Constructing QUD trees

Arndt Riester
University of Stuttgart

Abstract

We discuss and combine representation formats for discourse structure, in particular 'd-trees' from QUD theory and SDRT graphs. QUD trees are derived from SDRT graphs, while changes must apply to QUD theory in order to allow for representations of naturalistic data. We discuss whether QUDs can replace discourse relations. We apply a new method for the identification of implicit Questions under Discussion (QUDs) to examples from an interview, and we address the status of non-at-issue material within our framework.

Keywords: *discourse structure, information structure, QUD, SDRT, annotation, non-at-issue*

1 Introduction

In this article, we discuss the role of *Questions under Discussion (QUDs)* in analyses of both discourse structure and information structure, as well as possible ways to integrate them into discourse trees. The QUD notion is primarily associated with the work by Carlson (1983), Grosz and Sidner (1986), Ginzburg (1994, 1996), Roberts (1998, 2012) as well as Beaver and Clark (2008). It is assumed that (both conversational and single-speaker/-writer) discourse contains implicit questions for each of the assertions made, which are thereby turned into answers. Theories of discourse structure, from Hobbs (1985) to contemporary theories like Rhetorical Structure Theory (Mann and Thompson 1988, Taboada and Mann 2006) or Segmented Discourse Representation Theory (Asher and Lascarides 2003) predict a structural analysis of discourse which is based on the identification of rhetorical relations, not of questions. However, recently, a number of new suggestions have been made how QUDs could be integrated into discourse structures; see, for instance Onea (2016), Velleman and Beaver (2016) or Hunter and Abrusán (to appear) but, in fact, the idea to relate the structuring of discourse to questions (or *topics*, which many scholars essentially see as their mirror images) goes way back to proposals by Klein and von Stutterheim (1987), von Stutterheim and Klein (1989), Polanyi (1988) or van Kuppevelt (1995). In the introduction to his article, van Kuppevelt (1995: 110) sketches his plan and the state of the art at the time:

> [A]n explication is given of our main subject of investigation, namely the way in which discourse structure results from the process of questioning. This is illustrated mainly by examples involving a succession of explicit question-answer pairs as they appear in question-answer dialogues. All principles which hold for explicit questioning are claimed also to apply to implicit questioning. However, an explanation of how implicit questions are reconstructed lies beyond the scope of the present paper.

We would like to point out that explicit questions are, in fact, different entities from implicit questions in important respects because explicit questions can change the information status of the content they consist of (they can introduce new material into the discourse), while it is crucial for the reconstruction of implicit questions *not* to change the discourse. In our paper, we specify the principles according to which implicit questions are identified, making use of a new QUD-based methodology for the joint analysis of naturalistic text in terms of discourse structure and information structure, described in Reyle and Riester (2016) as well as Riester et al. (under revision). Furthermore, our interest is aimed at different ways of how to represent *QUD trees*, i.e. discourse trees supplemented with implicit questions. A theoretical analysis of a given discourse can be visualized in different ways, and we show that representation formats are not necessarily self-explanatory and need to be interpreted.

The paper is structured as follows: in Section 2, we introduce some relevant terminology and the discourse representation formats from QUD theory and SDRT, as well as existing ideas how to combine them. We also touch upon the issue of whether there is a direct correspondence between discourse relations and QUDs.

Section 3 provides a solution to the general problem of how to identify implicit QUDs in a non-circular way, without relying on prosodic or morphosyntactic cues. We demonstrate the major traits of the procedure on examples taken from a spoken German radio interview (*SWR2 public radio, Interview der Woche*).[1] For the sake of enhanced intelligibility, we will mostly work with (our own) English translations, but we assume that the underlying German version gives rise to exactly the same structures.[2] This follows from the fact that our analysis procedure is based on interpretation rather than form, and we claim that it is cross-linguistically applicable (provided, of course, that the analyst is able to understand the text); cf. Riester et al. (under revision). After the conclusions and outlook, the paper contains an appendix with the complete analyses of QUDs, information structure and discourse structure in German and English as well as a representation of the QUD tree.

2 Prerequisites for a QUD analysis of natural data

There is widespread consensus among linguists that the structure of natural discourse cannot be adequately represented by a mere linear sequence of assertions (or other speech acts), but that discourse (no matter if text or dialogue) is hierarchically organised in the form of *discourse trees* (or even more complex structures), which group the elementary units of the discourse into sections, subsections etc. and, thereby, highlight their hierarchical relations. Various theories of discourse structure, such as Mann and Thompson (1988), Polanyi (1988), van Kuppevelt (1995) or Asher and Lascarides (2003), have explicitly dealt with the question how to construe such trees and how to identify and interpret the so-called *discourse relations* between the nodes of these trees. Over the years, a substantial amount of corpus data have been annotated using RST, SDRT and other frameworks.

2.1 QUD theory and d-trees

Discourse structures, albeit of a different kind, also play a role in Roberts's (2012) account of information structure. According to Roberts, natural discourse in general serves to answer hi-

---

[1] All data cited in this work are used with the kind permission of *Südwestrundfunk (SWR)*.

[2] Of course, this requires that translations must closely follow the source text; potentially, at the expense of the elegance of formulation. At this point, we do not intend to enter a debate about the possibility of perfect translation but we assume that the pragmatic concepts that play a role in our work are indeed universal and fully translatable.

erarchically ordered *Questions under Discussion (QUDs)*. Speakers devise *strategies* to answer general questions by breaking them down into more specific subquestions and answering them one by one. The processing of discourse is represented in terms of a dynamic stack of discourse moves. An important incentive of Roberts's framework is that QUDs can be used to account for the information structure (Rooth 1992, Lambrecht 1994, Hajičová et al. 1998, Schwarzschild 1999, Krifka 2007, Beaver and Clark 2008, Büring 2016) of the discourse units in their context: that part of an assertion which is already contained in the formulation of the current question is called the *background* (or sometimes *topic*[3]) while the part which provides the actual answer is the *focus*. Moreover, Büring (2003) has used Roberts's framework to explain the phenomenon of *contrastive topics*. A question with two *wh*-elements (e.g. *Who ate what?*) is often successively settled by providing answers to subquestions *about* each of the individuals in the relevant domain of one of the *wh*-words. *(What did X eat? What did Y eat? ...)* These individuals – at the same time *topics* and *contrastive / focused* entities – are called *contrastive topics*, and the sequence of superquestion and subquestion-answer pairs is, again, arranged in the form of a "d-tree" (Büring 2003), which can be interpreted as a static representation of Roberts's dynamic *QUD stacks*. In the following, we would like to combine the main advantages of two frameworks: the coverage of natural discourse enabled by SDRT, and the insights about QUDs and their relation to the information structure of answers, as provided by Roberts (2012) and Büring (2003).
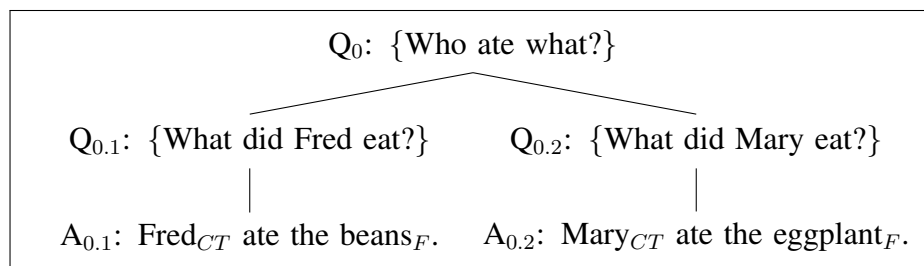
Q_0: {Who ate what?}

$Q_{0.1}$: {What did Fred eat?} $Q_{0.2}$: {What did Mary eat?}

$A_{0.1}$: Fred_{CT} ate the beans_F. $A_{0.2}$: Mary_{CT} ate the eggplant_F.

Figure 1: D-tree with subquestions and corresponding CT-F structures

We start by introducing Büring's *d-trees* consisting of a superquestion and several subquestions that are required to account for the contrastive topics in their answers. A key example from (Büring 2003: 520) is given in Figure 1. The tree is Büring's original depiction with just a few cosmetic changes. We have added *Q*s and *A*s to mark the questions and answers. Crucially, the subquestions in the tree are indicated by $Q_{0.1}$, $Q_{0.2}$, in order to express the fact that they are *entailed* by Q_0.[4] The contrastive topics ar those elements which occur in *(are topical with respect to)* the subquestions but are new *(focal)* with respect to the superquestion.

2.2 From SDRT graphs to QUD trees

The second type of discourse structure we consider is an SDRT graph. We use the often-cited mini-discourse example in (1), from Asher and Lascarides (2003: 8ff.) or Lascarides and Asher (2008), and its corresponding structural representation, given in Figure 2.

---

[3]There is still, to date, an unfortunate terminological confusion about these notions. In our annotations, topics are referential expressions contained in a background, see Section 3.4.

