



HAL
open science

Rhetorical Structure and QUDs

Julie Hunter, Marta Abrusan

► **To cite this version:**

Julie Hunter, Marta Abrusan. Rhetorical Structure and QUDs. M. Otake; S. Kurahashi; Y. Ota; K. Aatoh; D. Bekki. New Frontiers in Artificial Intelligence JSAI-isAI 2015 Workshops, LENLS, JURISIN, AAA, HAT-MASH, TSDAA, ASD-HR, and SKL, Kanagawa, Japan, November 16-18, 2015, Revised Selected Papers, Springer, pp.41-57, 2017, 10.1007/978-3-319-50953-2_4 . hal-03906547

HAL Id: hal-03906547

<https://hal.science/hal-03906547>

Submitted on 19 Dec 2022

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

Rhetorical Structure and QUDs*

Julie Hunter¹ and Márta Abrusán²

¹ IRIT, Université Paul Sabatier

² IRIT, CNRS

Abstract. We consider two hypotheses about how rhetorical structure and QUD structure might come together to provide a more general pragmatic theory. Taking SDRT ([2]) and some basic principles from [18]’s QUD framework as starting points, we first consider the possibility that rhetorical relations can be modelled as QUDs, and vice versa. We ultimately reject this hypothesis in favor of the possibility that QUDs correspond to topics that bind together the members of complex discourse units.

Theories of rhetorical structure [2, 13] and theories of discourse structure centered around a Question Under Discussion or QUD [8, 18] share many of the same principles.³ Both approaches hold that the interpretation of a given sentence or *elementary discourse unit* (EDU) depends in part on that EDU’s relation to other moves that have been made in the same discourse.⁴ Both also hold that a discourse context must therefore keep track of not only (some subset of the) prior discourse moves but also certain structural relations between these moves. These structural relations are believed to play an integral role in discourse coherence and the relevance of individual discourse moves, and as such, to influence various semantic and pragmatic phenomena, including ellipsis of various sorts, anaphora (rhetorical theories), and prosody (QUD theories).

To clarify their potential contribution to analyses of semantic and pragmatic phenomena, we need a better understanding of how rhetorical and QUD frameworks are related: are they fundamentally distinct, but complementary theories, or do they aim to model the same phenomena? If the latter, do they end up describing two sides of the same coin or are they in conflict? The goal of this paper is to propose and evaluate two hypotheses about how the two frameworks might correspond, so that we can eventually come to a better understanding of the respective roles that these frameworks play with regard to phenomena from the semantics-pragmatics interface. The first hypothesis, which we develop and ultimately reject in §2.1, is that there is a direct correspondence between instances of discourse relations and QUDs in a discourse. The second hypothesis, which we present in §2.2, is that QUDs correspond to *complex discourse units* in a discourse graph. We judge this hypothesis to be more promising.

Our discussion focusses on two particular theories. On the side of rhetorical structure, we adopt *Segmented Discourse Representation Theory* [2], which we briefly introduce in §1.1. Of all of the rhetorical theories, SDRT is the most developed from a

* We gratefully acknowledge support from the ERC grant 269427 and Marie Curie FP7 Grant, PCIG13-GA-2013-618550.

³ See also [21] for a framework that combines QUDs and rhetorical relations.

⁴ How a discourse is broken down into basic units can vary from theory to theory, but all rhetorical theories and QUD theories must take a stand on what constitutes a basic discourse move.

semantic point of view; it assigns semantics to each of its discourse relations and posits semantic constraints on how its hierarchical discourse representations, which capture the contents of full discourses, can be constructed. These choices are fuelled by concerns about anaphora resolution, presupposition, temporal interpretation and other phenomena relevant to the semantics-pragmatics interface, putting SDRT in an ideal position to be compared to alternative semantic/pragmatic theories of discourse structure. On the side of QUDs, we take inspiration from [18] whose proposal, we think, is closest in spirit to SDRT. Like SDRT it aims to capture relations between discourse moves and provide hierarchical discourse structures that model global features of the discourse context. However, because the account outlined in [18] less thoroughly developed than SDRT, the possible formulations of QUD frameworks that we consider in §1.2 sometimes move beyond basic principles laid out in [18] in directions that we think deserve consideration, but which may not be endorsed by Roberts.

1 Background on SDRT and QUD

1.1 A very brief introduction to SDRT

A fundamental principle of rhetorical theories, including SDRT, is that the relations that utterances stand in to one another affect on the one hand the interpretation of the discourse in which the utterances figure and on the other, the interpretation of the utterance contents themselves. Consider (1):

- (1) I missed my meeting this morning. My car broke down.

Seeking a connection between the two parts of (1), an addressee will naturally understand the content of the second sentence as providing an explanation for the content of the first. This interpretation of the discourse comes with its own truth conditions: it is true just in case the speaker was late for her meeting, her car broke down, and the latter event was the cause of the former. At the same time, the inferred causal connection affects the interpretation of the two sentences in (1). Both sentences are in the past tense, which indicates that the events that they describe occurred before the speech time. The inferred causal relation between the two events entails that in addition, the second event must have occurred before the first—a cause must occur before its effect. This sequence of events does not follow from the tense of the verbs and the order in which the events are described in the discourse, but only from understanding how the two sentences are coherently or rhetorically related. Of course, in the absence of an explicit marker for causality (e.g. *because*), this causal connection is at best an implicature; nevertheless, speakers regularly accept and act on such implicatures. SDRT aims to model such connections by developing rules for constructing logical forms for discourses and by providing models to interpret these logical forms.

