Attribution and the discourse structure of reports

Emar Maier
University of Groningen

Abstract I propose a discourse-level analysis of report constructions. Indirect discourse, mixed and direct quotation, free indirect discourse, and attitude ascriptions are all analyzed in terms of a discourse relation of ATTRIBUTION, connecting two propositional discourse units corresponding to (i) a frame segment (*he said*, *she dreamed*) and a (possibly complex, multi-sentence) report (*"I'm an idiot"*, *(that) she was president*). I provide a unified semantics for the discourse relation of ATTRIBUTION that invokes a flexible notion of 'characterization'. A discourse unit may characterize a speech event by reproducing its linguistic surface form (as in quotation) or its propositional content (as in indirect speech and attitude reports), or some mixture of both (as in mixed quotation or free indirect discourse). I formalize this unified discourse-level ATTRIBUTION approach to reporting within the general framework of SDRT, and apply it to direct, indirect, and free indirect reports that extend beyond the single embedded or quoted clause. The resulting account is the first to do justice to the complex internal dependencies within stretches of reported discourse.

Contents

1   Introduction: discourse, coherence, and reporting

A correct interpretation of a multi-sentence discourse includes more information than is contained in the interpretations of its individual sentences taken in isolation. Take the mini-discourse in (1).

(1)   John was biking home late. A police officer stopped him. She give him a fine. His lights were off.

We naturally infer that a police officer stopped John *while* John was biking home late *and then* the police officer gave John a fine *because* John's lights were off. The individual sentences themselves describe states and events, which we as interpreters try to combine into a coherent discourse by inferring various causal, temporal and other relations between these states and events (Hobbs 1979). These coherence inferences are generally defeasible and constrained by rationality, world-knowledge, a finite inventory of potential discourse relations (NARRATION, BACKGROUND, ELABORATION, EXPLANATION, etc.), and linguistic cues (an overt connective like *and then* would signal NARRATION, *because* would signal EXPLANATION).

    Now say the story continues with a question like (2).

(2)   What was he thinking?

In principle, (2) could represent a (genuine or rhetorical) question of the writer to the reader, but, in the given narrative context, another likely interpretation is that this is rather a report of a question that one of the characters is asking. It could be the police officer reprimanding John by asking, somewhat sarcastically, "What were you thinking?". Or perhaps it's John reflecting on his own actions, thinking to himself "What was I thinking?". In this paper I propose to account for these kinds of report readings at the level of discourse structure. My proposal will be couched in the general framework of Segmented Discourse Representation Theory (SDRT, Asher & Lascarides 2003). Crucially, my analysis revolves around a dedicated discourse relation called ATTRIBUTION. I will provide a very general semantics for ATTRIBUTION in terms of an underspecified notion of characterization that covers
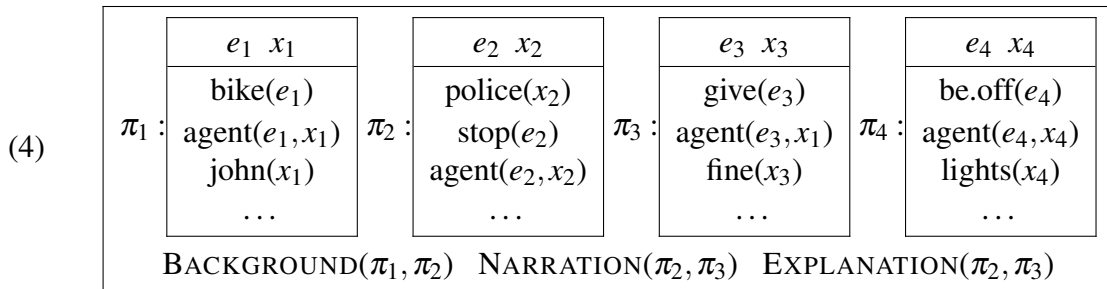
the full range of reporting types, from verbatim direct quotation to the paraphrasing of propositional content in attitude ascriptions.

## 2   Modeling coherence in SDRT

SDRT treats each individual clause in a discourse as contributing a separate discourse unit, and formulates a number of axioms that model the establishment of discourse relations, like NARRATION, RESULT, CONTRAST, and ELABORATION, between these discourse units. Unlike competing theories of discourse structure it gives these relations a model-theoretic semantic interpretation. For instance, the story in (1) gives rise to four elementary discourse units, typically labeled π_1, π_2, etc.

(3) π_1 : John was biking home late.
      π_2 : A police officer stopped him.
      π_3 : She gave him a fine.
      π_4 : His lights were off.

SDRT is compatible with any dynamic semantic interpretation for the individual discourse units, but in this paper I'll use DRT (Kamp & Reyle 1993), and extend its box-style notation to SDRSs as a whole, as illustrated in (4):[1]

(4)

e_1 x_1		e_2 x_2		e_3 x_3		e_4 x_4
bike(e_1)		police(x_2)		give(e_3)		be.off(e_4)
π_1 : agent(e_1,x_1)	π_2 :	stop(e_2)	π_3 :	agent(e_3,x_1)	π_4 :	agent(e_4,x_4)
john(x_1)		agent(e_2,x_2)		fine(x_3)		lights(x_4)
. . .		. . .		. . .		. . .

BACKGROUND(π_1,π_2)   NARRATION(π_2,π_3)   EXPLANATION(π_2,π_3)

Abstracting away from the semantic contents of the elementary units we can visualize just the global coherence structure of the discourse as a graph:

(5)

$$\pi_1 \xrightarrow{\text{BACKGROUND}} \pi_2 \xrightarrow{\text{NARRATION}} \pi_3$$
$$\Big\downarrow \text{EXPLANATION}$$
$$\pi_4$$

---
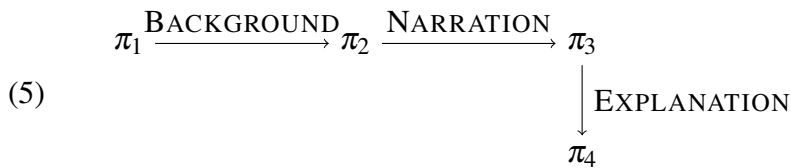
[1] To avoid formal clutter in notation I leave π_i discourse referents out of the DRS universes and ignore the top-level π_0 altogether. In the examples I discuss these can always easily be reconstructed unambiguously.

In these diagrams we stick with the standard SDRT convention of horizontal edges visualizing *coordinating* discourse relations, i.e., discourse relations like NARRATION and BACKGROUND that in some intuitive sense move the story forward and change the active topic, and vertical edges visualizing *subordinating* relations, i.e., relations like EXPLANATION or ELABORATION that don't move time and instead explore subtopics of the 'dominant' node.[2]

The two main questions for a formally precise and practically usable discourse semantics are: how do we derive a graph representation like (5) or (4) from a discourse like (1), and how exactly are we to interpret such formal structures? The SDRT framework provides two formal systems to answer these two questions. To start with the latter, the model-theoretic interpretation of an SDRT graph representation extends the standard DRT semantics for the graph's π_i-labeled DRS nodes with interpretation rules for the various discourse relations like in (6). Notation: K_{π_1} denotes the DRS unit that is labeled with proposition label π_1; e_{π_1} denotes the main eventuality introduced in the universe of the DRS unit labeled π_1; $[\![K]\!]$ denotes the dynamic semantic interpretation of a DRS (i.e., a context change potential, defined as a function from information states to information states, representing how an utterance affects an input context, à la Groenendijk & Stokhof 1991); $\circ$ denotes function composition (i.e., the dynamic semantic analogue of conjunction); the symbol $\bigcirc$ in a DRS condition denotes temporal overlap between eventualities; $\prec$ denotes immediate temporal precedence (the second eventuality occurs right after the first):

(6)   a. $[\![\text{NARRATION}(\pi_1, \pi_2)]\!] = [\![K_{\pi_1}]\!] \circ [\![K_{\pi_2}]\!] \circ [\![e_{\pi_1} \prec e_{\pi_2}]\!]$
      b. $[\![\text{EXPLANATION}(\pi_1, \pi_2)]\!] = [\![K_{\pi_1}]\!] \circ [\![K_{\pi_2}]\!] \circ [\![\text{cause}(e_{\pi_2}, e_{\pi_1})]\!]$
      c. $[\![\text{BACKGROUND}(\pi_1, \pi_2)]\!] = [\![K_{\pi_1}]\!] \circ [\![K_{\pi_2}]\!] \circ [\![e_{\pi_1} \bigcirc e_{\pi_2}]\!]$

In words, (6a) says that a NARRATION relation between two discourse units means that we have to update the context with the contents of both discourse units, in order, and moreover the main eventuality described by the second unit, follows immediately after the event described by that of the first.

Now for the first question, how to derive a discourse structure representation like the graph (5) and ultimately the full SDRS (4) from a sequence of utterances? Let's assume that the elementary discourse units are already identified and assigned DRS representations by the standard DRS construction algorithm (see Kamp & Reyle 1993). Now, SDRT's so-called Glue Logic provides inference rules that specify what discourse configurations trigger what discourse relations. For instance, a sequence

---

2 The main advantage of this convention is to visualize the so-called Right Frontier Constraint that relates anaphora resolution to discourse structure. In this paper we are not concerned with anaphora resolution so we'll skip over this (Asher & Lascarides 2003).

of two discourse units where the first contributes a state and the second an event licenses the inference that they are connected by a BACKGROUND relation – unless the resulting graph leads to an inconsistent or not maximally coherent final output representation. Similarly, a sequence of two eventive units defeasibly triggers ($\rightsquigarrow$) a NARRATION connection.