[4]There are two ways of defining *entailment* between questions: a question Q_0 entails its subquestion $Q_{0.1}$ iff every proposition that *completely* answers Q_0 also completely answers $Q_{0.1}$; alternatively, Q_0 entails $Q_{0.1}$ iff every assertion that *partially* answers $Q_{0.1}$ also *partially* answers Q_0, cf. Groenendijk and Stokhof (1984), Roberts (2012).
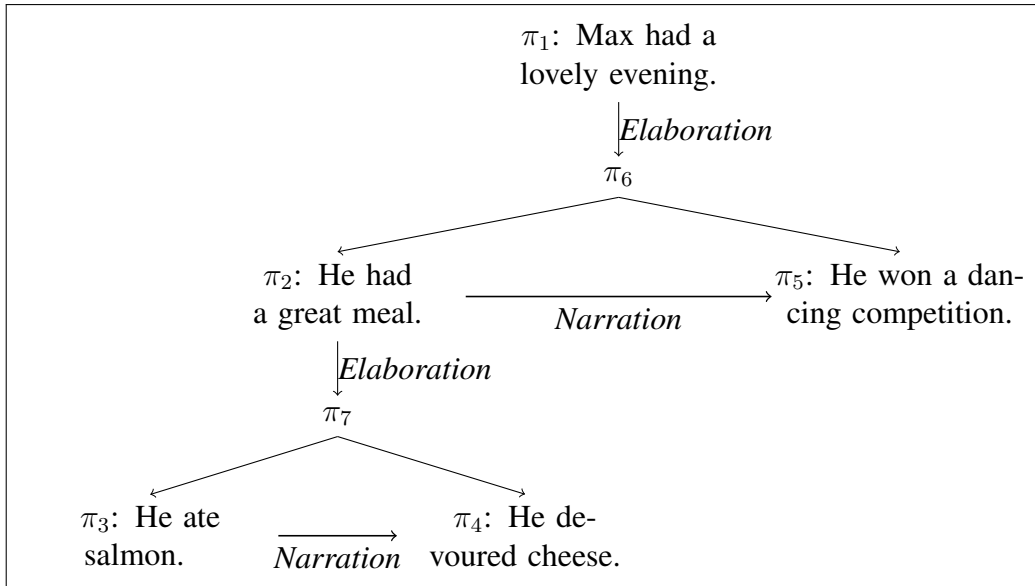
Figure 2: SDRT style discourse graph for (1)

(1) π_1: Max had a lovely evening.
          π_2: He had a great meal.
          π_3: He ate salmon.
          π_4: He devoured cheese.
          π_5: He won a dancing competition.

The structure in Figure 2 contains nodes for the assertions (or *discourse units*) in Example (1). It also contains the additional nodes π_6 and π_7, representing (implicit) *complex propositions* (Lascarides and Asher 2008) or *discourse topics*. Roughly, π_6 corresponds to Max's activities on that lovely evening, and π_7 to the courses that Max ate during his meal. The nodes of the graph are connected by vertical edges, representing subordinating discourse relations (e.g. *Elaboration*) or links between a complex unit and its parts, and horizontal edges representing coordinating relations (e.g. *Narration*). From the representation in Figure 2, it is only a small step to the *QUD tree* in Figure 3, which contains a QUD for each of the assertions.

To indicate question-answer congruence, we have changed the names of the nodes into A and Q. The discourse topics have been translated into their question equivalents, $Q_{0.1}$: {*What did Max do on that lovely evening?*} and $Q_{0.1.1}$: {*What did Max eat during his meal?*}. Moreover, the tree now contains a root question Q_0 for the intitial assertion, {*What is the way things are?*}, as suggested in Roberts (2012).[5] Answers share the index of their question parent. Parallel, partial answers are indicated as, for instance, $A_{0.1'}$, $A_{0.1''}$ etc. We have put the QUDs in curly brackets to indicate that they are implicit entities, while explicitly asked questions could occur – in a then different discourse – in the very same position without brackets. Note that, other than in Roberts (2012) and in Büring's *d-tree* (Figure 1) we allow for the situation that question nodes may occur as children of answer nodes, thereby siding with van Kuppevelt (1995), Onea (2016) or Velleman and Beaver (2016). Van Kuppevelt (1995: 119) refers to answer nodes that dominate a question as *feeders*. In turn, we would like to refer to questions that are dominated

---

[5]Our questions and the structural analysis are by and large comparable to the assumptions made by Hunter and Abrusán (to appear: Sect. 2.2). The slightly different formulation of our questions is justified by the constraints formulated in Section 3.2 below.
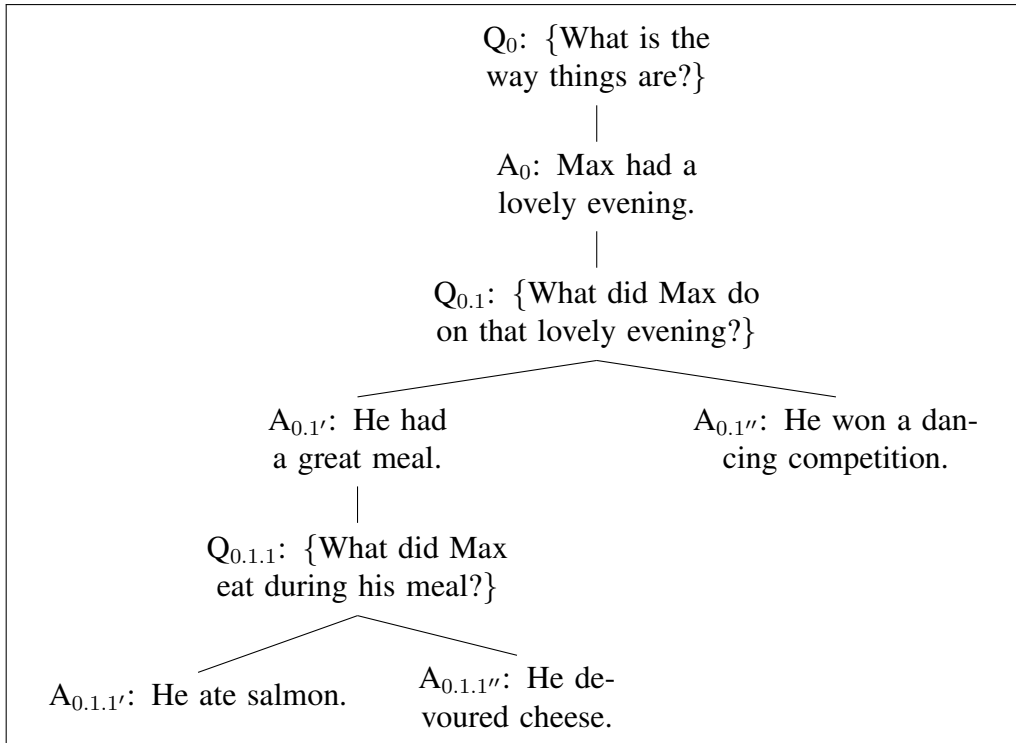
Figure 3: QUD tree for (1)

by answers as *anaphorically dependent questions*, since they necessarily build on given material from the feeder.

It is often the case that anaphorically dependent questions are entailed by the question of their feeder (e.g. in Figure 3, $Q_{0.1.1}$ is entailed by $Q_{0.1}$, and $Q_{0.1}$ is entailed by Q_0, hence the indices), but this is neither necessary to preserve coherence nor is it, to our mind, a realistic assumption for naturally occurring discourse. For instance, it would sound perfectly acceptable – and normal – to replace the branch below $A_{0.1'}$ with the one in Figure 4, headed by the, non-entailed, question Q_1.

This, in turn, amounts to yet another deviation from Roberts (2012), according to whom in-coming discourse moves have to be *relevant* and contribute to answering the questions which are already on the current QUD stack, but, as a matter of fact, people often insert strictly-speaking irrelevant things into the ongoing discourse, which nevertheless exhibit a topical connection to what has just been said. This insight, again, is already pointed out in van Kuppevelt (1995), Onea (2016) as well as Velleman and Beaver (2016).