To capture the rhetorical connections in a discourse, rhetorical theories must accomplish three tasks. First, the discourse must be segmented into EDUs according to rules set by the theory. In SDRT, the aim is to have each EDU denote a single eventuality. The next task is to attach each EDU to some other part of the discourse. SDRT maintains that a new move can only be a coherent extension of a given discourse if it is relevant to some other move that has been made previously; thus, each EDU will be attached to

at least one other EDU that was discourse prior to it, with the exception of the initial EDU, which will only have attachments to later moves. In this way, a rhetorical relation is a kind of *anaphoric* relation, but one that holds between discourse units rather than referring expressions and discourse referents. The final task to be accomplished, which is in practice accomplished in tandem with labelling and even segmentation at times, is the labelling of each discourse attachment with a rhetorical relation (Explanation, Elaboration, Contrast, Narration, etc.).

Attachment and labelling involve default reasoning about the contents of the EDUs involved and world knowledge. Importantly, however, they also take into account *global* features from the discourse. For example, because discourse relations are often implicatures, the inference of an attachment/label that might otherwise be justified can be blocked by information carried by other EDUs in the discourse. For this reason, SDRT provides defeasible rules for inferring discourse relations and stresses the need for global constraints on the development of a discourse structure.

One such global constraint, which affects EDU attachment, is the Right Frontier Constraint (RFC). SDRT, along with other discourse theories [16, 15], posits that even if the content of a new EDU satisfies the semantic conditions for being attached to some other EDU in a discourse, this connection is only coherent if the prior EDU is *accessible*⁵, that is, if it is on the Right Frontier (RF). Compare (2-a)-(2-c):

- (2) a. [John speaks German.] π_1 [He can translate for you while you're in Berlin.] π_2
 b. [John speaks German] $\pi_{1'}$ [and his sister speaks French.] $\pi_{2'}$?? [He can translate for you while you're in Berlin.] $\pi_{3'}$
 c. [John speaks German.] $\pi_{1''}$ [(because) He lived in Stuttgart for 10 years.] $\pi_{2''}$ [He can translate for you while you're in Berlin.] $\pi_{3''}$

While π_1 and π_2 in (2-a) are related by Result—a connection supported by the contents of the EDUs and world knowledge about what being able to translate in Berlin would normally require—the same connection is blocked for $\pi_{1'}$ and $\pi_{3'}$ in (2-b). (2-c), however, shows that it is not blocked by just any intervening EDUs.

SDRT defines the RF so as to reflect facts about accessibility like those illustrated in (2-a)-(2-c). We start by introducing SDRT's discourse structures, which are connected graphs rather than trees. Graphs are needed to model certain facts about discourse, two of which are relevant for this paper: (i) some units can have incoming links from two different EDUs, and (ii) multiple DUS can work together to form a *complex discourse unit* (CDU), which serves as a single argument to a discourse relation.

- (3) a. Sam is being punished. π_1 She took her parents' car without permission, π_2 so they've grounded her for 2 weeks. π_3
 b. *Explanation* (π_1, π_2), *Result* (π_2, π_3), *Elaboration* (π_1, π_3)
 (4) a. $\pi_{1'} + \pi_{2'}$, but their parents don't speak any foreign languages. $\pi_{3'}$
 b. *Continuation*($\pi_{1'}, \pi_{2'}$), *Contrast* ($[\pi_{1'}, \pi_{2'}], \pi_{3'}$)

In (3), π_3 is the second argument for an instance of Result and an instance of Elaboration, whose first arguments are distinct. In (4), which builds on $\pi_{1'}$ and $\pi_{2'}$ from (2-b),

⁵ SDRT does allow for violations of the RFC in cases that [1] calls *discourse subordination*, but such violations need to be explicitly signalled, e.g. *Let's go back to your first point*.

not only are both $\pi_{1'}$ and $\pi_{2'}$ needed to provide the necessary antecedent for *their* in $\pi_{3'}$, but these units are attached to each other and both satisfy the conditions needed to support the Contrast with $\pi_{3'}$. SDRT would therefore group them into a single, though internally complex, argument for Contrast.

SDRT's graphs thus contain two types of edges: (i) edges that are labelled with discourse relations, and (ii) edges that relate each CDU to each DU that it contains. Edges of type (i) can be further subdivided into *subordinating* and *coordinating* edges, governed by the semantics of the relations that label these edges. When a DU is attached to another DU via a subordinating relation, both arguments remain accessible for further attachments. In (2-c), for example, the role of $\pi_{2''}$ is to provide background information that explains how John came to speak German ($\pi_{1''}$) or to simply back up that claim. $\pi_{1''}$ therefore remains central to the discussion and salient enough to be accessible to $\pi_{3''}$ despite the intervening $\pi_{2''}$. Subordinating relations include Explanation, Elaboration, Background, and Question-Answer Pair, among others. When a DU is attached to another DU via a coordinating relation, the RF is pushed dynamically forward so that the second DU is on the RF, but the first is knocked off. In (2-b), for instance, $\pi_{2'}$ goes on to tell us that John's sister speaks French, so the discourse is no longer centered on John's German speaking abilities when it comes time to add $\pi_{3'}$. An attachment from $\pi_{3'}$ to $\pi_{1'}$ is thus difficult to achieve. Coordinating relations include Contrast, Continuation, Narration (Sequence), Result, and Conditional, among others.

We conclude with a description of SDRT's RF. A node π_x is on the RF of a graph G , i.e. $\text{RF}_G(\pi_x)$, just in case (a) π_x is *Last*, i.e. π_x is the EDU introduced most recently into G following the textual order of the EDUs in G , or (b) $\exists \pi_y(\text{RF}_G(\pi_y))$ such that (i) $e_s(\pi_x, \pi_y)$ for a subordinating edge e_s or (ii) $\pi_y \in \pi_x$ (i.e., π_x is a CDU). The Right Frontier Constraint (RFC) simply states that an incoming EDU that needs to be attached to a graph G should be attached to a node along the RF.