(7) a. $\text{state}(e_{\pi_1}) \wedge \text{event}(e_{\pi_2}) \rightsquigarrow \text{BACKGROUND}(\pi_1, \pi_2)$
     b. $\text{event}(e_{\pi_1}) \wedge \text{event}(e_{\pi_2}) \rightsquigarrow \text{NARRATION}(\pi_1, \pi_2)$

We will not go into the formal details of either model theory or Glue Logic, nor into the presupposed DRS construction algorithm and dynamic semantics in terms of context change potentials. I trust the above examples, diagrams and simplified formulas suffice to illustrate the basics of the SDRT discourse semantics framework to the uninitiated, and I defer to Asher & Lascarides (2003) for all formal details. In the following I provide an account of reported speech in this general framework, treating reporting as a discourse phenomenon, i.e., analyzing its semantic effects in terms of a semantically interpreted discourse relation ATTRIBUTION (Hunter 2016, Cumming 2020).

3 Indirect discourse

3.1 From operators to event modifiers

Attitude and speech reports have occupied a central position in semantic theory from its very beginnings (Frege 1892). In contemporary possible worlds semantics, the intensional operator approach (Hintikka 1969) and its descendants (Kaplan 1989, Schlenker 2003) are still dominant. Recently, there's been a rise in event-based versions, where the attitude or speech verb introduces an event of thinking, speaking, hoping, and the complement clause specifies the content of that event (Kratzer 2006, Hacquard 2010) (notation: $^{\wedge}\varphi$ refers to the possible worlds proposition expressed by φ, which is just a traditional Montagovian way of dealing with intensionality without introducing explicit possible worlds variables into the formal metalanguage).

(8) a. Mia said Don is a phony.
     b. $\exists e[\text{say}(e) \wedge \text{agent}(e, \text{mia}) \wedge \text{content}(e, {^{\wedge}}\text{phony}(\text{don}))]$
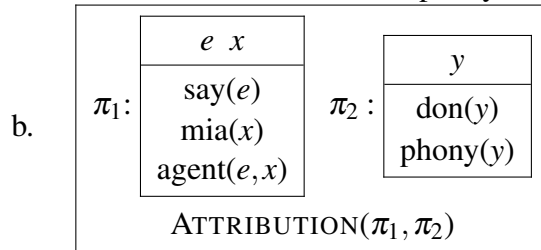
Such an analysis fits neatly in a more general neo-Davidsonian framework by treating subject and complement uniformly as event modifiers. Instead of treating speech and attitude verbs as special operators it relies on the idea that there are certain events that have propositional contents. In this section I'll adopt the event-based approach but move it from the syntax–semantics interface into the discourse/pragmatics level, where, I will argue in the remainder of the paper, it belongs.

3.2 From clausal complements to discourse units

When we look at a report like (8a) from the perspective of discourse structure, the first question that arises is whether we are dealing with a single elementary discourse unit (*Mia said Don is a phony*) or with two separate units (*Mia said (something)*, em Don is a phony) connected by a discourse relation. Hunter 2016 argues for the latter, on the basis of an ambiguity between regular (in)direct speech attributions and so-called parenthetical readings (also known as evidential or non-at-issue readings) of report constructions. In this paper I will provide a different, independent argument for this bipartite segmentation of indirect discourse, based on unembedded continuations of reports (§4). In the remainder of this subsection I first illustrate how a Hunter-style bipartite analysis could work for a simple report like (8a).

On Hunter's analysis, the two units in a report are connected by a discourse relation of ATTRIBUTION:[3]

(9) a. π_1: Mia said. π_2: Don is a phony.

b.

$$
\pi_1:
\begin{array}{|c|}
\hline
e \quad x \\
\hline
\text{say}(e) \\
\text{mia}(x) \\
\text{agent}(e,x) \\
\hline
\end{array}
\qquad
\pi_2:
\begin{array}{|c|}
\hline
y \\
\hline
\text{don}(y) \\
\text{phony}(y) \\
\hline
\end{array}
$$

$$\text{ATTRIBUTION}(\pi_1, \pi_2)$$

ATTRIBUTION is a non-veridical discourse relation, i.e., its truth does not presuppose the truth of both arguments. Specifically, π_2 serves to characterize what Mia said, not what the world is actually like. We build this into our semantics as follows, using the $\text{content}(e, p)$ relation from §3.1:[4]

(10) $[\![\text{ATTRIBUTION}(\pi_1, \pi_2)]\!] = [\![K_{\pi_1}]\!] \circ [\![\text{content}(e_{\pi_1}, {}^{\wedge}K_{\pi_2})]\!]$ (to be revised)

This definition presupposes that π_1 introduces a main eventuality (e_{π_1}) that can plausibly be said to have a propositional content, such as an utterance event, an occurrent thought, an attitudinal state, or a perceptual state/event. This requirement should ultimately be included in the antecedent of a defeasible Glue Logic axiom

---

3 A technical advantage of the Kratzerian event-based approach here over the classic Hintikkan intensional operator approach that Hunter uses is that a unit of the form 'Mia said' in (9a), without a grammatical object, is semantically speaking completely well-formed and interpretable.

4 The term 'attribution' is somewhat ambiguous: we usually say that we attribute an attitude or opinion to an individual, but strictly speaking the discourse relation of ATTRIBUTION here connects the content of the attitude/opinion to the event or state of an individual experiencing or expressing said attitude or opinion. Since there seems to be little risk of confusion, I've decided to stick with the now established SDRT terminology (e.g. Asher et al. 2006, Hunter 2016, Abrusán 2020a).

for inferring an ATTRIBUTION connection, of the form in (11), but we'll leave the precise conditions in the '...' for another occasion:

(11)   contentful.eventuality$(e_{\pi_1}) \wedge \ldots \rightsquigarrow$ ATTRIBUTION(π_1, π_2)

A defeasible inference rule of the form in (11) should allow us to infer ATTRI-BUTION in passages where we have one clause introducing a speech, thought, or attitude event, and another that could plausibly be interpreted as specifying that event's content. In the case of (8a) however we have a grammatical report construction that, arguably, forces an ATTRIBUTION connection between frame and complement. This situation is parallel to what we see with most other discourse relations. A CONTRAST may be left implicit, defeasibly inferred by the interpreter on the basis of various semantic, pragmatic, and discourse structural cues, but it may also be encoded directly in the grammar by means of an unambiguous, dedicated lexical item like *but*. Similarly, we have NARRATION, optionally marked by *and then*, or EXPLANATION by *because*. In SDRT, lexical items like these directly inform the Glue Logic, i.e., they simplify the SDRS construction process by filling in a fixed discourse relation. With ATTRIBUTION, we could assign this function to the complementizer *that* (which may be silent).[5]

In any case, whether marked on the surface as a report or inferred pragmatically on the basis of (11), π_1 and π_2 are going to be connected by ATTRIBUTION here. π_1 introduces a speech event, and (10) then tells us that the content of that event must be the proposition expressed by π_2. In other words, our bipartite discourse structure analysis gives us exactly the truth conditions that we also got from the compositional semantics in (8b).

3.3 Parenthetical reports

Before I introduce my own applications of the bipartite discourse analysis of reporting, let's briefly review Hunter's (2016) application to what she calls parenthetical indirect reports:

(12)   A: Why is Mia not in class?
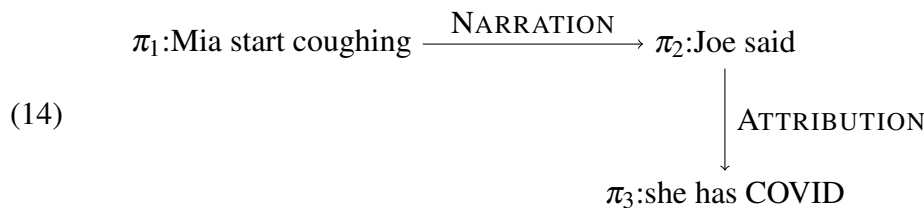   B: Joe said she has COVID.

A traditional Hintikka or Kratzer semantic analysis of B's answer gives us the proposition that Joe produced a speech act with a certain content, which hardly counts as an answer to A's question about Mia. But intuitively B does provide an

---

5 Alternatively, we can point out some other part of to the grammatical structure of a communication or attitude verb plus subordinated complement clause to encode the Glue Logic restriction to ATTRIBUTION.
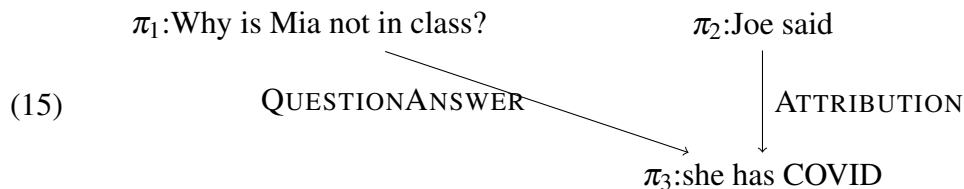
acceptable answer. According to Hunter, this is because B's indirect discourse report here allows a so-called parenthetical reading, presenting the reported information (that Mia has COVID) as the primary, at-issue contribution, with the reportative information (that Joe said something) serving as a not-at-issue meaning supplement. The current discourse approach to reports, where we analyze a report as consisting of two distinct discourse units, seems ideally suited to account for such parenthetical readings without having to assume a syntactic ambiguity. To prove this, let's first consider the exact same report in a different, more narrative context like (13).