Another issue that we haven't addressed so far is the fact that the horizontal *Narration* edges got lost during the transition from Figure 2 to Figure 3, thereby turning the initial graph structure into a proper tree. This touches upon the debate whether discourse relations can be replaced by QUDs or not. Hunter and Abrusán (to appear: Sect. 2.1, 2.2) investigate two versions of this hypothesis. They reject a direct one-to-one correspondence between discourse relations and QUDs, partly for the limitations of Roberts's strict QUD framework discussed above, and partly for the fact that SDRT enables discourse units to have two incoming edges, see e.g. π_4 and π_5 in Figure 2, while an answer can only have one immediate QUD as its parent. Instead, Hunter and Abrusán suggest that QUDs should, above all, occur when they address issues that can be represented by a *complex discourse unit*, "having a topic that glues its members together".
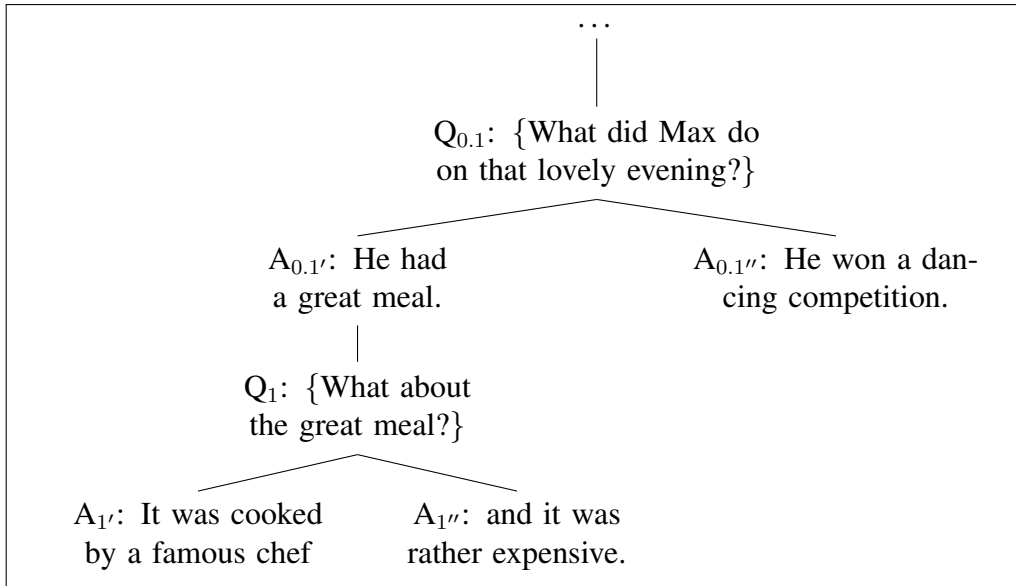
Figure 4: A non-entailed, anaphorically dependent question (Q_1)

While we follow Hunter and Abrusán in most of their arguments, the conclusions that we draw are slightly different. We claim that, for a purely structural purpose, *subordinating* discourse relations can be replaced by questions (for instance, an *Explanation* can be formulated using a *why* question, an *Elaboration* by means of a *what-about* question. This does not mean that the two are always equally informative. As Hunter and Abrusán point out, different discourse relations can sometimes translate into the same QUD. On the other hand, the QUD defines the information structure of an answer while the discourse relation does not necessarily do so.

It is sometimes suggested that also structurally coordinating relations like *Narration* can be treated as simple *and-then-what* (or *what-happened-next*) questions, cf. Velleman and Beaver (2016). In this case, however, it is unclear how the resulting "QUD graph" should be interpreted; certainly not by use of a stack. By contrast, we think that there is actually a gain in going from discourse graphs to actual discourse trees, which express the hierarchy of discourse moves in a unanimous way. This, however, means that a simple *and-then-what* question will not suffice. Instead, in order to maintain the coordinating nature of a narrative sequence, we draw on a proposal made already in von Stutterheim and Klein (1989: 43), according to which temporal progression can be represented by a series of parallel questions about different times, see Figure 5.

Note that, while Figure 5 reflects the order of the overt discourse moves according to their time of utterance, the representation says nothing about the temporal order of the reference times (t_2, t_3 etc.) and is therefore compatible with both the default situation of temporal progression from left to right and with other temporal configurations. This information must be encoded separately. Moreover, since we are primarily interested in the information structure of the assertions, the additional, temporally located questions of Figure 5 can be considered as optional because there are no overtly contrastive temporal expressions in the assertions about the eating or dancing events which could participate in a marked information-structural contrast like the one shown in Example (2).

(2) $First_{CT}$ he ate $salmon_F$.
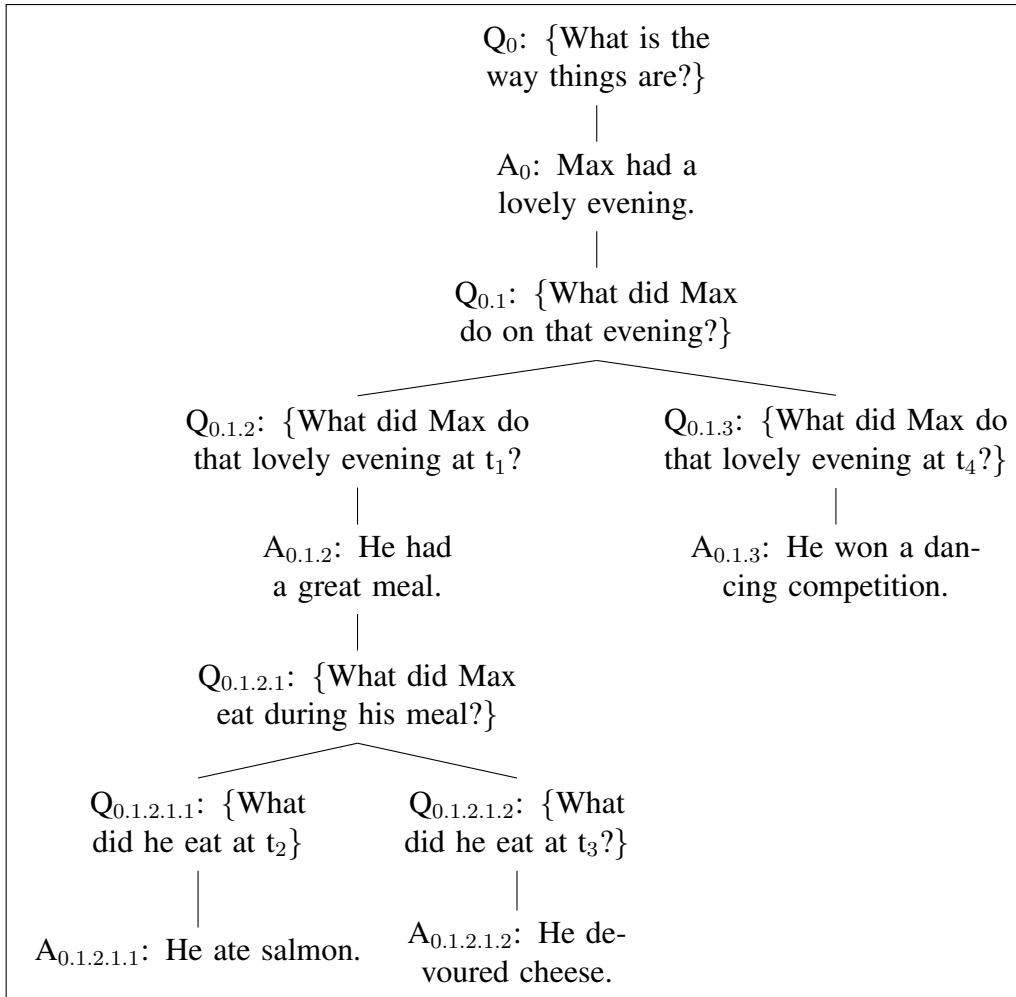 $Then_{CT}$ he devoured $cheese_F$.

6

Figure 5: QUD tree with temporally located questions

A legitimate question, of course, is: based on what rules have the QUDs in the above figures been formulated, in the first place? This is precisely what we are going to deal with in the next section.

3 Information-structural constraints on QUDs

In the previous section, we have mentioned a few essential aspects of discourse trees and the role of questions in the structuring of discourse. We now shift our attention away from discourse theory and towards the practical task of reconstructing QUDs in natural discourse. In doing so, we follow the spirit of van Kuppevelt (1995: 109), who proposed that topicality (or questions) should be seen as "the general organizing principle of discourse". However, we go a step further in that our method not only accounts for the structure of discourse but, at the same time, for the information structure of the individual assertions the discourse consists of. In this article, we confine ourselves to the basic traits of the procedure. For a comprehensive annotation manual, see Riester et al. (under revision), and Reyle and Riester (2016) for a formal account in terms of UDRT. To illustrate the constraints, we choose examples from a section of a spoken German

radio interview.[6] The entire section with annotations, English translation and a QUD tree can be found in Appendix A.

3.1 Segmentation

The first step of the analysis procedure is the segmentation of the text into separate *discourse moves*, which are mostly assertions. Each assertion is marked by a letter *A*. Segmentation will, to a certain extent, ignore punctuation or a syntactic notion of clause. It will come as no surprise that assertions are split at clause-level coordinations. However, a split is sometimes also required below the clause level, as in the case of VP or DP conjunctions or disjunctions, see (3).[7] In cases like these, it may be handy to reconstruct the elided (crossed out) material, in order to make each single assertion transparent.

(3) A: Da   sollen Fenster in Teeküchen   vorgeschrieben werden,
         There shall   windows in staff.kitchens prescribed   be
         'It [the bill] will prescribe having windows in staff kitchens'

     A: oder auch die Helligkeit am   Heimarbeitsplatz soll   vorgeschrieben werden.
         or   also the brightness at.the home.workplace shall preserved   be
         'and even the brightness of home workplaces.'

