas (2013). Protagonist projection. *Mind & Language 28*(2), 204–232.

Stokke, Andreas (this volume). Protagonist projection, character focus and mixed quotation. In E. Maier and A. Stokke (Eds.), *The Language of Fiction*. Oxford: Oxford Univ Press University Press.

Wiebe, Janyce M. (1990). *Recognizing subjective sentences: A computational investigation of narrative text.* Ph. D. thesis, Technical Report 90-03. (Buffalo: SUNY Buffalo Dept. of Computer Science).

Wiebe, Janyce M. (1994). Tracking point of view in narrative. *Computational Linguistics 20*(2), 233–287.

Wiebe, Janyce M. and Rebecca F. Bruce (2001). Probabilistic classifiers for tracking point of view. *PROGRESS IN COMMUNICATION SCIENCES*, 125–142.

Wiebe, Janyce M. and William J. Rappaport (1988). A computational theory of perspective and reference in narrative. In *Proc. 26th Annual Meeting of the Assoc. for Computational Linguistics (ACL-88)*, pp. 131–138.