(13)   So we're sitting in that waiting room, when Mia starts coughing, and then Joe said she has COVID. All hell broke loose.

In (13), as in the cases we'll be discussing in the next sections, it's the reporting segment, that Joe said something, that is at-issue, or as Hunter operationalizes it in SDRT, it's the reporting segment that directly connects to the previous discourse (via NARRATION, in this case): first Mia coughs, and then Joe says something (and as a result all hell breaks loose).

(14)

$$\pi_1\text{:Mia start coughing} \xrightarrow{\text{NARRATION}} \pi_2\text{:Joe said}$$
$$\downarrow \text{ATTRIBUTION}$$
$$\pi_3\text{:she has COVID}$$

Now back to the trickier case of the dialogue in (12). Here, it's the information contributed by the reported clause (that Mia has COVID), that directly connects to the previous discourse (i.e., A's question about Mia's whereabouts), in this case via the QUESTIONANSWER relation. To a first approximation, the discourse seems to be structured like this:

(15)

$$\pi_1\text{:Why is Mia not in class?} \qquad \pi_2\text{:Joe said}$$
$$\text{QUESTIONANSWER} \searrow \qquad \downarrow \text{ATTRIBUTION}$$
$$\pi_3\text{:she has COVID}$$

One complication that arises is that B probably uses the report embedding as way to hedge their own commitment to the truth of the complement, while the use of the veridical relation QUESTIONANSWER in (15) entails full commitment on A's part. Hunter's solution involves the introduction of 'modalized discourse relations', such as $\Diamond$QUESTIONANSWER, to decorate diagonal connections between material above

and below an ATTRIBUTIONarrow (as we see in (15)). In the following we'll ignore this and other complications (e.g., relating to Hunter's syntactically parenthetical reports or Bary & Maier's (2020) evidential reports) as we focus on what Bary & Maier call 'at-issue eventive' (uses of) reports, exemplified by (13), where it's the report frame that's directly connected to the previous discourse.

The upshot of this section is that we can emulate the classic semantic analysis of indirect discourse in a discourse framework, effectively moving the intensional embedding semantics from the syntax/semantics interface to the level of discourse structure. We've seen one distinctive benefit of the discourse approach, due to Hunter, viz. that it allows us to capture discourse parenthetical readings of reports without postulating ad hoc syntactic ambiguities.

4 Reports beyond the clause

On the discourse approach to reporting, the attitude or speech verb plus clausal complement construction is treated as a cue that informs the pragmasemantic Glue Logic of SDRS construction to infer the discourse relation of ATTRIBUTION between two discourse units. In the remainder of the paper I show that the powerful added machinery of the discourse-level approach is warranted by cases that are not overtly marked as reports but nonetheless interpreted as such. The most salient example of this is probably free indirect discourse, to be discussed in section 6. Below we first discuss another case that has received far less attention: indirect report continuations beyond the overtly embedded complement clause.

Consider the following extended dream report:

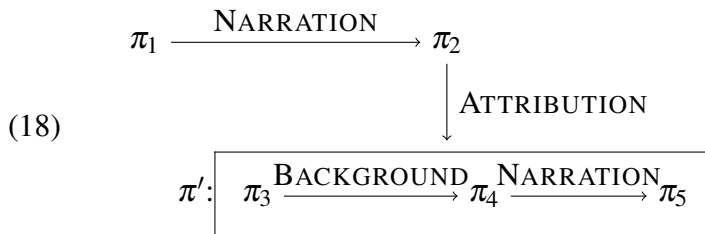(16)   Dan went to bed early. He dreamed that he was a frog. He jumped around a bit and then he was eaten by a stork.

On our discourse-level approach we parse this discourse as consisting of 5 segments.

(17) π_1 : Dan went to bed early
        π_2 : He dreamed
        π_3 : (that) he was a frog
        π_4 : He jumped around a bit
        π_5 : (and then) he was eaten by a stork

The corresponding discourse units can be straightforwardly connected by discourse relations to create an interpretable discourse graph. Note that in this example, two discourse relations are arguably encoded grammatically: the complement construction in *dreamed that* encodes ATTRIBUTION and *and then* encodes NARRATION. The rest can be defeasibly inferred by existing Glue Logic axioms, such as the

sequence of events in π_1 and π_2 giving rise to a likely NARRATION inference, and the sequence of state and event in π_3 and π_4 giving rise to a likely BACKGROUND inference.

Crucially, when we read this story, we take π_3-π_5 together to form a complex description of a single dream (that consists of an internally coherent sequence of events). To correctly represent this reading we need to construct a so-called complex discourse unit (Asher & Lascarides 2003). We then take that complex unit (rather than just π_3) as the second argument of the ATTRIBUTION. In our graph notation, we draw a labeled box to indicate the scope of a complex unit (here: π'): [6]

(18)

$$\pi_1 \xrightarrow{\text{NARRATION}} \pi_2$$

$$\downarrow \text{ATTRIBUTION}$$

$$\pi': \quad \boxed{\pi_3 \xrightarrow{\text{BACKGROUND}} \pi_4 \xrightarrow{\text{NARRATION}} \pi_5}$$

When we try to describe the discourse interpretation process leading up to this graph procedurally, the question arises. at what point do we create the complex unit? This is a thorny and quite general question for SDRT, but to keep it simple I propose to stipulate that the second argument of ATTRIBUTION[7] *always* comes with a complex unit. Thus, officially, the simple report *He dreamed that he was a frog* in isolation would already lead to the creation of a complex unit π' with π_3 as it's sole contents. Following the standard SDRT procedural attachment rules, this embedded unit π_3 is available for subsequent discourse units to attach to (it's on the so-called Right Frontier, Asher & Lascarides 2003). In the case of an isolated, single report clause, the stipulated extra layer of embedding is semantically superfluous, so we might as well introduce a notational shorthand to the effect that these extra embedding boxes are not drawn (resulting in familiar graph diagrams like (9b)) until they contain more than one discourse unit and thereby become semantically relevant (as in (18)).

The eventual discourse graph in (18) straightforwardly captures the 'modal subordination' (Roberts 1989) reading, where π_4 and π_5 are interpreted as describing the content of the dream, despite being syntactically outside the scope of the attitude

---

6 Interestingly, when we continue the discourse in (16) with *He woke up screaming*, we should attach that via RESULT to π_2 at the top-level, because it's not part of the dream description. He went to bed, and then had a dream (with such and such content), and as a result of the dream he woke up screaming. As one referee points out, we could construct another complex unit around π_2 and π' and connect that to the waking up screaming. Either way, since ATTRIBUTION is non-veridical, we can't connect the screaming directly to the being eaten with a veridical discourse relation.

7 Probably we can generalize this to any non-veridical argument of a coherence relation, as that's where complex discourse units are crucial.

verb. By contrast, the only way for a traditional sentence-level report semantics to deal with this would be to assume a silent dream operator in front of every proposition interpreted as a dream description. Note that such a sequence of hidden operators would ultimately still fail to capture the obvious discourse structural, temporal, and anaphoric relations between these segments (e.g., π_4 and π_5 could not be connected to each other by NARRATIONif they were each syntactically, semantically, or even discourse structurally, embedded by their own separate intensional operator.

Similar unmarked continuations of reports occur with other attitude and speech reports. In some languages, such syntactically unembedded continuations of speech reports can be marked with a reportative subjunctive mood on the verb:

(19)   Sie sagte sie habe   keine Zeit. Sie müsse   noch 86 Prüfungen
         She said she have-SUBJ *no   time. She must*-SUBJ *still   86 exams*
         bewerten.
         grade
         'She said she has no time. She still has 86 exams to grade (she said)'
         (German, Bary & Maier 2020)

In such constructions, the traditional, compositional approach would take the subjunctive morpheme as a separate semantic report operator (which causes significant complications for dealing with the overtly embedded subjunctive in the first sentence in (19), see Fabricius-Hansen & Saebø 2004). On the current approach, we take the subjunctive merely as a grammatical cue that constrains the Glue Logic to block attachment of the current unit to a top-level unit, i.e., forcing it to attach to a unit under an ATTRIBUTION.

In English, where we have no subjunctive inflection to mark something as reported content, we occasionally find unmarked free standing clauses that are interpreted as speech report continuations:

(20)   Trump says he'll cut inflation in half. He'll also create record numbers of jobs and beat COVID before Christmas.

As in (18), by connecting the propositions about inflation, record job numbers, and COVID together into a complex unit (using coordinating, veridical relations like LIST or CONTINUATION between them), we automatically get the most likely reading where all three together are semantically interpreted as describing what Trump said, without relying on any covert operators in the syntax.

5 Quotation

The above event-based implementation of Hunter's (2016) discourse-structural approach to indirect discourse applies to both speech and attitude reports in the indirect mode, i.e., where we are reporting the content of another person's speech or attitudinal state in our own words. I propose to generalize the semantics of ATTRIBUTION in order to cover also quotation and free indirect discourse reports, which seem to exhibit similar sensitivity to discourse structure, like allowing complex report continuations far beyond the sentence level.

5.1 Direct discourse and pure quotation

We start with a simple, clausal, direct quotation. On an event-based account we can treat direct and indirect speech uniformly as event modifiers, one that characterizes a speech event by its propositional content, and one that characterizes it by its linguistic form (Maier 2017):

(21)   a.   Mia said, "Don is a phony"
           b. $\exists e[\text{say}(e) \wedge \text{agent}(e, \text{mia}) \wedge \text{form}(e, \text{'Don is a phony'})]$

As with indirect reports I now propose a discourse-level alternative to this type of (near-)compositional account that retains the idea of treating quotation as event modification. We parse the quotation and the frame as distinct discourse units, connected by ATTRIBUTION.