On the other hand, we will not separate sentential arguments from their embedding matrix verbs, as shown in (4). The embedding part of such constructions will, later on, typically be analyzed as *non-at-issue* material, cf. Section 3.6 and Riester et al. (under revision) for further details.

(4) A: Jetzt heißt es, das Kanzleramt hat diese Verordnung gestoppt.
         Now means it   the Chancellery has this   bill   stopped
         ' Now they are saying that the Chancellery has stopped the bill.'

3.2 QUDs consist of given material

Suppose that the segmentation of a discourse has yielded a list of assertions. It is now the task of the analyst to provide, for each assertion, a QUD which is in fact answered by that particular assertion, and to arrange the questions and assertions in order to form a discourse tree. The identification of implicit QUDs has often been dismissed as a hopeless, or at least highly speculative, task. This is understandable if assertions are considered in isolation. Without knowing the information structure of a sentence (i.e. its focus span), there are (at least) as many possible QUDs as there are syntactic constituents in the sentence. Hence, in Reyle and Riester (2016) and Riester et al. (under revision), we identify a number of discourse-contextual *constraints* that make the QUD identification task tractable and deterministic. In the following, we give a short sketch of the account. The constraints are derived from the focus literature of the past decades, in particular, Rooth (1992), Schwarzschild (1999) and Büring (2008). The first constraint on the formulation of QUDs is given below:

---

[6]*SWR2 Interview der Woche*, 2015/02/28 (6:30 p.m.), with Andrea Nahles (SPD), Federal Minister for Labour and Social Affairs. The interview is part of the *Stuttgart SFB732 Silver Standard Collection* (Eckart and Gärtner 2016).

[7]For a DP coordination, see the answers to $Q_{4.1.2}$ in the appendix.

Q-A-Congruence
QUDs must be answerable by the assertion(s) that they immediately dominate.

Q-A-CONGRUENCE is, of course, rather trivial. It means, for instance, that a sentence like (7) is an answer to any of the questions in (5) but not to (6).

(5)   a.   {What happened?}
        b.   {What about him?}
        c.   {Who literally suffocated?}

(6)   {Who owns a bicycle?}

(7)   A: He literally suffocated.

Note that our formulation of Q-A-CONGRUENCE is deliberately kept weak and does not make any reference to *alternative sets* (Rooth 1992) or *question sets* (Hamblin 1973, Groenendijk and Stokhof 1984) precisely because we do not want to presuppose the focus-background structure of the answer. All we want is a question that can be answered by the assertion *once its information structure has been determined*. A second constraint, which takes the discourse context into account, is the following one:

Q-Givenness
Implicit QUDs can only consist of given (or, at least, highly salient) material.

Q-GIVENNESS restricts implicit QUDs in a way that renders them less flexible than explicit ones: they are not allowed to introduce new material into the discourse themselves. The only semantic material they can consist of – ignoring function words; in particular, *wh*-pronouns – is material that is already *given*[8] in the discourse. We should point out, that Q-GIVENNESS is, of course, the flipside of Schwarzschild's (1999) GIVENNESS constraint, which does not mention QUDs but requires that *discourse-new* material be *F(ocus)-marked*. Since the focus-marked material of an answer corresponds to the *wh*-expression of its QUD, it follows that new material is banned from occurring inside the QUD.[9] Let us have a look at a simple example.

(8)   a.   A: And all I can say is that his condition was extremely bad during his last years.
        b.   A: He literally suffocated.

As we said, when sentence (8b) is considered in isolation, it can function as the answer to any of the questions in (5). However, in the context of sentence (8a), it is obvious that (5c) is ruled out as QUD, since it would introduce the sequence *literally suffocated* as new information and fail to provide a link to the previous discourse. Neither does (5a) contain any discourse-given material, but it is not ruled out by the above constraint. In order to capture the intuition that question (5b), which does in fact connect to the previous discourse via the pronoun *him*, should be preferred over (5a), we need a third constraint:

---

[8]We assume a strict definition of *givenness* in terms of referential or lexical availability of a word/phrase or discourse referent in the previous discourse context, which goes hand in hand with cognitive activation, cf. Baumann and Riester (2012), Riester and Baumann (2017). Note that activation or salience can occasionally arise without explicit mention, see Q_5 and Q_7 in the appendix.

[9]Occasionally, in particular at the beginning of a discourse, Q-GIVENNESS is overruled by the effect of *presupposition triggers* (e.g. clefts). Presupposition accommodation (Beaver 1999, Beaver and Zeevat 2007) can lead to a narrower – though never wider – focus than predicted by our principles. In order to capture accommodation phenomena, our approach must be enhanced by language-specific rules.

Maximize-Q-Anaphoricity
Implicit QUDs should contain as much given (or salient) material as possible.

MAXIMIZE-Q-ANAPHORICITY corresponds to what has been discussed in the literature under the names AVOIDF (Schwarzschild 1999) or MAXIMIZEANAPHORICITY (Büring 2008). Their purpose is to ensure *discourse coherence*. In other words, two consecutive assertions are expected to be related to each other to a substantial extent, and this is mediated via the QUD. Note that the constraints, taken together, prefer question (5b) above either (5a) or (5c). However, they do not actually dictate that (5b) is the optimal QUD in general. In principle, it is possible to increase the degree of discourse coherence, by formulating other questions that contain even more given material, e.g. the ones in (9).

(9)   a.   How bad was the situation[10] for him?
        b.   How bad was his condition?
        c.   What was his bad condition like?

The fact that there are still so many possible ways of formulating the actual QUD may have added to the impression that the identification of QUDs is a highly speculative matter. However, if we abstract away a bit and look for a common pattern, we note that questions (9a)-(9c), as well as (5b), all ask for an answer of property type; in other words, they give rise to the same information structure of their answer. The lesson we learn from the current discussion is that there normally isn't a single way of formulating implicit QUDs – which are semantic objects, after all – since language allows for synonymous formulations. What is essential however, is that it is possible to identify a QUD as an object of *meaning* that does account for the discourse structure of a sequence of assertions, and for the information structure of the assertion that answers it.

What is meant in the definition of MAXIMIZE-Q-ANAPHORICITY by "as much given (or salient) material as possible" is mainly that all the given material *in the answer* must show up in the question. (Certainly not all material available in the previous discourse context!) The interplay of the constraints also gives an answer to the question about the nature of QUDs. It is sometimes suggested that QUDs "arise" or "follow" from the previous context and then "guide" the way how the subsequent discourse is going to evolve. Nothing could be more wrong than that. In fact, the only rule that speakers or writers must observe when formulating their next move is to think about some topical connection to whatever was said before,[11] but in all other respects they are free to formulate their own continuation of the plot. So, if ever QUDs can be said to "arise" then they do so only out of the interplay between the current discourse and the next assertion made by the speaker.

3.3   Compact discourse trees

The discourse constellation that is associated with Example (8) is shown in Figure 6, in two variants. A practical problem when analysing larger sections of text in terms of QUD trees is that the trees quickly grow very deep. In order to minimize the problem, we introduce a new

---

[10]We assume that very general expressions like *situation* are always salient, i.e. *quasi-given*, although they might not be *given* in the strict sense of having been previously mentioned. Corpus examples of *quasi-given* (but strictly speaking *discourse-new*) nouns are discussed in Riester and Piontek (2015: pp. 243ff.).