1.2 QUD

The concept of a QUD is not a homogenous one; we will aim in this section to clarify two ways of understanding this concept from the literature. Perhaps the widest application of QUD has been in the context of focus phenomena ([5, 6, 18]). Consider (5) and (6) (where the capital letters indicate the desired intonational emphasis):

- (5) a. John gave the cake to TOM.
 b. John gave the CAKE to Tom.

(5-a) and (5-b) have the same syntactic structure and the same truth conditions: each is true just in case John gave the cake in question to Tom. Yet utterances of (5-a) and (5-b) are not appropriate in the same contexts. (5-a), for example, would be an appropriate answer to the question (Q1) *To whom did John give the cake?* but not to (Q2) *What did John give to Tom?*; (5-b) would be an appropriate answer to (Q2) but not to (Q1).

[19] proposed that focus is licensed if there is a suitable set of anaphorically available alternative propositions in the context. Since questions are standardly analysed as introducing a set of alternative questions (*Alt-Q*) (cf. [11]), the anaphoric requirement of focus can be satisfied by a *congruent question*, a question whose alternative set *Alt-Q* is a subset of *Alt-F*. The *Alt-F* sets for (5-a) and (5-b) are as follows:

- (6) a. $Alt-F$ of (5-a) = $\{x \in D_e: \text{John gave the cake to } x\}$
 b. $Alt-F$ of (5-b) = $\{x \in D_e: \text{John gave } x \text{ to Tom}\}$

(Q1) is a suitable antecedent for the alternatives introduced by (5-a) but not by (5-b) because $Alt-(Q1) = \{x \in D_e: \text{John gave the cake to } x\}$ is a subset of $Alt-F$ of (5-a) but not $Alt-F$ of (5-b). Conversely, (Q2) is a suitable antecedent for (5-b) but not (5-a).

Sometimes the QUD for a discourse can be a sub-question of a larger question that needs to be addressed, as illustrated in (7). Such hierarchical structures or *stacks* of questions have been proposed by [21, ?, ?, ?], among others.

- (7) John gave the CAKE to [Tom]_{CT} and the ICE cream to [Linda]_{CT}.

[6] would analyze (7) as follows. The first unit can be taken to answer the question (Q3) *What did John give to Tom?*, and the second, the question (Q4) *What did John give to Linda?*. The presence of contrastive topic marking on *Tom* and *Linda* indicates that the two questions should be understood as addressing the super-question (Q5), *John gave what to whom?*. In this case, (Q3) and (Q4) form a *strategy* for answering (Q5). The QUD for an utterance *u* is always the open question that is the most salient at the time *u* is made. Thus (Q3) is the QUD for *John gave the CAKE to Tom*, but once this question has been answered, the next subquestion of (Q5), namely (Q4), is addressed.

While the above notion of QUD is helpful for the analysis of focus, contrastive topic and other focus-related phenomena, it is a very local notion in the sense that it cannot tell us what accounts for the incoherence of the following mini-discourse:

- (8) JOHN ate the beans. My sister is a NURSE. Rose was gone for 10 MINutes!

We can retrieve a QUD for each sentence in (8) based on its focus structure: *Who ate the beans?*, *What does your sister do?*, and *How long was Rose gone?*, respectively. Moreover, each sentence provides a complete answer to its QUD, so moving from one sentence to the next should be a perfectly fine strategy. Yet (8) is not a good discourse.

[18] aims to push the notion of QUD further in order to account for the felicity of the focus of utterances in the contexts in which they are actually used. In a nutshell: the sentences in (8) do not hang together well because we do not understand them as participating in a strategy to answer the same, larger question. In more detail, [18] posits that the structure of open questions in a discourse can be represented with a *stack*. In (7), for example, when (Q3) is the QUD, it is on the top of the stack, just above (Q5). Once (Q3) is answered, it is popped from the stack and (Q4) is pushed on to the top. Once all sub-questions of (Q5) have been addressed, and (Q5) has therefore been answered, (Q5) is also pushed from the stack. The heart of Roberts's idea is to push the stack paradigm further: don't assume in a discourse like (7) that (Q5) is the question on the bottom. In a real discourse, (Q5) would in turn figure in a strategy to answer a yet larger question, and so on, all leading up (or down) to the Big Question: *What is the way things are?*. Of course, the Big Question is far too large to answer in one discourse, so interlocutors adopt the strategy of breaking it down into smaller questions. Exactly which sub-questions are relevant in a given discourse are determined by the *conversational goals* of the interlocutors; the sub-questions chosen signify a strategy

for achieving this goal.⁶ What is wrong with (8), then, is that it's not clear what sub-question of the Big Question is at issue; the individual segments do not form a clear strategy for answering an obvious question.

The rules of the language game constrain how different types of linguistic structures update the discourse context, with the following principal effects: (a) If an assertion is accepted by the interlocutors in a discourse, it is added to the common ground. (b) If a question is accepted by the interlocutors in a discourse, then it is added to the set of questions under discussion, and it becomes the immediate topic of the discussion. This in turn commits the interlocutors to a common goal, namely, finding the answer. By the principle of Relevance (described in (9) below), the interlocutors should attempt to answer the question as soon as it is asked. A member of the set of questions under discussion in a discourse is removed from that set iff its answer is entailed by the common ground or it is determined to be unanswerable.

- (9) A move m is **Relevant** to the question under discussion q , i.e., to $\text{last}(\text{QUD}(m))$, iff m either introduces a partial answer to q (m is an assertion) or is part of a strategy to answer q (m is a question). ([18], p. 21)

As a result of the above rules, all discourse moves must be relevant to the QUD, though they need not all answer (completely or partially) the QUD: it is enough if they are part of the strategy for answering the QUD.