(22)
$$\pi_1\text{:Mia said}$$
$$\Big\downarrow \text{ATTRIBUTION}$$
$$\pi_2\text{:"Don is a phony"}$$

Now, to get the right truth conditions we could technically admit two distinct, primitive types of attribution: one defined as in (10), contributing $[\![\text{content}(e_{\pi_1}, {}^{\wedge}K_{\pi_2})]\!]$, and one, say QATTRIBUTION, contributing instead something like $[\![\text{form}(e_{\pi_1}, \sigma_{\pi_2})]\!]$ (with σ_{π_2} denoting the linguistic/graphemic/phonological surface form of speech act π_2). However, this move will lead us down a path of multiplying discourse relations for each type of reporting, including, beyond direct and indirect discourse, mixed quotation, free indirect discourse, speech balloons, etc. In this paper I explore an alternative route, where we stick with a single discourse relation of ATTRIBUTION. To make this work we have to generalize its semantic contribution so that it subsumes both form- and content-based reporting.

5.2 Attribution as underspecified event characterization

I propose to replace our original definition of the semantics of content-based ATTRI-BUTION in (10) with (23), which invokes a distinct notion of 'characterization'. In this definition the notation 'Char$(\mathscr{F}(\pi),e)$' means that 'discourse unit π characterizes event e', using Asher & Lascarides's (2003) official SDRT notation, in which $\mathscr{F}$ denotes the function that maps labels in an SDRS to the SDRS constituents that they label – that is, $\mathscr{F}(\pi)$ is effectively a notational variant for what we've been denoting as K_π (we'll rely on this more abstract notation below in making precise what characterization does).

(23) $[\![\text{ATTRIBUTION}(\pi_1,\pi_2)]\!]=[\![K_{\pi_1}]\!]\circ[\![\text{Char}(\mathscr{F}(\pi_2),e_{\pi_1})]\!]$

The idea behind (23) is that languages may allow different ways of characterizing what someone said, thought, or dreamed. We can characterize what someone said by reproducing its propositional content in our own words. That is what happens in indirect discourse reports, and it is exactly this type of 'loose' characterizing that is formalized explicitly in our original formulation of the semantics of ATTRIBUTION in (10). But we can also characterize what someone said by reproducing the exact words uttered. This is what happens in direct discourse.[8]

   The proposed general approach to ATTRIBUTION leaves us with the question of what to do with the actual quotation marks. Are they merely a cue to enforce the inference of an underspecified ATTRIBUTION– the way we suggested treating the reportative subjunctive mood in (19) above –, or are they a genuine semantic quotation operator applied to the second ATTRIBUTION argument? Applied to the current SDRT setting, the first option – in the spirit of pragmatic accounts of quotation like Gutzmann & Stei 2011 – would mean that at the level of semantic representation quoted sentences are treated just like any other discourse unit, i.e., parsed and assigned their regular DRS representation. But for reports with quotation marks we need more than just the semantic representation of the complement, we need access to the actual form of the words used to express it. I propose that's what quotation marks do: they tell the DRS construction algorithm to introduce a surface form into the semantic representation. For reasons to be discussed below we'll assume that we also construct the regular DRS representation of the quoted material, where possible. Hence, in the full SDRS representation of (22), the report frame π_1 is represented as just a content DRS, while the quoted unit π_2 is represented as a form–content pair, consisting of a copy of the quoted surface form along with a DRS representation of its content.

---

8 Below we'll briefly survey some other forms of characterization, such as simultaneous form and content characterization, diagonal characterization, and iconic characterization.

(24)

$$\pi_1 : \boxed{\begin{array}{c} \text{e} \quad \text{x} \\ \hline \text{mia(x)} \quad \text{say(e)} \\ \text{agent(e,x)} \end{array}} \quad \pi_2 : \left\langle \text{Don is a phony,} \boxed{\begin{array}{c} \text{y} \\ \hline \text{don(y)} \quad \text{phony(y)} \end{array}} \right\rangle$$

$$\textsc{Attribution}(\pi_1, \pi_2)$$

We can now be more precise about the two most salient types of characterization that figure in the semantic definition of ATTRIBUTION. First, propositional characterization: A DRS K propositionally characterizes a contentful eventuality e if the proposition expressed by K matches the propositional content of e.[9] Second, formal characterization: a form–content pair formally characterizes a speech or thought event e if the form component matches the linguistic form of the reported speech event.[10] We can rephrase this more formally as in (25), using the following notational conventions: $[\![\varphi]\!]_w^{f,c}$ is the (static)[11] semantic interpretation of an atomic DRS condition φ, i.e., its truth value relative to an assignment f, a Kaplanian context c and a possible world index w; $[\![\varphi]\!]^{f,c} = \lambda w[[\![\varphi]\!]_w^{f,c}]$, i.e., the proposition expressed by φ; and *Content* and *Form* are the by now familiar functions mapping certain events to their propositional contents and surface forms, respectively.

(25) a. $[\![\text{Char}(K,e)]\!]_w^{f,c}$ is defined iff $f(e)$ is a contentful eventuality (speech event, belief state, etc.). If defined, $[\![\text{Char}(K,e)]\!]_w^{f,c} = 1$ iff *Content*$(f(e))$ $= [\![K]\!]^{f,c}$

  b. $[\![\text{Char}(\langle \sigma, K \rangle, e)]\!]_w^{f,c}$ is defined iff $f(e)$ is a linguistic speech act or language-like occurrent thought. If defined, $[\![\text{Char}(\langle \sigma, K \rangle, e)]\!]_w^{f,c} = 1$ iff *Form*$(f(e)) = \sigma$ (to be revised)

The definition of characterization in (25) together with the general semantics of ATTRIBUTION from (23) allows us to model some standard forms of direct and indirect discourse adequately. It effectively recreates the truth-conditional predictions of a traditional account of direct discourse as pure quotation, and a traditional

---

9 I'm assuming here that propositional matching means identity between sets of possible worlds. This is an oversimplification. The original speech act may in fact have been quite different from the reported complement (e.g., I can report that Mary said that she's coming if she literally said something more specific, like "I'll be at the party between 9 and 10PM" (von Stechow & Zimmermann 2005, Abreu Zavaleta 2019).

10 Again, for simplicity I'll assume matching means identity between strings of letters or phonemes, though to model judgments regarding natural language quotation more realistically we have to make room for cleaning up false starts and filled pauses and allow literal translations, at the very least.

11 In DRT we typically use essentially static truth definitions for conditions as part of a definition of dynamic context change potentials for DRSs. See Kamp, van Genabith & Reyle (2003) for details.

account of indirect discourse as an intensional operator (or rather, as contentful event modifier) (Kaplan 1989, Brasoveanu & Farkas 2007, Maier 2017).
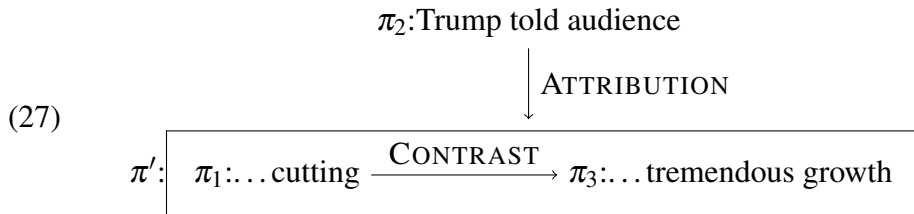
Note that (25b) effectively ignores the second component of the form–content pair in a quotation, which entails that quoted words are never really interpreted at all, they just contribute their form, i.e., their 'shape' (D. Davidson 1979), to the eventual interpretation. This would be fine if all we're interested in are the kinds of pure and direct quotations discussed in the philosophical literature, like *'Boston' is a six letter word* and *Otto said "I'm a fool"*. But when we're interested in more global discourse structures in actual text, this will prove unsatisfactory.[12]

5.3 Complex quotations

Take a, still very simple, quotation like (26).

(26)   "Oh, we'll be cutting," Trump told the audience. "But we're also going to have tremendous growth."

On the one hand, the two quoted fragments flanking the report frame should somehow be linked together, because together they characterize the form of Trump's speech. On the other hand, they clearly contribute two distinct discourse units of their own that are moreover meaningfully connected by CONTRAST (as evidenced by the overt connective *But*). In other words, we want a discourse graph like (27):

(27)

$$\pi_2:\text{Trump told audience}$$

$$\Big\downarrow \text{ATTRIBUTION}$$

$$\pi':\boxed{\pi_1:\dots\text{cutting} \xrightarrow{\text{CONTRAST}} \pi_3:\dots\text{tremendous growth}}$$

In order to correctly infer coherence (and anaphoric) connections between multi-sentence quotations and derive graph structures like (27), the Glue Logic needs to have some access to the semantic content of quoted discourse units as well as their forms. The first step we already took is to represent both form and content in the

---

12 Partee (1973) and others have already provided well-known arguments against the pure quotation approach to direct discourse on the basis of anaphora and ellipsis dependencies between quotation and surrounding discourse, as in:

(i)   "Don't worry, my boss likes me! He'll give me a raise" said Mary, but given the economic climate I doubt that he can. (Maier 2015)

SDRS representation of a quote, but now we still have to revise the form–content interpretation rule in (25b) to take advantage of that two-dimensional representation.