[11]But even when this rule is not observed, a speaker can strictly-speaking never entirely "drop out of discourse". The worst thing that can happen is that a speaker utters an all-new statement without any connection to the previous context, by which she simply jumps back to answering the root question *What is the way things are?* and, thereby, loses access to all previously given information.
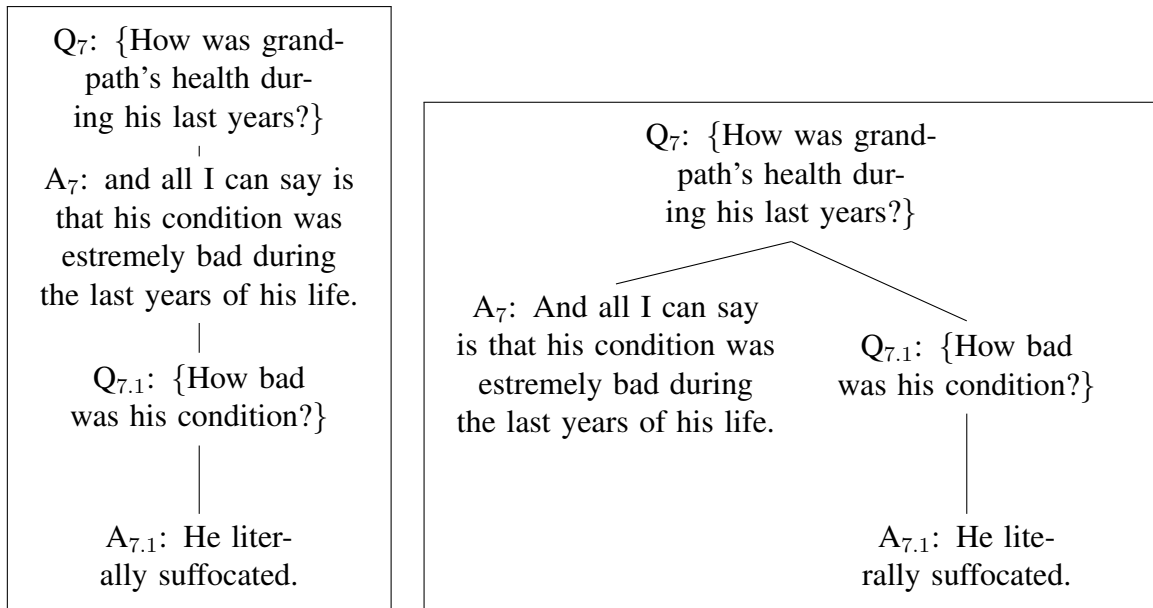
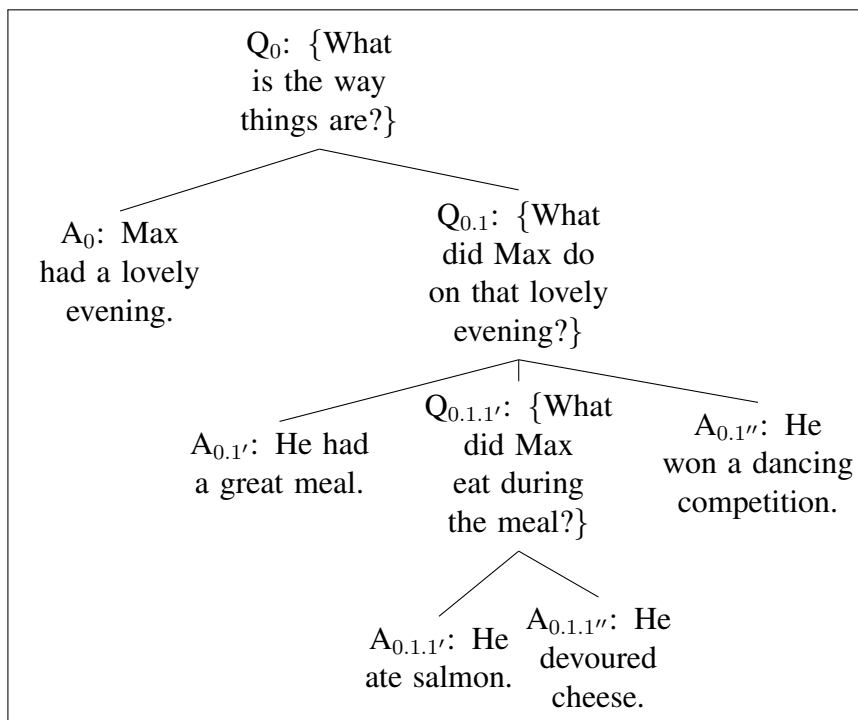Figure 6: QUD trees (deep / compact) of the sequence in (8)

Figure 7: Compact QUD tree for (1)

representation format for discourse trees, which is also used (in textual form) in the analyses given in the appendix. The left "deep" tree in Figure 6 is the SDRT-based format used already in Section 2.2 (as, for instance, in Figures 3 or 4). The tree on the right is meant to express the same analysis in the *compact QUD-tree format*, in which anaphorically dependent questions (here, $Q_{7.1}$) occur to the right of their feeder (A_7). Note that in this format it does not matter, whether the anaphorically dependent questions are entailed, like in the current example, or not

(as any of the questions Q_1-Q_8 in the appendix).[12] A compact QUD-tree representation of the SDRT example is shown in Figure 7.

Apart from the space-saving aspect of the compact representations, a big advantage is the clear spearation between questions, forming the non-terminal nodes, and assertions, which constitute the terminal nodes of trees, and which can be read off from left to right.[13]

The three QUD constraints defined so far describe how Questions under Discussion should be formulated: they should allow for Q-A-CONGRUENCE, they should not contain any non-salient information (Q-GIVENNESS) and they should consist of a maximal amount of given information (MAXIMIZE-Q-ANAPHORICITY). Taken together, this also does have a crucial influence on the attachment site of an incoming QUD and its answer: a QUD must be subordinate to its feeder A, i.e. the assertion which provides its respective antecedental material, where *subordinate to A* means *below A* in the deep format and *to the right of A* in the compact format. Following common assumptions in discourse semantics, the only material that is accessible to anaphoric retrieval is material at the *right frontier* (Polanyi 1988, Webber 1991, Asher 2008, Afantenos and Asher 2010) of the current discourse tree. Any new piece of information should therefore attach *as low as necessary* in order to enable anaphoric retrieval, but *as high as possible* in order to allow the speakers to conclude a discourse section, to return to the starting point of a conversation or the main question of a text, and to raise new issues. This is captured by the constraint BACK-TO-THE-ROOTS.

Back-to-the-Roots
In the compact QUD-tree format, an incoming QUD (and its answers) must attach to the right of the lowest antecedent of its given content, but otherwise as high as possible.

3.4 Information structure

In the previous sections, we have been talking about information structure in a loose way. We would now like to define it more precisely. Assuming the question-answer sequence from Figure 6, the information structure of assertion A_7 will be the one shown in (10).

(10)   [He$_{BG}$ [literally suffocated]$_F$]~

Following standard assumptions in the literature, we define the *focus (F)* of an assertion as that part of an assertion which actually answers the QUD. Moreover, the remaining part, i.e. the material which already occurs in the QUD, is called the *background (BG)*. Note that our framework differs terminologically from the one by van Kuppevelt (1995) or Hajičová et al. (1998), in which backgrounded material is generally called *topic*.[14]

Together, the background (which may be empty) and the focus (which cannot be empty) form a *focus domain* (Rooth 1992, Büring 2008), which is marked using the ~ operator. Focus domains do double duty. On the one hand, a focus domain matches the structure of the QUD, since the backgrounded part of an assertion must recur inside the corresponding QUD. In other words, a QUD must have the same focus domain as its answer, except for the *wh-*/focus part.

---

[12]In Figure 12 in the appendix, we attach non-entailed questions via dashed lines and entailed questions via solid lines but this is actually redundant, since non-entailment is already indicated by means of the different indices.

[13]The only questions which are also terminal nodes are overt questions that remain unanswered and are therefore "left dangling". See $Q_{2.1}$ and Q_3 in the appendix.

[14]In the appendix, we have optionally marked referring expressions in the background as *aboutness topics (T)* (Reinhart 1981, Krifka 2007). Note that we assume that not every background has to contain an aboutness topic, but that every (non-contrastive) aboutness topic is necessarily backgrounded.

On the other hand, focus domains also play a role in the highlighting of parallel structures, as discussed in the next section.

3.5 Parallelism

Besides the givenness-related constraints already defined, there is another important cue for the reconstruction of QUDs, viz. the identification of recurring patterns in the discourse, i.e. the search for semantically parallel structures. This is formulated in the following constraint.[15]

Parallelism
The background of a QUD with two or more parallel answers consists of the (semantically) common material of the answers.

Two types of parallelisms play a role: simple ones, which semantically differ in only one position per assertion, and complex ones, which differ in two positions. An example of a simple parallelism is given in Example (3), reproduced below as (11).

(11) a. $A_{1'}$: It will prescribe having windows in staff kitchens
      b. $A_{1''}$: and ~~it will~~ even ~~prescribe~~ the brightness of home workplaces.

Identifying semantically identical material or, as in Example (11), reconstructing elliptical material, in turn, helps us formulate the shared QUD: the constant material (i.e. the background) must re-occur inside the QUD, while the alternating parts (the foci) correspond to a wh-word in the QUD, which we formulate, for the current example, as the root of the tree in Figure 8, which indicates both discourse and information structure.