2 Locating the correspondence

[18] has hypothesized that rhetorical theories and QUD-based theories are compatible and that understanding how they work together might give us a more complete pragmatic theory. The goal of this paper is to determine, on the assumption that both theories are correct, how the theories might be related. We assume that they cannot be independent: if a discourse move m_n , say, explains the content of a prior discourse move m_k , and both m_n and m_k are relevant to the QUD, then the fact that m_n explains m_k should also be relevant to the QUD—relating relevant discourse moves in irrelevant or off-topic ways would presumably lead to a very odd discourse. In other words, we assume that if the two are compatible, the discourse structures predicted by QUD-theories should correspond in some way to the structures posited by SDRT, such that the discourse graph for a whole discourse should shed light on the question structure for that discourse.

In what follows we explore two hypotheses about how the theories might be related: the first, which we will reject, assumes that rhetorical relations correspond to QUDs; the second, which will be shown to be more promising, assumes that QUDs can be associated with complex discourse units in SDRT graphs.

⁶ That the conversational goals and intentions of speakers are relevant for computing the pragmatic meaning of an utterance goes back at least to [9]. See also [10, 20, 7], and [12]. See [8] for a QUD-based theory that shares many features with [18], but is importantly different in ways we cannot consider in this paper.

2.1 Rhetorical relations as QUDs

An intuitive starting point for the comparison of SDRT and QUD is the hypothesis that rhetorical relations can be analyzed as QUDs, and vice versa. For example, when two discourse units, π_i and π_j , are related by Narration, the QUD associated with π_i , to which π_j would provide an answer, would be *What happened next?*. Similarly, if π_i and π_j are related by Explanation, the QUD associated with π_i , to which π_j would provide an answer, would be *Why?*. This hypothesis, which casts SDRT and QUD as two sides of the same coin, is frequently raised in our discussions with colleagues on this topic, and it appears to inform the work of [21] and [14].

We call the hypothesis that there is a one-to-one correspondence between discourse relation tokens and QUDs (or, more precisely, that each relation instance can be analyzed as a QUD at the point that the relation is added to the graph) **R-QUD**:

- (10) **R-QUD**: For discourse units π_i and π_j and rhetorical relation R in an SDRT graph, if $R(\pi_i, \pi_j)$, then there is some question $QUD(\pi_i)$ that π_i gives rise to and that π_j answers fully or partially.

Structural correspondence. R-QUD also suggests a structural correspondence between the hierarchical structures of SDRT and those of QUD. The hierarchical structures in SDRT, for example, are fully determined by the instances of rhetorical relations in the discourse, so if each rhetorical relation in an SDRT graph corresponds to a question in the QUD structure for the same discourse, then we might expect the set of questions to display a hierarchical structure that echoes that of the SDRT graph. This expectation is reinforced by the fact that SDRT's Right Frontier (RF) and Roberts's QUD stacks reflect (a part of) the hierarchical structure of a discourse and are designed to perform the same task, namely that of tracking live and salient issues in a discourse. Indeed, the two structures are similar, at least at first glance. The RF of a graph G consists of the last EDU introduced in G as well as any DU that is super-ordinate to a node on the RF via a chain of subordinating relations. Likewise, a QUD stack will include the most recent open question introduced in the discourse as well as any question that is a super-question of a question on the stack via a chain ordered by the sub-question relation.⁷ In SDRT, a new DU in a graph G that attaches to a node m along the RF of G will knock any other node that is subordinate to m off of the RF. Similarly, a new question Q in a QUD stack S that is a sub-question of another question Q_m on the stack, entails that any question Q_n , such that Q_n is higher on S than Q_m be popped from the stack (unless Q is a sub-question of Q_n as well). Finally, both SDRT and QUD posit constraints that require new moves to attach to/address a move on the RF/stack.

Despite these apparent similarities, there are also some obvious, and deep, dissimilarities between the RF and stacks. Most notably, the RF orders DUs, not relations, while a QUD stack orders questions, not answers. Given that R-QUD associates relations and questions (and the arguments to relations with answers), R-QUD will inevitably limit the role that either QUD stack or the the RF can play in determining the coherence and relevance of discourse moves. From the point of view of QUD, a new move, m , in a discourse d must be relevant to one of the open questions on the QUD-stack for d at the time that m is made. Yet if we assume R-QUD and SDRT's RF, then the nature and

⁷ This is the ordering adopted by [18] on page 15, clause (g.iii).

order of the open questions on a QUD stack for a discourse d would themselves be derived through rhetorical reasoning—there are no independent principles that would lead to exactly this set of questions for d in exactly the desired order. From the perspective of SDRT, a new EDU e in a discourse d must be attachable to some node along the RF of the graph for d at the time that the utterance that introduces e into the discourse is made. Yet if we assume R-QUD and QUD stacks, then coherence and relevance cannot be driven by the arguments to the rhetorical relations, for these correspond to *answers* not questions on the QUD stack. Given R-QUD, then, QUD and SDRT must be understood as giving two very different and incompatible models of what drives coherence and relevance in discourse.

Another structural consequence of assuming SDRT and R-QUD would be that the stack derived from the RF of a given discourse graph would not necessarily be ordered by the sub-question relation. (And conversely, a consequence of assuming QUD and R-QUD would be that the relations connecting nodes along the RF would be severely limited by the sub-question ordering on the QUD stacks.) Again, it is in part the QUD and the need to answer that QUD that is supposed to determine whether a given discourse move is coherent or not in a QUD framework; the sub-question relation is important because it keeps conversation on topic by simply breaking down the QUD into smaller QUDs. Examples like (11) show, however, that a combination of R-QUD and SDRT exclude the possibility of a QUD stack ordered by the sub-question relation.