On a more technical note, when we spell out the full semantic (S)DRS box representation corresponding to the abstract discourse graph in (27), we get a complex unit, π', as the second component of our ATTRIBUTION, but as it stands this will be just a box around two form–content pairs, which is not itself a form–content pair yet,[13] and hence will not even trigger a quotational interpretation in the first place. To remedy this technical problem first, we add a 'Form-Projection' Rule: when we attach a form–content pair to another form–content pair inside a complex unit (which, we stipulated in §4, is always present under ATTRIBUTION), the embedded form components project up to the complex discourse unit containing them, where they are concatenated (notation: $^\cap$).

(28)   Form-Projection Rule:

$$\pi': \boxed{\pi_1 : \langle \sigma_1, K_1 \rangle \quad \pi_2 : \langle \sigma_2, K_2 \rangle} \quad \rightsquigarrow \quad \pi' : \left\langle \sigma_1 {}^\cap \sigma_2, \boxed{\pi_1 : K_1 \quad \pi_2 : K_2} \right\rangle$$

Applying form-projection to our example we get the following full SDRS representation for (27):

(29)

$$\pi' : \left\langle \begin{array}{l} \text{Oh, we'll be cutting.} \\ \text{But we're also going to} \\ \text{have tremendous growth} \end{array} , \begin{array}{c} \pi_2 : \boxed{\begin{array}{c} e_2 \\ \hline \text{tell}(e_2) \end{array}} \\ \\ \boxed{\pi_1 : \boxed{\begin{array}{c} e_1 \\ \hline \text{cut}(e_1) \end{array}} \quad \pi_3 : \boxed{\begin{array}{c} e_3 \\ \hline \text{have.growth}(e_3) \end{array}}} \\ \text{CONTRAST}(\pi_1, \pi_3) \end{array} \right\rangle$$

$$\text{ATTRIBUTION}(\pi_2, \pi')$$

In sum, a straightforward form-projection mechanism thus puts complex quotations like (26) in the right format to feed into our semantics, as laid out in (23) and (25).

Now to make our quotation semantics sensitive to both form and content (i.e., as philosophers put it, treat direct quotation as simultaneous mention and use (D. Davidson 1979, Cappelen & Lepore 1997)), I'll follow a straightforward implementation based on the two-dimensional account of direct quotation of Potts (2007): a form–content pair $\langle \sigma, K \rangle$ characterizes a speech or thought event e if the first component σ formally characterizes e and the second component K propositionally characterizes e.

---

13 The simple pure quotation analysis of §5.1 already can be seen as suffering from a milder version of this technical issue, if we had strictly followed our official stipulation that the second argument of ATTRIBUTION is always a complex discourse unit.

One complication we run into when we spell this out is that we have to incorporate a context shift in the content-matching criterion: propositional characterization in the case of direct discourse must compare the content of the speech/thought event e to the content of the complement K relative to the shifted, reported context of utterance, not relative to the actual, reporting context of utterance (as in regular indirect discourse) (Potts 2007). A context shift is necessary in order to get the reference of indexicals right – in direct discourse, all indexicals are systematically shifted. I'll assume a function *Context* mapping a speech/thought event to the context in which it takes place (Eckardt 2015).[14] In sum, we replace the second clause, (25b), in our general definition of characterization with a stricter definition that demands matching of form and content simultaneously, like this:[15]

(30) $[\![\text{Char}(\langle\sigma,K\rangle,e)]\!]_w^{f,c}$ is defined iff $f(e)$ is a linguistic speech act or language-like occurrent thought. If defined, $[\![\text{Char}(\langle\sigma,K\rangle,e)]\!]_w^{f,c}=1$ iff $Form(f(e))=\sigma$ and $Content(f(e))=[\![K]\!]^{f,Context(f(e))}$

As a further illustration of this rather technical, auxiliary notion of form–content characterization, let me show how it can be used to analyze mixed quotation.

5.4   Mixed quotation

Mixed quotation typically involves an indirect discourse where part of the clause is quoted directly (D. Davidson 1979):

(31)   Biden said that Putin "totally miscalculated"

Following the so-called presuppositional analysis of the phenomenon (Geurts & Maier 2005, Maier 2014), the intended interpretation can be schematically represented as involving two meaning components: an assertion of an (underspecified) indirect discourse, (32a), and a metalinguistic presupposition, (32b):[16]

(32)   a.   assertion: Biden said that Putin has property X.
         b.   presupposition: Biden used the words 'totally miscalculated' to express property X.

---

14 $Context(e)=\langle w,t,x\rangle$ iff e occurs in w at time t and the agent of e is x. This is assuming events are world-bound particulars. If we instead assume that a single event can occur in different possible worlds we would have to add the world as an extra parameter, i.e., $Context(e,w)$.

15 If we allow the content compartment to be empty, and in such cases disregard it semantically, we get a way to account for the intuitive well-formedness and interpretability of quoting gibberish (*She was like "Shis thewgg"*.
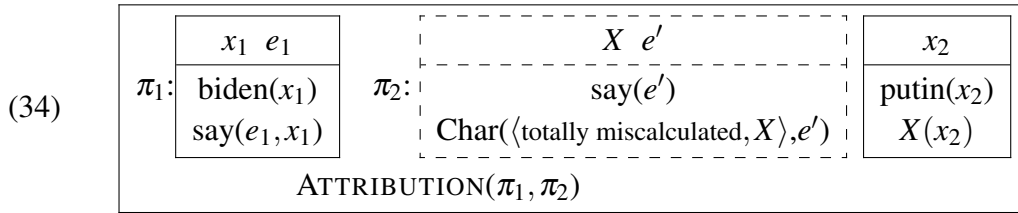
16 For arguments that the second component, in (32b), really is a presupposition and not some other type of (not-at-issue) content, I refer to (Maier 2014).

Maier's (2014) DRT implementation of this idea involves a primitive three-place relation E(xpress): $E(x,\text{totally miscalculated},X) \approx x$ uses the linguistic expression *totally miscalculated* to express semantic property X. When we port the ideas behind the presuppositional account of mixed quotation over to the current SDRT framework we can actually reduce this primitive E-relation to the independently motivated Char(acterization) relation introduced above. To make this precise, let's work out the interpretation of the simple example in (31).

   We start by segmenting the basic biclausal report as consisting of two units:

(33) π_1: Biden said π_2: Putin "totally miscalculated"

For the compositional semantic interpretation of π_2 let's follow Maier's syntactic parse and DRS construction, where the mixed quoted VP leads to the introduction of a discourse referent X (of type $\langle e,t \rangle$, i.e., ranging over properties), together with a metalinguistic presupposition, viz. that X is the property s.t. there was a saying event e' and X matches the content of e' and *totally miscalculated* matches the form of e'. We can capture this combination of form and content matching in a simple DRS condition using the Char relation (as defined in (30)). Further notational convention: unresolved presuppositions are represented by dashed boxes that sit between the triggering DRS box and its label.

(34)

$$
\pi_1: \begin{array}{|c|}
\hline
x_1 \;\; e_1 \\
\hline
\text{biden}(x_1) \\
\text{say}(e_1,x_1) \\
\hline
\end{array}
\quad
\pi_2:
\begin{array}{|c|}
\hline
X \;\; e' \\
\hline
\text{say}(e') \\
\text{Char}(\langle\text{totally miscalculated},X\rangle,e') \\
\hline
\end{array}
\begin{array}{|c|}
\hline
x_2 \\
\hline
\text{putin}(x_2) \\
X(x_2) \\
\hline
\end{array}
$$

$$\text{Attribution}(\pi_1,\pi_2)$$

We can resolve the metalinguistic presupposition in the (accessible) π_1 box, by accommodating both X and e' there. Note that while we can't directly bind e' to e_1 and fully equate them (because the content of e_1 is a full proposition, and that of e' is just a property), we can add a bridging inference to the effect that e' is a subevent of e_1 (notation: $e' \sqsubset e_1$).

(35)

$$
\pi_1: \begin{array}{|c|}
\hline
x_1 \;\; e_1 \;\; e' \\
\hline
\text{biden}(x_1) \\
\text{say}(e_1) \\
\text{say}(e') \quad e' \sqsubset e_1 \\
\text{Char}(\langle\text{totally miscalculated},X\rangle,e') \\
\hline
\end{array}
\quad
\pi_2: \begin{array}{|c|}
\hline
x_2 \\
\hline
\text{putin}(x_2) \\
X(x_2) \\
\hline
\end{array}
$$

$$\text{Attribution}(\pi_1,\pi_2)$$
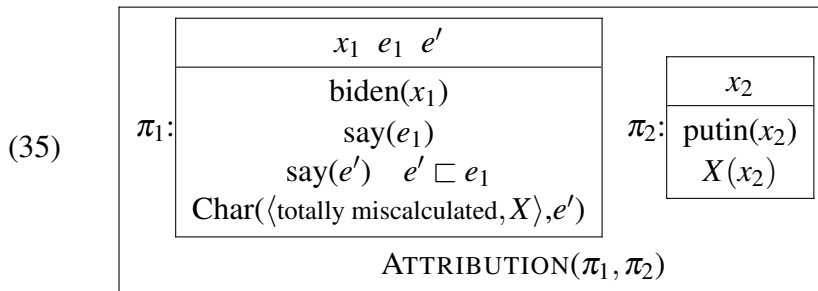
We've seen here how characterization plays an important role in capturing the metalinguistic presupposition triggered in mixed quotations. This is in line with the idea that characterization is a useful concept in its own right, beyond an auxiliary technicality that allows a more unified simple statement of the general meaning of ATTRIBUTION (viz. ATTRIBUTION(α,β) means that β characterizes the main eventuality of α). In fact, the definition of ATTRIBUTION in terms of Char opens up a variety of potential further extensions. Let me end this section with a few directions for future extensions of the framework, based on intuitively plausible extensions of the notion of characterizing.