Q_1: {What will the bill prescribe?}

$A_{1'}$: [It will prescribe [having windows in staff kitchens]$_F$]~

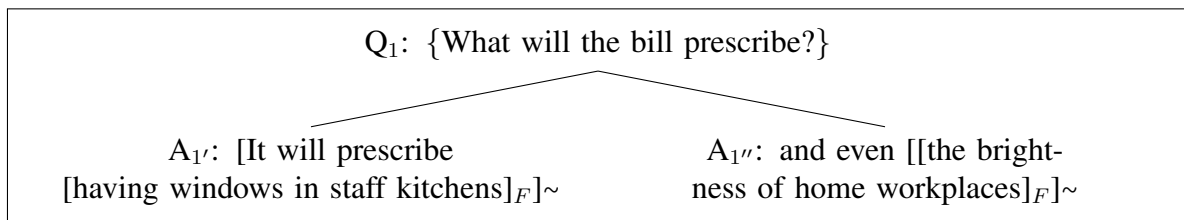$A_{1''}$: and even [[the brightness of home workplaces]$_F$]~

Figure 8: Discourse tree for the parallel answers in (11)

A second, more complex form of parallelism, which we already encountered in Section 2.1, is the one in which two assertions differ in two positions. Unless we have reasons to believe that we are merely dealing with the special case of a complex (Krifka 1992: 21) or discontinuous focus (e.g. Gussenhoven 1999: 50), i.e. a single focus that merely consists of two disjoint parts, we follow the approach by Büring (2003) in assuming that such structures give rise to pairs of contrastive topic and focus. Beyond Büring's simple examples (Figure 1), contrastive topics may, for instance, take the form of antecedent clauses of (hypothetical) conditionals (Iatridou 1991: 51). In those cases, the information structure is of the form: *[If A_{CT} then B_F]~; [if C_{CT} then D_F]~*. A corpus example of this kind is shown in Figure 9.

3.6 Non-at-issue material

Not all information contained in a sentence takes part in the actual focus-background divide signalled by the QUD. Potts (2005) has drawn the attention to the phenomenon of *conventional*

---

[15]See Reyle and Riester (2016: 18) for an example showing that PARALLELISM can override Q-GIVENNESS.
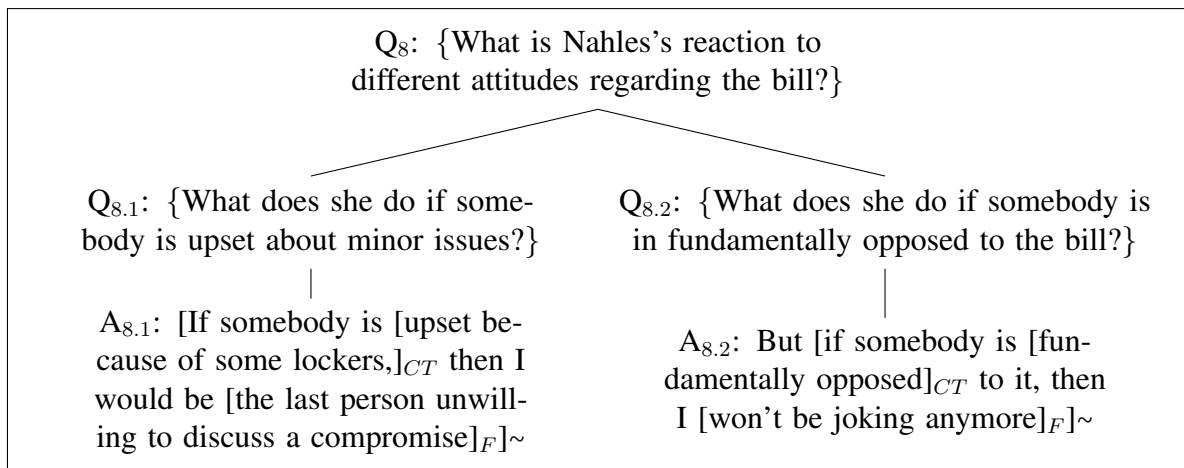
Q8: {What is Nahles's reaction to different attitudes regarding the bill?}

Q8.1: {What does she do if somebody is upset about minor issues?}

A8.1: [If somebody is [upset because of some lockers,]$_{CT}$ then I would be [the last person unwilling to discuss a compromise]$_F$]~

Q8.2: {What does she do if somebody is in fundamentally opposed to the bill?}

A8.2: But [if somebody is [fundamentally opposed]$_{CT}$ to it, then I [won't be joking anymore]$_F$]~

Figure 9: Conditionals that take the form of contrastive-topic/focus pairs

implicatures, i.e. meaning components that are projective in the sense that their contribution survives after being embedded under negation and other operators, but which are nevertheless distinct from presuppositions in that conventional implicatures do not impose any identifiability constraints on the common ground or discourse context. On the contrary, they seem to represent genuinely new information. Among triggers of conventional implicatures, Potts (2005) originally lists *supplements* (appositions, non-restrictive relative clauses, speaker-oriented adverbs) as well as *expressives* (honorifics and expressions describing the speaker's positive or negative attitude towards the asserted content). He makes it clear that a crucial trait of conventional implicatures is that they are independent from the *at-issue*, or regularly asserted, content. Another very frequent phenomenon that can presumably be subsumed under conventional implicatures are *evidentials* (Faller 2002, Murray 2010, Tonhauser 2012), which are grammaticalized in many languages or expressed by means of *I think that, John said that* etc. Examples (12) and (13) show corpus examples of non-at-issue expressions (in grey).

(12)   A2: Now they are saying that, the Chancellery has stopped the bill.   (EVIDENTIAL)

(13)   A4.1.2.1: [. . .] and it is due to numerous measures in the past 40 years called occupational safety measures and workplace regulations – that's indeed how they are called – that, Thank God, the numbers of casualties, injuries and accidents has decreased massively.   (PARENTHESIS, EXCLAMATION)

The distinction between *at-issue* and *non-at-issue* content has received considerable attention within the semantics-pragmatics community in recent years. For instance, the work by Simons et al. (2010) and Tonhauser et al. (2013) discusses *projective content*, subsuming conventional implicatures, presuppositions as well as anaphora and backgrounded material. In the light of this, it seems worthwhile to point out that using the term *non-at-issue content* interchangeably with *conventional implicatures*, as it has become quite popular and as we will also do in the current paper, carries a mild terminological conflict since *backgrounded content* is, of course, also not "at issue" in the strict sense of not providing the answer to the QUD, but it is, of course, to be kept distinct from triggers of conventional implicatures.

However, we argue that another phenomenon should be treated as non-at-issue material, viz. *adjunct phrases*, i.e. certain locative, temporal and other adverbials, which are traditionally distinguished from *argument phrases* with respect to their syntactic optionality. Adjunct

phrases (unless in focus) are optional in the sense that they can be left out without affecting the interpretability and the truth value of the main proposition denoted by the utterance. In fact, we propose that optionality could figure as the defining criterion for the type of non-at-issue material that we have in mind: to be non-at-issue with respect to the current Question under Discussion means nothing else than to be optional in the current assertion. In semantic terms, we therefore define non-at-issue material as follows:

Non-at-issue (relative to Q)
An expression X whose denotation is discourse-new and which is contained in in an assertion A is NON-AT-ISSUE with respect to the current Question under Discussion Q iff X is optional with respect to Q, where optional means that under deletion of X, A is still an answer to Q.

There is a point that we need to clarify, though: to be non-at-issue (with respect to the current QUD) does not – in the least – mean to represent unimportant information.[16] If a speaker formulates an utterance that contains non-at-issue material, she is simply choosing a compact way of conveying two (or more) assertions in one go, and the most important piece of information may very well be formulated in the form of a supplement. To understand this point, consider example (14), overheard from German radio news.[17]

(14)   In Deutschland sterben viele Patienten gegen   ihren Willen im   Krankenhaus.
       In Germany   die   many patients against their will   in.the hospital.
       'In Germany, many patients die at the hospital(,) against their will.'

Example (14) has two drastically different interpretations that can be disambiguated in terms of information structure, i.e. in terms of assumptions about their implicit QUD. The rather worrisome interpretation is the one in (15a), which suggests that German hospitals are horrifying places in which constantly people are getting murdered, although they explicitly fought against it. The more harmless – hopefully intended – reading is the one in (15b), which merely expresses the sad fact that in Germany a lot of people end up dying in hospitals despite their explicit wish to die at home.

(15)   a.   Q: {How do many patients die in German hospitals?}
            A: [In D. sterben viele Patienten [gegen ihren Willen]$_F$ im Krankenhaus.]~
       b.   Q: {Where do many patients die in Germany?}
            A: [In D. sterben viele Patienten gegen ihren Willen [im Krankenhaus.]$_F$]~

The information-structural difference between the two readings consists in the choice of QUD and corresponding focus constituent. While the prepositional phrase *against their will* is the, non-optional, focus in (15a), it is merely an optional adjunct, and therefore non-at-issue, in (15b). We should, finally, point out that under the current interpretation of 'non-at-issue material = optional material' adjuncts, evidentials and evaluatives can change their (non-)at-issue status depending on the context, while e.g. appositions are presumably always non-at-issue.