- (11) We had so much fun in London! _{π_1} We got to see the Lion King! _{π_2} I've been wanting to go for a really long time _{π_3} and my mom finally gave me tickets for my birthday! _{π_4} We also got to ride on the big Ferris wheel _{π_5} ...

SDRT would predict the following structure for (11):

- (11)' Elaboration(π_1, π_2), Background($\pi_2, [\pi_3, \pi_4]$), Continuation(π_3, π_4),
Elaboration(π_1, π_5), Continuation(π_2, π_5).

From the first two relation instances, R-QUD would yield questions: (q1) *What did you do?* and (q2) *What makes that so exciting?* (or something like that). Suppose we stack these questions so that (q2), the more discourse-recent question, is on top of (q1). The fact that we (eventually) have Elaboration(π_1, π_5) and Continuation(π_2, π_5) shows that (q1), *What did you do?*, has not yet been fully answered when π_3 and π_4 are uttered, so it must still be on the stack of open questions. (q2), however, is not a subquestion of (q1). Still, the utterances of π_3 and π_4 are coherent and relevant in the discourse.

A third consequence of R-QUD is that it would result in a loss of information for SDRT. A node in an SDRT graph, as noted in section 1.1, can have incoming links from two different nodes, though it is unclear how such dependencies could be modelled using stacks. In (3), repeated here, π_2 appears to explain π_1 , and so could be taken to answer the question (q1) *Why?* (or *Why is Sam being punished?*), while π_3 appears to elaborate on π_1 , and so could be taken to answer the question (q2) *How is Sam being punished?* At the same time, π_3 also describes a result of π_2 and so could be taken to answer the question (q3) *what happened as a result?* or something like that.

- (3) Sam is being punished. _{π_1} She took her parents' car without permission, _{π_2} so they've grounded her for 2 weeks. _{π_3}

The stack architecture does not tell us where (q3) fits in to the stack; stacks are more naturally compatible with a tree-based structure, not the graph-based structure of SDRT. Nor does the stack architecture predict that reversing the order of (q1) and (q2) (and so that of π_2 and π_3) would make a difference to the discourse content, though clearly such a reversal would make a significant impact.

Another source of information loss for SDRT, given R-QUD, relates to CDUs. The need for CDUs is illustrated in (12):

- (12) I finally figured out the right baking process for cannelés. π_1 I left them at 210 for 20 minutes, π_2 then I turned them down to 190 for 30 minutes, π_3 and then I finished at 180 for the last 15 minutes. π_4

In (12), π_2 - π_4 together describe the correct baking process introduced in π_1 ; no single unit alone provides this information. In terms of questions, no single unit answers the question: *What is the right baking process?*. In other words, π_2 - π_4 form a CDU that serves as the second argument to the Elaboration relation that intuitively holds between π_1 and π_2 - π_4 . At the same time, the CDU composed of π_2 - π_4 has its own internal structure. Its component EDUs are related by Narration or Sequence, which imposes temporal constraints on the EDUs—the cooking steps cannot be followed in just any order. QUD stacks derived through R-QUD (or otherwise) do not provide the structure needed to capture the layers of discourse graphs that result from CDUs.

Relation instances. Even if we abandon a correspondence between the RF and stacks (which would render R-QUD far less interesting), the correspondence posited in R-QUD breaks down for several reasons. First, for certain rhetorical relations at least, we run into a problem of circularity in which the intuitive question has to mention the associated discourse relation or marker directly. For example, SDRT would posit the relation Contrast(π_1 , π_2) for (13), but it's unclear what question could replace the Contrast relation that wouldn't itself presuppose the same rhetorical relation.

- (13) [Pat]_F came to the party π_1 but [Mel]_F didn't. π_2

The most suitable implicit question to posit in this case would be *What does π_1 contrast with?*. Similar remarks can be made for instances of Parallel, Result (*What happened as a result?/And so?*), and Narration (*What happened then/next/after that?*). This calls into question the possibility of truly translating relations into independent questions.

An alternative question for (13), which would be more in line with the ideas presented in [17] and [6], is that π_1 and π_2 in (13) could be taken as addressing the same question *Who came to the party?*. Yet the fact that π_1 and π_2 both answer the same question doesn't justify the use of the contrastive marker *but*. In fact, a variant of (13) in which *but* is replaced by *and*, and π_1 and π_2 are related with Parallel, is acceptable as well, though the intuitive background question for such a variant would be the same: *Who came to the party?*. Similar remarks can be made for certain examples of Contrast that have possible Continuation-based variants. For instance, (14-a) and (14-b) could both be understood as answering the question: *What's going on with John?*, but this doesn't capture the intuitive difference between the examples, brought about by the presence of the contrastive marker in (14-a).

- (14) a. John called but he said he's running late.

- b. John called and he said he's running late.

In sum, if we assume R-QUD, the QUD framework runs into a problem because it does not easily extend to handle most co-ordinating relations.

Summary. The forgoing discussion reveals that a correspondence like that assumed by R-QUD would lead to a loss of information and an abandonment of basic principles from both SDRT and QUD. From the point of view of SDRT, if we assume R-QUD and QUD, then we must abandon the role that rhetorical reasoning plays in determining the coherence and relevance of individual discourse moves. We also lose information because QUD is not designed to handle triangular discourse structures (like that underlying (3)), CDUs, or coordinating relations. Nor is the correspondence posited by R-QUD favorable for QUD. If we assume SDRT + R-QUD, we must abandon the hypothesis that QUD-stacks of the sort posited in [18] guide the coherence and relevance of individual discourse moves. Moreover, the QUDs predicted from R-QUD would not look like the QUDs that a QUD-framework would predict otherwise. In [18] and elsewhere, it is suggested that focus marking indicates the QUD. In the case of (15), the second sentence has focus marking, but the question that can be derived from this marking (*What did John buy?*) is much more fine-grained than the question type that seems to correspond to the rhetorical relation Elaboration (e.g., *Can you tell me more?*). In fact, the question generated from focus is compatible with many rhetorical relations.