First, we could define an intermediate mode of characterization, somewhere in between formal and propositional characterization, viz. characterization at the level of Kaplanian character or its diagonal (Kaplan 1989, Stalnaker 1978, Zimmermann 1991). This would be useful for capturing monstrous or *de se* reports.[17]

Second, we could extend characterization to the visual modality, to analyze distinctively visual conventions for representing characters' speech, thoughts, dreams, or hallucinations as instances of ATTRIBUTION, as part of an overall SDRT approach to analyzing the narrative structure of sequential visual media like comics and film (Bateman & Wildfeuer 2014, Cumming, Greenberg & Kelly 2017).[18]

Third, we might extend characterization beyond contentful events to model demonstrations more generally. For instance, the semantics of *Mary ate like <gobbling gesture>* (K. Davidson 2015) would involve an event of eating being 'iconically characterized' by a gobbling gesture.

Incorporating such extensions and comparing various implementations is beyond the scope of this paper, which focuses on the general account of reporting as a discourse-structural phenomenon.

---

17 To define this concisely assume that contexts ($c \in C$) and indices ($w \in W$) are tuples of the same type, i.e., indices are contexts with unused coordinates for agent, addressee, location etc., so that $C \subseteq W$ (von Stechow & Zimmermann 2005). Then we can easily define diagonal content as a a set of contexts:

(i)   diagonal DRS content: $\backslash K \backslash^f = \lambda c. [\![K]\!]_c^{f,c}$

We could now say that a discourse unit π with a DRS component K_π *diagonally characterizes* a contentful eventuality e if the '*de se* content' of e (the set of contexts 'compatible with e', Lewis 1979, Schlenker 2003) corresponds to the diagonal content of K_π.

18 Interestingly, the seemingly distinctive visual technique of the 'blended perspective shot' (e.g., presenting a character's internal perceptual hallucinations from a seemingly objective, neutral observer viewpoint) may already be captured by the plain content matching clause in (25a) that we've used for interpreting regular indirect discourse (Maier & Bimpikou 2019, Maier 2022).

6 Free indirect discourse

Free indirect discourse is a form of reported speech or thought that shows characteristics of both direct and indirect discourse (Banfield 1982). Take (36).

(36)   Sue stared at the calendar. Oh no, she had to hand in that damn paper today! She'd never make it...

The first sentence is just a description of what's going on in the story world, but the next two seem to describe what's going on inside Sue's head. The way this 'perspective shift' is marked linguistically is often subtle but it involves a combination of the use of expressive and indexical elements (*oh no*, *damn*, *today*, *!*) directly representing the protagonist Sue's point of view (i.e., as in direct speech), and the regular narrative past tense and third person pronouns (*she had to*, *she'd*) representing the thinking protagonist from the narrator's 'third person' perspective (i.e., as in indirect speech).[19]

Linguists have examined the semantic properties of free indirect discourse in some detail, and have proposed various competing semantic analyses, e.g., in terms of monstrous indirect discourse (Sharvit 2008), the addition of an extra context parameter (Schlenker 2004, Eckardt 2014), and quotation plus unquotation (Maier 2015, 2017).

Some salient features of free indirect discourse that are often overlooked by semanticists are (i) that these types of reports tend to span several sentences or even entire paragraphs, and (ii) that it may require intricate textual analysis to pinpoint exactly where such a report starts or ends. These neglected features however are exactly the type of thing we would expect on a discourse-structural approach. On our ATTRIBUTION-based approach, once we have established that there's an ATTRIBUTION, we get for each new incoming discourse unit a choice: do we attach it to the complex unit underneath that ATTRIBUTION(i.e., treat it as a continuation of the report), or to the main story line above it (i.e., treat it as a narrative description of the story world)? This choice is guided by often subtle considerations of global discourse coherence, i.e., which attachment generates a more coherent overall output SDRS (Asher & Lascarides 2003). Combined with the lack of clear, overt cues like quotation or (in English) subjunctive mood marking, this explains the observed difficulty of determining the exact boundaries of free indirect discourse passages.

Let me now flesh out the proposed discourse-structural ATTRIBUTION account of free indirect discourse by applying it to the example in (36). Attuned to the grammatical cues for free indirect discourse detection, sketched above, we can recognize three discourse units, of which two form a complex node that is connected

---

19 See Abrusán (2020a) for discussion of a more comprehensive algorithm for detecting 'perspective shift' based on grammatical, lexical and discourse-level cues.

to the previous discourse via ATTRIBUTION. But strictly speaking, ATTRIBUTION can't have the staring eventuality as its first argument, because staring is not in any way a contentful or linguistically structured event that can sensibly be characterized by a form or a content. Following recent discourse-structural analyses of free indirect discourse (Abrusán 2020b, Bimpikou, Maier & Hendriks 2021, Altshuler & Maier 2020) I propose that we may in such cases accommodate a simple discourse unit, π_3, to introduce the required thought event.

(37) π_1 : Sue stared at the calendar.
        π_2 : Oh no, she had to hand in that damn paper today!
        π_3 : (she thought.)
        π_4 :She'd never make it. . .

(38)

$$\pi_1 \xrightarrow{\ \text{BACKGROUND}\ } \pi_3$$

$$\downarrow \text{ATTRIBUTION}$$

$$\pi': \boxed{\pi_2 \xrightarrow{\ \text{RESULT}\ } \pi_4}$$
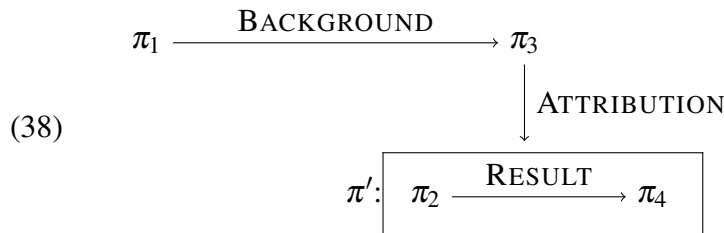
Due to the inherent underspecification in the semantics of ATTRIBUTION, this graph is in principle compatible with the various competing semantic analyses of the interpretation of free indirect discourse constructions. All that (38) tells us about the reports is that π_2 and π_4 together characterize the (accommodated) thought event in π_3. In its abstract graph form it doesn't specify what kind of characterization this is – simultaneous use/mention quotation, indirect discourse, or something else. But if we want to spell out the full SDRS box corresponding to the graph, and its interpretation, we'll eventually have to settle on a specific semantic theory. I'll explore here my own quotation-plus-unquotation approach.[20]

Let's assume, following the argumentation of Maier 2015, that the DRS construction algorithm treats a free indirect discourse segment – recognized as such – as essentially quoted. This means that we introduce corresponding form layers for π_2 and π_4. But, still following Maier 2015, pronouns and tenses are to be treated as 'unquoted'.[21] Technically, that means these pronouns and tenses are 'moved' out of the reports and interpreted separately, leaving (metalinguistic) traces (Maier 2014).

---

20 A monstrous account à la Sharvit 2008 would involve defining a mode of characterization that preserves the character or diagonal for most of the report, but preserves only content for pronouns and tenses, presumably relying on some feature deletion mechanism already at the syntax/semantic level of DRS construction.

21 Maier 2017 seeks to derive the unquote-pronouns-and-tenses assumption from general pragmatic interpretation and production principles.
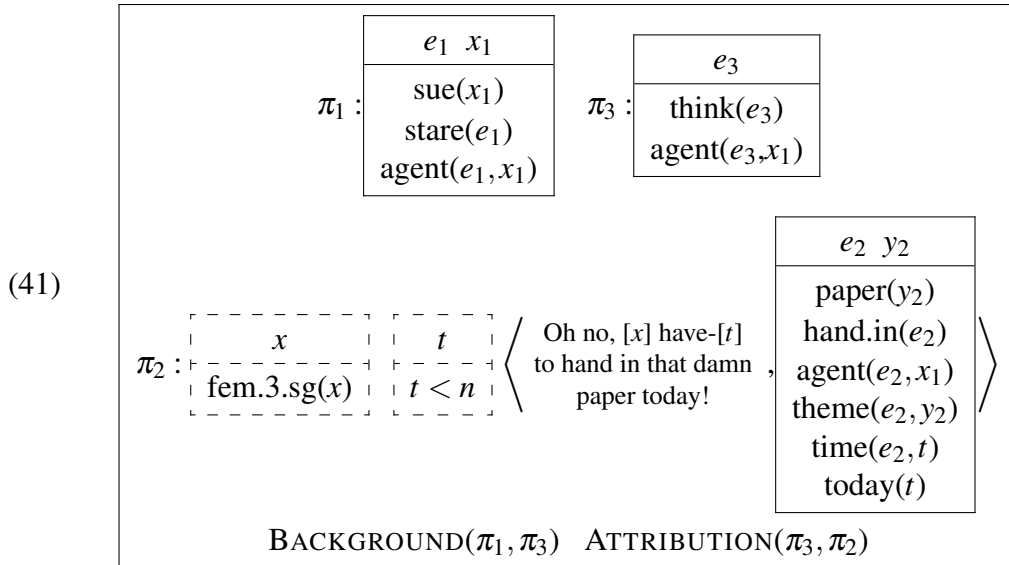
Let's go through the steps of the DRS construction algorithm for the first part of our example. First, we assume a (usually covert) quotation with (covert) unquotation of all pronouns and tenses, (39b). To interpret this semantically we first move the unquoted elements out of the quotation, (39c).