3.7   The discourse status of non-at-issue material

Being non-at-issue with respect to the current QUD does not mean *not being an answer at all*. This may seem paradoxical at first, and yet it is a consequence of the property of non-at-issue material to denote a separate assertion. Like any other assertion, non-at-issue material must

---

[16]We are obliged to Maribel Romero (p.c.) for raising this issue.
[17]*Deutschlandfunk Nachrichten,* Nov. 02, 2015

contain a focus of its own, which serves as the answer to a subordinate question, cf. Riester and Baumann (2013: Sect. 2). This focus may be called a *non-at-issue focus*. In other words, (any) piece of non-at-issue material is only non-at-issue with respect to the *current QUD* but it is, in fact, *at-issue* with respect to a subordinate question. Putting these insights into practice, we analyze the discourse structure of the reading in (15b) as shown in Figure 10.
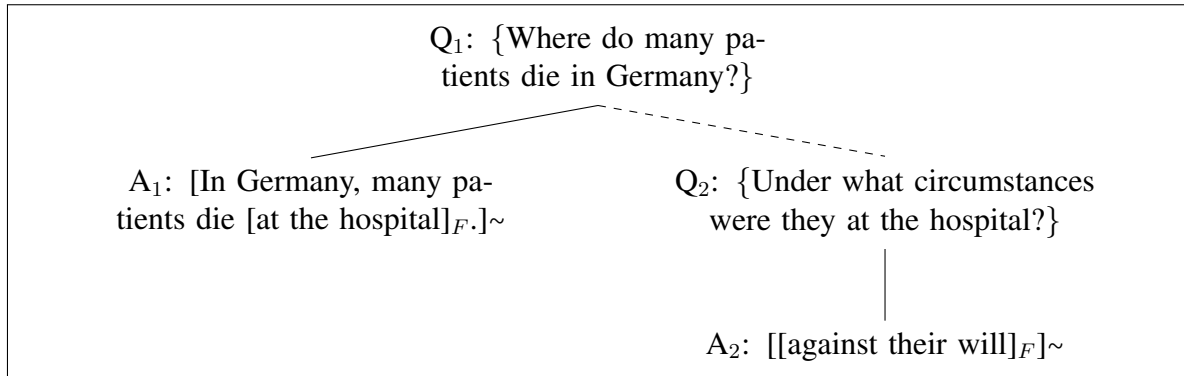
Q_1: {Where do many patients die in Germany?}

A_1: [In Germany, many patients die [at the hospital]$_F$.]~

Q_2: {Under what circumstances were they at the hospital?}

A_2: [[against their will]$_F$]~

Figure 10: Treatment of sentence-internal non-at-issue material, Example (15b)

As we said in Section 3.3, an advantage of representing discourse trees in the compact format is that the textual order is preserved when the assertions are read off from left to right. The treatment of sentence-internal non-at-issue material as a separate, subordinate assertion, as shown in Figure 10, is the only case that is disrupting the linear textual order, since the non-at-issue assertion is extracted from its original location. However, if non-at-issue material occurs in sentence-final position, no restructuring is necessary and the material can simply be treated as a separate assertion right away, compare Example (16) and its corresponding discourse tree.

(16)   a. A_5: He was working in a slate mine after the war,
          b. A_6: in Mayen.

Q_5: {What about Nahles's grandpa after the war?}

A_5: [He [was working in a slate mine]$_F$ after the war]~

Q_6: {Where was the slate mine?}
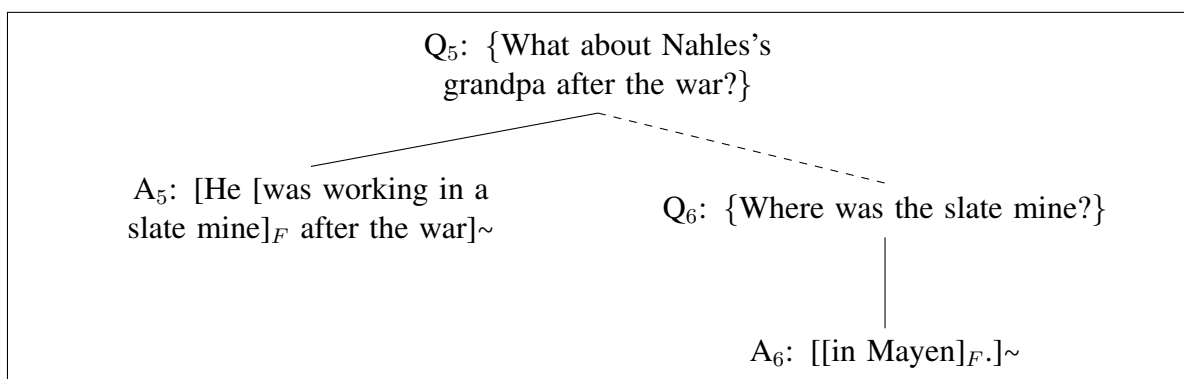
A_6: [[in Mayen]$_F$.]~

Figure 11: Treatment of sentence-final non-at-issue material

Note that the discourse constellation in Figure 11 allows the speaker to continue the discourse in (at least) two ways: she may either continue talking about grandfather and his work, which would amount to a continuation at the level of the at-issue content (A_5), or about the town of *Mayen,* at level A_6. The question whether sentence-final non-at-issue material, e.g. appositive

16

relative clauses, can behave like ordinary at-issue material is also addressed, for instance, in AnderBois et al. (2010) and Syrett and Koev (2014).

4 Conclusions and outlook

We have presented a comprehensive framework for the analysis of dialogue and text data in terms of discourse structure and information structure. We have, furthermore, discussed representation formats of discourse trees from QUD theory and SDRT, and presented ways of how to combine them in so-called QUD trees, thereby integrating QUDs into SDRT-based structures, which allow for a more flexible analysis of discourse, mainly for the reason that subordinated discourse units can answer questions which are not (entailed) subquestions to other questions higher up in the tree. We have argued that it is necessary to loosen Roberts's constraints on entailment and relevance. Furthermore, we adopt van Kuppevelt's concept of questions that are triggered by (and therefore subordinate to) answer nodes (so-called feeders). However, in order to arrive at more compact tree representations (in which all answers are terminal nodes) we give up the graphical representation of answers dominating questions, while keeping the idea of questions being anaphorically dependent of previous answers.

We argue that QUDs can replace subordinating discourse relations, and that ways can be found to represent coordinated assertions (e.g. those standing in a relation of *Narration, Parallel* or *Contrast*) as parallel siblings below a common QUD and possibly parallel subquestions. Finally, we argue that the identification of QUDs is not arbitrary but is contrained by a number of contextual constraints derived from information-structure theory: implicit QUDs must be congruent, non-novel, maximally given and attach as high as possible while maintaining givenness. The identification of parallelisms is a second way of determining QUDs in corpus data.

The benefit of creating a pool of discourse- and information-structurally labelled data is to obtain a means for the study of prosodic and morpho-syntactic variation within one language, or across several ones. Furthermore, such analyses would allow for a quantitative assessment of the coherence of texts or argumentation patterns. A further application, in the long run, is the establishment of training data for machine learning tasks like discourse parsing, co-reference resolution, text-to-speech synthesis and the like. And finally, since the method is pragmatic in nature and does not presuppose any specific morphosyntactic knowledge, it can also, in principle, be used to describe information-structural effects in lesser-studied languages.

5 Acknowledgements

A Appendix: corpus data

Stuttgart SFB732 Silver Standard Collection, cf. Eckart and Gärtner (2016)
Südwestrundfunk, SWR2 Interview der Woche

Parts of the corpus (13 interviews, 1,356 sentences, 24,114 tokens) are currently being annotated for QUDs, information structure and discourse structure according to the guidelines in Riester et al. (under revision).
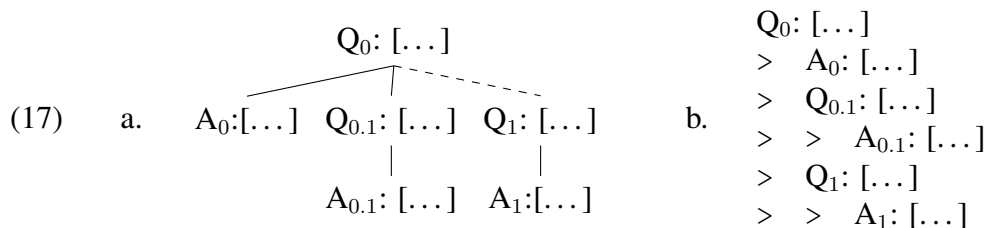
Section taken from:
Date, time: 2015/02/28, 6:30–6:40 p.m.
Interview partner: Andrea Nahles (SPD), German Federal Minister for Labour and Social Affairs

Representation:
Note that, in the following, we represent tree edges (in the compact tree format) by use of indentation marks (>), i.e. the following two representations in (17) are equivalent:

(17) a.

$$Q_0: [\ldots]$$
$$A_0:[\ldots] \quad Q_{0.1}: [\ldots] \quad Q_1: [\ldots]$$
$$A_{0.1}: [\ldots] \quad A_1:[\ldots]$$

b.

$Q_0: [\ldots]$
$>\quad A_0: [\ldots]$
$>\quad Q_{0.1}: [\ldots]$
$>\quad >\quad A_{0.1}: [\ldots]$
$>\quad Q_1: [\ldots]$
$>\quad >\quad A_1: [\ldots]$

Notes to the interview:

(a) We assume here that the pronoun *da 'there'* refers anaphorically to *Arbeitsstättenverordnung 'workplace regulation bill'* and is, therefore, topical.