(15) John went to the store (π_1). He bought [apples]_F (π_2).

(16) Mary is mad at John (π_1). He bought [apples]_F (π_2). (instead of pears.)

SDRT would posit Elaboration(π_1, π_2) for (15), but Explanation(π_1, π_2) in (16). Yet in both cases, the question generated from the focus structure of π_2 is: *What did John buy?* Thus, in adopting R-QUD, we lose information from QUD as well.

2.2 Complex discourse units as QUDs

In this section, we take a step back and consider another hypothesis about how SDRT and QUD might interact. One of the intuitions behind [18]'s QUD account is that QUD stacks model the plan structure of a discourse. The rough idea is that given a specific discourse goal, which either triggers a QUD or is itself modelled as a QUD, a speaker or set of interlocutors will aim to answer this QUD by breaking it down in smaller and smaller QUDs until the QUDs become more manageable. As these smaller QUDs are answered, their collective answers work together to provide answers to larger QUDs and so on until the QUD triggered by the original discourse goal is answered. The question that we pose in this section is whether there are features of an SDRT graph that might reveal the planning structure of a discourse, yielding a correspondence more in touch with the planning aspect of [18]'s QUD stacks. The hypothesis that we consider is that CDUs and the relations between them yield the desired planning structure. In short, the idea is that CDUs manifest a certain topical cohesion that allows us to treat their contents as complex answers to implicit questions.

CDUs. A CDU, as briefly described in §1.1, is a collection of discourse units (EDUs, CDUs or a mixture of both) that work together to function as an individual argument to a discourse relation. (4) provides an example of a CDU in a Contrast relation; (12)

provides an example in an Elaboration relation. Further empirical motivation for CDUs comes from the behavior of anaphora:

- (17) One plaintiff was passed over for promotion three times (π_1). Another didn't get a raise for five years (π_2). A third plaintiff was given a lower wage compared to males who were doing the same work (π_3). But the jury did not believe this (π_4). ([2], p. 15)

As noted in [2], the anaphor *this* in π_4 can be resolved to π_3 or to the set $\{\pi_1, \pi_2, \pi_3\}$, but not to, say, π_2 alone. The second interpretation of (17), represented in Figure 1, requires the construction of a CDU $\pi_{t(1-3)}$ formed from $\{\pi_1, \pi_2, \pi_3\}$.

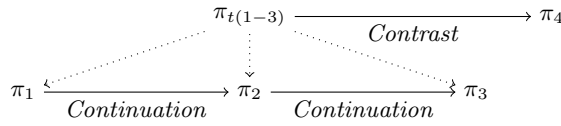


Fig. 1. SDRT graph for (17)

Thus CDUs are constructed when called for by the demands of rhetorical relations or anaphora, but we can also think of CDUs in a more general way as collections of DUs that share a thematic or rhetorical coherence. Suppose that two people are discussing a recent democratic presidential debate in the U.S. and one speaker is arguing that Bernie Sanders won the debate. She might provide a multi-step argument for her position, providing numerous justifications for it and elaborating on each justification in turn. We can think of her whole argument as yielding one large CDU whose members work together to provide support for the claim that Bernie Sanders won. Within this larger CDU, each justification for her main position will also yield a CDU whose members are the set of DUs that participate in the argument for this justification. For example, the speaker might offer as justification of her main claim the argument that Sanders provided a superior plan for attacking Wall Street. The members of the CDU associated with this justification would be the set of DUs that work together to support the claim that Sanders provided a superior plan for Wall Street. Within this CDU, the speaker might go on to provide multiple reasons for why Sanders' plan was superior, thereby giving rise to smaller CDUs and so on.

In this way, we can think of each CDU in a discourse as having a topic that glues its members together,⁸ although the topic of a given CDU is not (necessarily) derivable from the content of any explicit move in the discourse; it is rather a topic that emerges from considering what the discourse moves work together to accomplish. In addition, because CDUs are collections of *connected* DUs—i.e. not just arbitrary sets of DUs from the discourse—the set of CDUs in a discourse will naturally give rise to a partial order based on the subset relation: if a CDU π_i includes a CDU π_j , the set of members of π_j will be a strict subset of the set of members of π_i . A CDU cannot merely overlap another CDU without one being entirely included in the other.

⁸ Though see [3] for a discussion of the difficulties of defining discourse topic in SDRT.

CDU-QUD. We summarize the hypothesis that the content of a CDU fully answers an implicit QUD as follows:

- (18) **CDU-QUD_a:** For every CDU π_i in an SDRT graph there is some question QUD(π_i) that the the discourse units in π_i answer fully.⁹

To illustrate the correspondence, consider the textbook example (19) and the associated SDRT graph in Figure 2.

- (19) 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).