(39) a. Oh no, she had to hand in that damn paper today!
      b. "Oh no, [she] have-[past] to hand in that damn paper today!"
      c. she_x $past_t$ "Oh no, [x] have-[t] to hand in that damn paper today!"

Now we apply the standard DRS construction algorithm to the expressions in (39c). The two extraposed elements she_x and $past_t$ are anaphoric in nature and hence trigger presuppositions, the quotation will give rise to a labeled form–content pair consisting of the surface form (with two indexed holes) and a DRS box. The only new feature we have to add to the construction algorithm is a way to deal with indexed holes in a surface form. Since the traces tie each hole to a corresponding presupposition trigger, we can simply represent the contributions of the holes as the corresponding presupposed discourse referents, i.e., x and t, respectively.

(40) $\pi_2 :$ $\boxed{\begin{array}{c} x \\ \hline \text{fem.3.sg}(x) \end{array}}$ $\boxed{\begin{array}{c} t \\ \hline t < n \end{array}}$ $\left\langle \begin{array}{c} \text{Oh no, [x] have-[t]} \\ \text{to hand in that damn} \\ \text{paper today!} \end{array} , \boxed{\begin{array}{c} e_2 \ \ y_2 \\ \hline \text{paper}(y_2) \\ \text{hand.in}(e_2) \\ \text{agent}(e_2,x) \\ \text{theme}(e_2,y_2) \\ \text{time}(e_2,t) \\ \text{today}(t) \end{array}} \right\rangle$
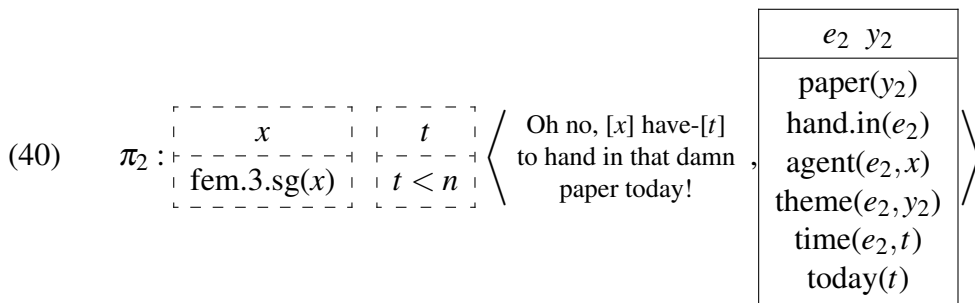
We can now add (40) to the SDRS under construction by connecting its discourse label to a suitable existing label (e.g., to a thought event, via ATTRIBUTION, or to another quoted or otherwise reported event already under an ATTRIBUTION). Looking at the earlier graph structure in (38), we have neither a suitable ATTRIBUTION nor a thought event, so we'll have to accommodate a thought event unit π_3 and attach (40) to that with an ATTRIBUTION to get the following SDRS:

(41)

$$\pi_1 : \boxed{\begin{array}{c} e_1 \ \ x_1 \\ \hline sue(x_1) \\ stare(e_1) \\ agent(e_1,x_1) \end{array}} \qquad \pi_3 : \boxed{\begin{array}{c} e_3 \\ \hline think(e_3) \\ agent(e_3,x_1) \end{array}}$$

$$\pi_2 : \boxed{\dashv x \vdash} \ \boxed{\dashv t \vdash} \ \left\langle \begin{array}{c} \text{Oh no, } [x] \text{ have-}[t] \\ \text{to hand in that damn} \\ \text{paper today!} \end{array} \right. , \boxed{\begin{array}{c} e_2 \ \ y_2 \\ \hline paper(y_2) \\ hand.in(e_2) \\ agent(e_2,x_1) \\ theme(e_2,y_2) \\ time(e_2,t) \\ today(t) \end{array}} \left.\right\rangle$$

$$\boxed{\dashv fem.3.sg(x) \vdash} \ \boxed{\dashv t < n \vdash}$$

$$\text{BACKGROUND}(\pi_1,\pi_3) \quad \text{ATTRIBUTION}(\pi_3,\pi_2)$$

Now we can resolve the presuppositions: x (*she*) binds to x_1, the only salient female third person, and t binds to the time of the thinking (e_3).[22] Now we add the final unit, π_4. We'll assume this is fed to the construction algorithm as a free indirect discourse, i.e., with quotation marks and unquotation holes, yielding a form–content pair with presuppositions, as in (40). We attach this π_4 to the existing form–content pair, π_2, within the existing (but previously invisible by the notational convention of Section 4) complex discourse unit π' under the existing ATTRIBUTION; project and concatenate the form components following (28); and bind π_4's unquoted tense and pronoun presuppositions. This gives the final output SDRS in (42), ascribing to Sue a complex thought whose form and content is characterized by two coherently connected discourse units.

---

22 More precisely, the antecedent time t_3 is introduced in K_{π_3} via a bridging inference. And this is still a simplification, as we have occurrences of both t_3 and x_1 in π' that are bound from outside the thought representation, leading to traditional philosophical worries about 'quantifying in', whose resolution is entirely orthogonal to the matters at hand.

(42)

$$\pi_1: \begin{array}{|c|}\hline e_1 \ \ x_1 \\\hline \text{sue}(x_1) \ \ \text{stare}(e_1) \\ \text{agent}(e_1,x_1) \\\hline\end{array} \quad \pi_3: \begin{array}{|c|}\hline e_3 \ \ t_3 \\\hline \text{think}(e_3) \\ \text{time}(e_3,t_3) \\ \text{agent}(e_3,x_1) \\\hline\end{array}$$

$$\pi': \left\langle \begin{array}{c}\text{Oh no, } [x_1] \text{ have-}[t_3] \\ \text{to hand in that damn} \\ \text{paper today!} \\ [x_1] \text{ will-}[t_3] \text{ never} \\ \text{make it in time}\end{array} \ , \ \pi_2: \begin{array}{|c|}\hline e_2 \ \ y_2 \\\hline \text{paper}(y) \\ \text{hand.in}(e_2) \\ \text{agent}(e_2,x_1) \\ \text{theme}(e_2,y_2) \\ \text{time}(e_2,t_3) \\ \text{today}(t_3) \\\hline\end{array} \ \pi_4: \begin{array}{|c|}\hline e_4 \\\hline \text{make.it}(e_4) \\ \text{agent}(e_4,x_1) \\ \text{time}(e_4,t_3) \\ \text{in.time}(t_3) \\\hline\end{array} \right\rangle$$

$$\text{RESULT}(\pi_2,\pi_4)$$

$$\text{BACKGROUND}(\pi_1,\pi_3) \quad \text{ATTRIBUTION}(\pi_3,\pi')$$
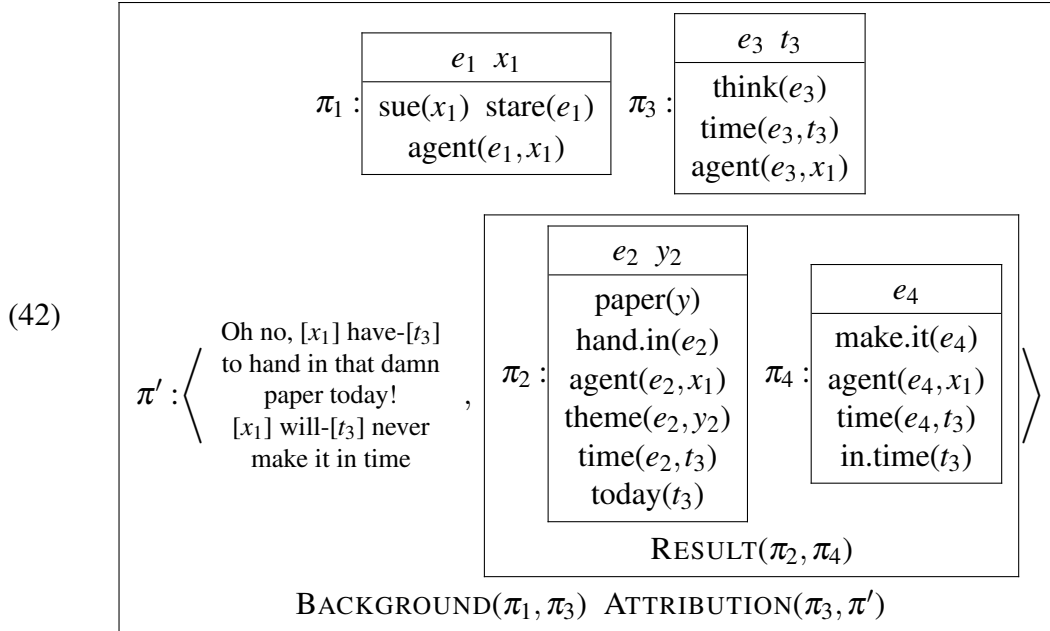
7 Conclusion

I have proposed abandoning attempts to model reporting constructions in terms of various clausal operators integrated in a compositional semantics. Instead, we should model them at the level of discourse structure. More specifically, I have proposed a discourse-structural account of all reporting in terms of a discourse relation of ATTRIBUTION connecting two distinct discourse units: one contributed by a frame segment (*she said*, *he dreamed*) and one complex report unit contributed by, for instance, a clausal complement (*that he was unhappy*), or a complex multi-sentence quotation (*"I'll beat COVID. But not global warming. That's still a hoax"*). I have proposed a simple semantics for the discourse relation of ATTRIBUTION that relies on the notion of a speech/thought/attitude eventuality being 'characterized' by a surface form or a propositional content, or both.

The proposed discourse-structural account is embedded in the general discourse semantics framework of SDRT. Clausal complements are simply analyzed as contributing their own discourse units, represented by a labeled DRS in the discourse-level 'logical form' (the SDRS). Quotation marks serve to introduce a surface form layer on top of the DRS representation of a quoted unit. These straightforward assumptions allow us to implement simultaneous use and mention for direct quotation, which I motivate with cases where multiple quoted sentences together form a complex discourse unit describing an internally coherent multi-sentence quotation.

More generally, it is such cases of extended direct, indirect, and free indirect reports, beyond the single reported clause, that have been the blind spots of traditional semantic accounts of attitude reports and quotation and that motivate the proposed shift from the syntax/semantics interface, to the level of discourse structure when it comes to understanding reports.

References

Abreu Zavaleta, Martín. 2019. Weak speech reports. *Philosophical Studies* 176(8). 2139–2166. https://doi.org/10.1007/s11098-018-1119-2.

Abrusán, Márta. 2020a. Computing perspective shift in narratives. In Emar Maier & Andreas Stokke (eds.), *The language of fiction*, forthcoming. Oxford: OUP.

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

Altshuler, Daniel & Emar Maier. 2020. Coping with imaginative resistance. Ms. https://philpapers.org/rec/ALTCWI.

Asher, Nicholas, Julie Hunter, Pascal Denis & Brian Reese. 2006. Evidentiality and intensionality: Two uses of reportative constructions in discourse. In *Workshop on Constraints in Discourse Structure*.

Asher, Nicholas & Alex Lascarides. 2003. *Logics of conversation*. Cambridge University Press. 556 pp.

Banfield, Ann. 1982. *Unspeakable sentences: narration and representation in the language of fiction*. London: Routledge & Kegan Paul.

Bary, Corien & Emar Maier. 2020. The landscape of speech reporting. Ms. Nijmegen/Groningen. https://ling.auf.net/lingbuzz/003526 (27 September, 2018).

Bateman, John & Janina Wildfeuer. 2014. A multimodal discourse theory of visual narrative. en. *Journal of Pragmatics* 74. 180–208. https://doi.org/10.1016/j.pragma.2014.10.001.

Bimpikou, Sofia, Emar Maier & Petra Hendriks. 2021. The discourse structure of free indirect discourse. In Mark Dingemanse & Elena Tribushinina (eds.), *LInguistics in the netherlands*. Amsterdam: John Benjamins.

Brasoveanu, Adrian & Donka Farkas. 2007. Say reports, assertion events and meaning dimensions. *Pitar Mos: A Building with a View. Papers in Honour of Alexandra Cornilescu*. 175–196.

Cappelen, Herman & Ernest Lepore. 1997. Varieties of quotation. *Mind* 106(423). 429–450. https://doi.org/10.1093/mind/106.423.429.

Cumming, Sam. 2020. Narrative and point-of-view. In Emar Maier & Andreas Stokke (eds.), *The language of fiction*, forthcoming. Oxford: OUP.

Cumming, Samuel, Gabriel Greenberg & Rory Kelly. 2017. Conventions of view-point coherence in film. *Philosophers' Imprint* 17(1). 1–29. http://hdl.handle.net/2027/spo.3521354.0017.001.

Davidson, Donald. 1979. Quotation. *Theory and Decision* 11(1). 27–40. https://doi.org/10.1007/BF00126690.

Davidson, Kathryn. 2015. Quotation, demonstration, and iconicity. *Linguistics and Philosophy* 38(6). 477–520. https://doi.org/10.1007/s10988-015-9180-1.

Eckardt, Regine. 2014. *The semantics of free indirect speech. how texts let you read minds and eavesdrop*. Leiden: Brill.

Eckardt, Regine. 2015. Utterance events and indirect speech. *Grazer Linguistische Studien* 83. 27–46.

Fabricius-Hansen, Cathrine & Kjell Johan Saebø. 2004. In a mediative mood: the semantics of the german reportive subjunctive. *Natural Language Semantics* 12(3). 213–257. http://link.springer.com/article/10.1023/B:NALS.0000034514.27887.d9 (27 June, 2015).

Frege, Gottlob. 1892. Uber Sinn und Bedeutung. *Zeitschrift für Philosophie und philosophische Kritik* 100(1). 25–50.

Geurts, Bart & Emar Maier. 2005. Quotation in context. *Belgian Journal of Linguistics* 17(1). 109–128. https://doi.org/10.1075/bjl.17.07geu.

Groenendijk, Jeroen & Martin Stokhof. 1991. Dynamic predicate logic. *Linguistics and Philosophy* 14(1). 39–100. https://doi.org/10.1007/BF00628304.

Gutzmann, Daniel & Erik Stei. 2011. How quotation marks what people do with words. *Journal of Pragmatics* 43(10). 2650–2663. https://doi.org/10.1016/j.pragma.2011.03.010.

Hacquard, Valentine. 2010. On the event relativity of modal auxiliaries. *Natural Language Semantics* 18(1). 79–114. https://doi.org/10.1007/s11050-010-9056-4.

Hintikka, Jaakko. 1969. Semantics for propositional attitudes. In *Models for modalities*, 87–112. Dordrecht: Reidel.

Hobbs, Jerry R. 1979. Coherence and coreference. *Cognitive Science* 3(1). 67–90. https://doi.org/10.1207/s15516709cog0301_4.

Hunter, Julie. 2016. Reports in discourse. *Dialogue & Discourse* 7(4). http://dad.uni-bielefeld.de/index.php/dad/article/view/3695 (13 April, 2017).

Kamp, Hans, Josef van Genabith & Uwe Reyle. 2003. Discourse representation theory. In Dov Gabbay & Franz Guenthner (eds.), *Handbook of philosophical logic*, vol. 10, 125–394. Heidelberg: Springer.

Kamp, Hans & Uwe Reyle. 1993. *From discourse to logic: an introduction to modeltheoretic semantics in natural language, formal logic and discourse representation theory*. Vol. 1. Dordrecht: Kluwer.

Kaplan, David. 1989. Demonstratives. In Joseph Almog, John Perry & Howard Wettstein (eds.), *Themes from kaplan*, 481–614. New York: Oxford University Press.

Kratzer, Angelika. 2006. Decomposing attitude verbs. Handout. Honoring Anita Mittwoch on her 80th birthday. The Hebrew University of Jerusalem.

Lewis, David. 1979. Attitudes de dicto and de se. *The Philosophical Review* 88(4). 513–543. http://www.jstor.org/stable/2184843.

Maier, Emar. 2014. Mixed quotation: the grammar of apparently transparent opacity. *Semantics and Pragmatics* 7(7). 1–67. https://doi.org/10.3765/sp.7.7.

Maier, Emar. 2015. Quotation and unquotation in free indirect discourse. *Mind and Language* 30(3). 235–273. https://doi.org/10.1111/mila.12083.

Maier, Emar. 2017. The pragmatics of attraction: explaining unquotation in direct and free indirect discourse. In Paul Saka & Michael Johnson (eds.), *The semantics and pragmatics of quotation*. Berlin: Springer. http://ling.auf.net/lingbuzz/002966 (3 May, 2016).

Maier, Emar. 2022. Unreliability and point of view in filmic narration. *Epistemology and Philosophy of Science* 59(2). 23–37. https://doi.org/10.5840/eps202259217.

Maier, Emar & Sofia Bimpikou. 2019. Shifting perspectives in pictorial narratives. *Sinn und Bedeutung* 23.

Partee, Barbara. 1973. The syntax and semantics of quotation. In S. Anderson & Paul Kiparsky (eds.), *A festschrift for Morris Halle*, 410–418. New York: Holt, Rinehart & Winston.

Potts, Christopher. 2007. The dimensions of quotation. In Chris Barker & Pauline Jacobson (eds.), *Direct compositionality*, 405–431. New York: Oxford University Press.

Roberts, Craige. 1989. Modal subordination and pronominal anaphora in discourse. *Linguistics and Philosophy* 12(6). 683–721. https://doi.org/10.1007/BF00632602.

Schlenker, Philippe. 2003. A plea for monsters. *Linguistics and Philosophy* 26(1). 29–120. https://doi.org/10.1023/A:1022225203544.

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. https://doi.org/10.1111/j.1468-0017.2004.00259.x.

Sharvit, Yael. 2008. The puzzle of free indirect discourse. *Linguistics and Philosophy* 31(3). 353–395. https://doi.org/10.1007/s10988-008-9039-9.

Stalnaker, Robert. 1978. Assertion. In Peter Cole (ed.), *Syntax and semantics 9: pragmatics*, 315–332. New York: Academic Press.

von Stechow, Arnim & Ede Zimmermann. 2005. A problem for a compositional treatment of de re attitudes. In Greg Carlson & Francis Pelletier (eds.), *Reference and quantification: the partee effect*, 207–228. Stanford: CSLI.

Zimmermann, Thomas Ede. 1991. Kontextabhängigkeit. In Arnim von Stechow &
Dieter Wunderlich (eds.), *Semantik: ein internationales handbuch der zeitgenös-
sischen forschung*, 156–229. Berlin/New York: Walter de Gruyter.