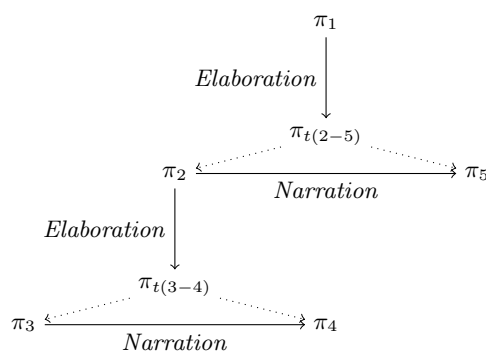


Fig. 2. SDRT graph for (19)

Putting aside deeper questions about how exactly to predict the construction of CDUs and their associated implicit topics, which would be an interesting topic for future research, let us assume that there are three CDUs in this structure: $\pi_{t(1-5)}$ ¹⁰, i.e., the discourse as a whole, $\pi_{t(2-5)}$, and $\pi_{t(3-4)}$. We derive $\pi_{t(2-5)}$ from the fact that if π_2 participates in the second argument of an Elaboration relation off of π_1 , then $\pi_{t(3-4)}$ does as well because $\pi_{t(3-4)}$ is simply an Elaboration of π_2 . These three CDUs can each be associated with a question as follows:

1. $\pi_{t(1-5)}$ is associated with a Q₁, e.g.: *What was John's evening like?*
2. $\pi_{t(2-5)}$ is associated with a Q₂, e.g.: *What did he do in the evening?*
3. $\pi_{t(3-4)}$ is associated with a Q₃, e.g.: *What did he eat?*

Note that the CDUs posited for (19) are ordered by the subset relation (if $\pi_i \in \pi_{t(3-4)}$, then $\pi_i \in \pi_{t(2-5)}$, etc.). This relation is moreover reflected in what appears to be a

⁹ Although what counts as a ‘full’ answer can depend on the context and interlocutors’ interests as in [18].

¹⁰ The CDU that represents the discourse as a whole is identical to the discourse graph as a whole, so we do not use special notation to label it.

sub-question relation that holds between Q_1 , Q_2 , and Q_3 (if an assertion answers Q_3 , it partially answers Q_2 , etc.).

Given the correspondence between the CDU structure of (19) and the proposed associated QUD structure, a reasonable hypothesis is that we can extend the correspondence posited by CDU-QUD_a with CDU-QUD_b.

- (20) **CDU-QUD_b**: If a CDU π_2 is a member of another CDU π_1 then QUD(π_2) is a subquestion of QUD(π_1).

CDU-QUD_b runs into problems when we consider a wider range of rhetorical relations than the ones in the above example, however. As mentioned above, $\pi_{t(3-4)}$ figures in $\pi_{t(2-5)}$ because $\pi_{t(3-4)}$ is related to π_2 via an Elaboration relation. This Elaboration relation in turn naturally gives rise to a set of questions ordered by the sub-question relation: even once we get down to $\pi_{t(3-4)}$, we're still talking about John's evening. The situation is less clear for CDUs figuring in other subordinating relations, e.g. Explanation, Consequence, Commentary, and so on.

Consider (21) and its discourse graph in Figure 3, in which the topic of the CDU provides an Explanation for π_1 :

- (21) Yesterday John and his wife went to the fanciest restaurant in Paris (π_1). It was John's birthday (π_2) and his wife wanted to spoil him (π_3).

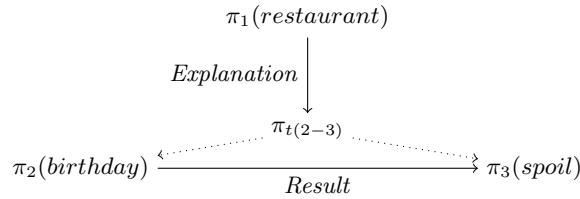


Fig. 3. SDRT graph for (21)

There are two CDUs in the above structure: $\pi_{t(1-3)}$ and $\pi_{t(2-3)}$. Because $\pi_{t(1-3)}$ is simply the graph as a whole, $\pi_{t(2-3)}$ is a subset of $\pi_{t(1-3)}$. CDU-QUD_b therefore predicts that the question associated with $\pi_{t(2-3)}$ will be a sub-question of the question associated with $\pi_{t(1-3)}$. This is incorrect: $\pi_{t(1-3)}$ would intuitively be associated with a question such as *What did John do yesterday?*, while $\pi_{t(2-3)}$ would likely be associated with a question such as *Why did he do this?*. Yet the latter question is not a sub-question of the former.

(22) illustrates the same point but with a different discourse relation:

- (22) The Fed lowered the prime interest again today for the third time in a month (π_1). Most economists greeted the move with skepticism (π_2) but were afraid to express this publicly (π_3). (Modified version of example (5) from [4])

Structurally (22) is like (21); only the rhetorical relations are different. Again, the question intuitively associated with $\pi_{t(2-3)}$ (e.g. *What was the economists reaction to this?*) does not seem to be a sub-question of the question associated with the whole graph (e.g. *What did the FED do today?*).

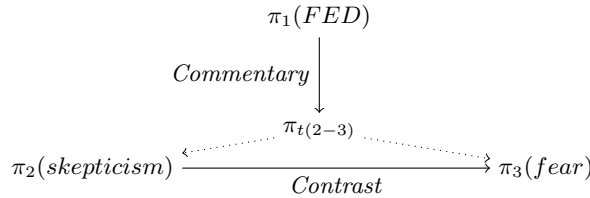


Fig. 4. SDRT graph for (22)

Because the structure of the attachment between $\pi_{t(1-5)}$ and $\pi_{t(2-5)}$ in (19) is identical to the structure of the attachments in (21) and (22), the sub-question relation posited by $CDU-QUD_b$ to hold between QUDs is not entailed by the structural relations in the discourse graph. The nature of the rhetorical relations, which are represented by the labels of the structural relations, must also be taken into account. As a result, $CDU-QUD_b$ is probably only applicable when the $CDU \pi_2$ (from the definition of $CDU-QUD_b$) attaches to the top node of the $CDU \pi_1$ with a relation that implies subeventhood. The semantics of the Elaboration relation, for example, entails that the eventuality described by the second argument of an Elaboration is a sub-eventuality of the eventuality described by the first.

Future directions. If the applicability of $CDU-QUD_b$ is indeed as restricted as the preceding discussion suggests, then this hypothesis entails that many QUDs derived from $CDU-QUD_a$ will not be linked to other QUDs semantically, though QUDs might still be ordered by a precedence relation. This is an unsatisfying result because it potentially prevents many QUDs in a discourse from working with other QUDs to form a plan or strategy for achieving a discourse goal.

One way to save the idea of a $CDU-QUD$ correspondence, which we ultimately find promising for at least certain kinds of discourse, is to loosen the requirement on the QUD side that QUDs be ordered by the sub-question relation and allow strategies for achieving discourse goals to be more complex than sequences of sub-questions. That is, we should abandon $CDU-QUD_b$ and replace it with a more general hypothesis. One idea, given the proposed correspondence in $CDU-QUD_a$, is that a QUD might be retrieved from the rhetorical structure of a discourse much like a question can be retrieved from focus structure in frameworks like [5, 6] and [18]. After all, CDUs are sets of connected nodes from a discourse graph, so the possible CDUs for a given discourse must be determined by the rhetorical connections at work in the discourse.

Associating QUDs with CDUs could have a value for SDRT as well. In particular, QUDs could be used to capture the topical cohesion of CDUs. This would give us a means for abstracting over the details of an SDRT graph to look at the higher-level strategy that a speaker, or group of interlocutors, has adopted to achieve her/their discourse

goal. We assume that this strategy will be apparent at least in cooperative, information-seeking discourses.

This high-level approach to thinking about the relation between QUDs, strategies for achieving discourse goals, and rhetorical structure reflects, we think, the spirit of [17] (last paragraph before the conclusion) and [18] (pp. 62-63) where the idea that each explicit move in a discourse should address a QUD, as suggested by R-QUD, is de-emphasized or outright rejected, and it is suggested that the connection between the rhetorical structure of larger chunks of discourse and strategies for achieving discourse goals should be further explored. Future work on the connection between CDUs and QUDs will depend, however, on these notions being more clearly defined in their respective frameworks than they currently are. What we suggest is that these aspects of QUD and SDRT might most beneficially be developed in tandem.

3 Conclusion

Adopting SDRT and some basic notions from [18]’s QUD framework as starting points, we have explored two hypotheses about how rhetorical structure and QUD structure might come together to provide a more general pragmatic theory. The first hypothesis, which posits a correspondence between rhetorical relations and QUDs, was rejected. The second hypothesis posits a correspondence between QUDs and more global features of an SDRT graph, namely how discourse units are thematically grouped together into larger, complex chunks of discourse. This hypothesis is, we have argued, more promising.

References

1. Asher, N.: Reference to Abstract Objects in Discourse. Kluwer Academic Publishers (1993)
2. Asher, N., Lascarides, A.: Logics of Conversation. Cambridge University Press (2003)
3. Asher, N.: Discourse topic. *Theoretical Linguistics* 30, 163–201 (2004)
4. Asher, N., Vieu, L.: Subordinating and coordinating discourse relations. *Lingua* 115, 591–610 (2005)
5. Beaver, D., Clark, B.: Sense and Sensitivity. Blackwell, Oxford (2008)
6. Büring, D.: On d-trees, beans, and b-accents. *Linguistics and Philosophy* 26, 511–545 (2003)
7. Carlson, L.: Dialogue games: an approach to discourse analysis (Synthese Language Library 17). D. Reidel, Dordrecht, The Netherlands (1982)
8. Ginzburg, J.: The interactive stance. Oxford University Press (2012)
9. Grice, H.P.: Studies in the way of words. Harvard University Press, Cambridge, MA (1989)
10. Grosz, B.J., Sidner, C.L.: Attention, intentions, and the structure of discourse 12(3), 175204 (1986)
11. Hamblin, C.: Questions in Montague English. *Foundations of Language* pp. 41–53 (1973)
12. Lewis, D.: Scorekeeping in a language game. *Journal of Philosophical Logic* 8(1), 339359 (1979)
13. Mann, W., Thompson, S.: Rhetorical structure theory: Towards a functional theory of text organization. *Text* 8, 243–281 (1988)
14. Onea, E.: Potential Questions in Discourse and Grammar. Ph.D. thesis, University of Göttingen (2013), habilitation Thesis

15. Polanyi, L.: A theory of discourse structure and discourse coherence. In: W. H. Eilfort, P.D.K., Peterson, K.L. (eds.) *Papers from the General Session at the 21st Regional Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society (1985)
16. Polanyi, L., Scha, R.: A syntactic approach to discourse semantics. In: *Proceedings of the 10th International Conference on Computational Linguistics (COLING84)*. pp. 413–419. Stanford (1984)
17. Roberts, C.: Context in dynamic interpretation. *Handbook of contemporary pragmatic theory* pp. 197–220 (2004)
18. Roberts, C.: Information structure in discourse: Towards an integrated formal theory of pragmatics. *Semantics & Pragmatics* 5, 1–69 (2012)
19. Rooth, M.: A theory of focus interpretation. *Natural Language Semantics* 1(1), 75–116 (1992)
20. Thomason, R.H.: Accommodation, meaning, and implicature: interdisciplinary foundations for pragmatics. In: Cohen, P.R., Morgan, J., Pollack, M.E. (eds.) *Intentions in communication (System Development Foundation Benchmark Series)*, p. 325363. The MIT Press, Cambridge, MA (1990)
21. Van Kuppevelt, J.: Directionality in discourse: Prominence differences in subordination relations¹. *Journal of semantics* 13(4), 363–395 (1996)