(b) We assume that, in the respective contexts, the phrases *nach dem Krieg 'after the war'* and *in den letzten Jahren 'in his last years'* are salient and, therefore, backgrounded althought they are not literally given in the previous discourse.

(c) The sentence is ungrammatical in the German original.

Summary of the context:

The previous discussion between the journalist and the politician is about a conflict between the Social-Democratic Minister and the more business-friendly Christian Democratic Party (CDU) about the Minister's newly introduced statutory minimum wages. The CDU is complaining that the new law means a lot of bureaucracy for employers. Nahles is defending herself against the accusations. End of the minium-wage topic.

A.1 Analysis of original German data

Journalist:

_(Notes) Q_0: {*Bei welchem anderen Projekt wird Nahles von den Arbeitgebern Bürokratisierung vorgeworfen?*}

> A_0: [Ein [anderes]$_{CT}$ Projekt, bei dem die Arbeitgeber Sie mit dem Bürokratievorwurf überziehen, ist [die Arbeitsstättenverordnung]$_F$.]~

> Q_1: {*Was soll verordnet/vorgeschrieben werden?*}

(a) > > $A_{1'}$: [[Da]$_T$ sollen [Fenster in Teeküchen]$_F$ vorgeschrieben werden,]~

> > $A_{1''}$: oder auch [[die Helligkeit am Heimarbeitsplatz]$_F$]~.

> Q_2: {*Wie ist der Status der Verordnung?*}

> > A_2: <mark>Jetzt heißt es,</mark> [[das Kanzleramt hat]$_F$ [diese Verordnung]$_T$ [gestoppt]$_F$]~.

> > $Q_{2.1}$: Können Sie das bestätigen?

> Q_3: Sind Sie da gescheitert?

Andrea Nahles:

> Q_4: Also zunächst muss man vielleicht mal sagen, worum es da eigentlich geht. {= *Worum geht es bei der Verordnung eigentlich?*}

> > $Q_{4.1}$: {*Wie war der Arbeitsschutz früher und wie ist er heutzutage?*}

> > > $Q_{4.1.1}$: {*Wie war der Arbeitsschutz früher?*}

> > > > $Q_{4.1.1.1}$: {*Welche persönlichen Beispiele hat Nahles?*}

> > > > > $A_{4.1.1.1}$: [[Mein]$_T$ [Opa hat eine Staublunge gehabt]$_F$,]~

(b) > > > > > Q_5: {*Was war mit Nahles' Opa nach dem Krieg?*}

> > > > > > A_5: [[der]$_T$ war [im Schieferbergwerk]$_F$ nach dem Krieg]~

> > > > > > Q_6: {*Wo war das Schieferbergwerk?*}

> > > > > > > A_6: [[in Mayen]$_F$]~

(b) > > > > > Q_7: {*Wie ging es dem Opa in den letzten Jahren?*}

> > > > > > A_7: und <mark>ich kann nur sagen,</mark> [das war [schlimm]$_F$ in den letzten Jahren, wie es [dem]$_T$ ging]~.

> > > > > > $Q_{7.1}$: {*Wie schlimm ging es ihm?*}

> > > > > > > $A_{7.1}$: [[Der]$_T$ [ist richtig erstickt]$_F$.]~

> > > $Q_{4.1.2}$: {*Wie ist der Arbeitsschutz heutzutage?*}

> > > > $Q_{4.1.2.1}$: {*Die Anzahl wovon ist durch die Arbeitsschutzmaßnahmen massiv zurückgegangen?*}

(c) > > > > > $A_{4.1.2.1'}$: Und [es ist [durch viele Maßnahmen in den letzten 40 Jahren, die man unter dem Stichwort "Arbeitsschutz und Arbeitsstättenverordnung", <mark>so heißt das Ding nun mal,</mark> firmiert]$_T$, sind <mark>Gott sei Dank</mark> die Anzahl [der Toten]$_F$,]~

> > > > > $A_{4.1.2.1''}$: [[der Kranken,]$_F$]~

> > > > > $A_{4.1.2.1'''}$: [[der Unfälle]$_F$ massiv zurückgegangen]~.

> A_4: Und deswegen [ist [das]$_T$ schon [etwas sehr Wertvolles]$_F$, [diese Arbeitsschutzpolitik und auch diese Arbeitsstättenverordnung]$_T$]~.

> Q_8: {*Was ist Nahles' Reaktion zu Vorbehalten gegenüber der Verordnung?*}

> > $Q_{8.1}$: {*Was tut sie, wenn sich jemand über Kleinigkeiten aufregt?*}

> > > $A_{8.1}$: [Wenn man [da]$_T$ jetzt [sich aufregt wegen einem abschließbaren Spind]$_{CT}$, <mark>na Gott,</mark> dann wäre ich [der Letzte der sagt, da kann man nicht über einzelne Punkte reden]$_F$.]~

> > $Q_{8.2}$: {*Was tut sie, wenn jemand die ganze Verordnung in Frage stellt?*}

> > > $A_{8.2}$: Aber [wenn man [das]$_T$ [grundsätzlich in Frage stellt]$_{CT}$, dann werde ich [allerdings doch ernst]$_F$.]~

A.2 English translation

Journalist:

Q_0: {*For which other projects is Nahles accused of bureaucratisation by the employers?*}

> A_0: [[Another]$_{CT}$ project for which employers are accusing you of bureaucratisation is [the workplace regulation bill]$_F$.]~

> Q_1: {*What will be prescribed?*}

> > $A_{1'}$: [[The bill]$_T$ will prescribe [having windows in staff kitchens]$_F$]~

> > $A_{1''}$: and even [[the brightness of home workplaces]$_F$]~.

> Q_2: {*What is the current status of the bill?*}

> > A_2: Now they are saying that [[the Chancellery has stopped]$_F$ [the bill]$_T$]~.

> > $Q_{2.1}$: Can you confirm this?

> Q_3: Have you failed there?

Andrea Nahles:

> Q_4: All right, first I need to say what this is all about. {= *What about this bill?*}

> > $Q_{4.1}$: {*What about occupational safety in earlier times and nowadays?*}

> > > $Q_{4.1.1}$: {*What about occupational safety in earlier times?*}

> > > > $Q_{4.1.1.1}$: {*What personal examples does Nahles have?*}

> > > > > $A_{4.1.1.1}$: [[My]$_T$ [grandpa suffered from silicosis]$_F$.]~

(b) > > > > > Q_5: {*What about Nahles's grandpa after the war?*}

> > > > > > A_5: [[He]$_T$ [was working in a slate mine]$_F$ after the war,]~

> > > > > > Q_6: {*Where was the slate mine?*}

> > > > > > > A_6: [[in Mayen]$_F$]~

(b) > > > > > Q_7: {*How was grandpa's health during his last years?*}

> > > > > > A_7: and all I can say is that [his$_T$ condition was [extremely bad]$_F$ during his last years]~.

> > > > > > $Q_{7.1}$: {*How bad was his condition?*}

> > > > > > > $A_{7.1}$: [[He]$_T$ [literally suffocated]$_F$.]~

> > > $Q_{4.1.2}$: {*What about occupational safety nowadays?*}

> > > > $Q_{4.1.2.1}$: The number of what has massively decreased due to occupational
> > > > safety measures?

> > > > > $A_{4.1.2.1'}$: and [it is [due to the numerous measures in the past 40 years called occupational safety measures and workplace regulations]$_T$ – that's indeed how they are called – that, Thank God, the numbers of [casualties,]$_F$]~

> > > > > $A_{4.1.2.1''}$: [[injuries]$_F$]~

> > > > > $A_{4.1.2.1'''}$: and [[accidents]$_F$ has decreased massively]$_F$]~.

> A_4: And, therefore, [[these safety measures and the workplace regulation bill]$_T$ are [something very valuable]$_F$]~.

> Q_8: {*What is Nahles's reaction to different attitudes regarding the bill?*}

> > $Q_{8.1}$: {*What does she do if someone is upset about minor issues?*}

> > > $A_{8.1}$: [If someone is [upset because of some lockers]$_{CT}$, oh well, then I would be [the last person unwilling to discuss a compromise]$_F$.]~

> > $Q_{8.2}$: {*What does she do if someone is fundamentally opposed to the bill?*}

> > > $A_{8.2}$: But [if someone is [fundamentally opposed]$_{CT}$ to [it]$_T$, then I [won't be joking anymore]$_F$.]~

A.3   QUD tree

The tree in Figure 12 makes the structure of the textual analyses (of both German and English) transparent. Questions nodes are rectangular, answer nodes are oval. Entailment between question nodes is represented by solid edges, while non-entailment is represented by dashed edges. Questions attached via non-entailed edges must be anaphorically dependent on a previous answer to the superquestion (the feeder).
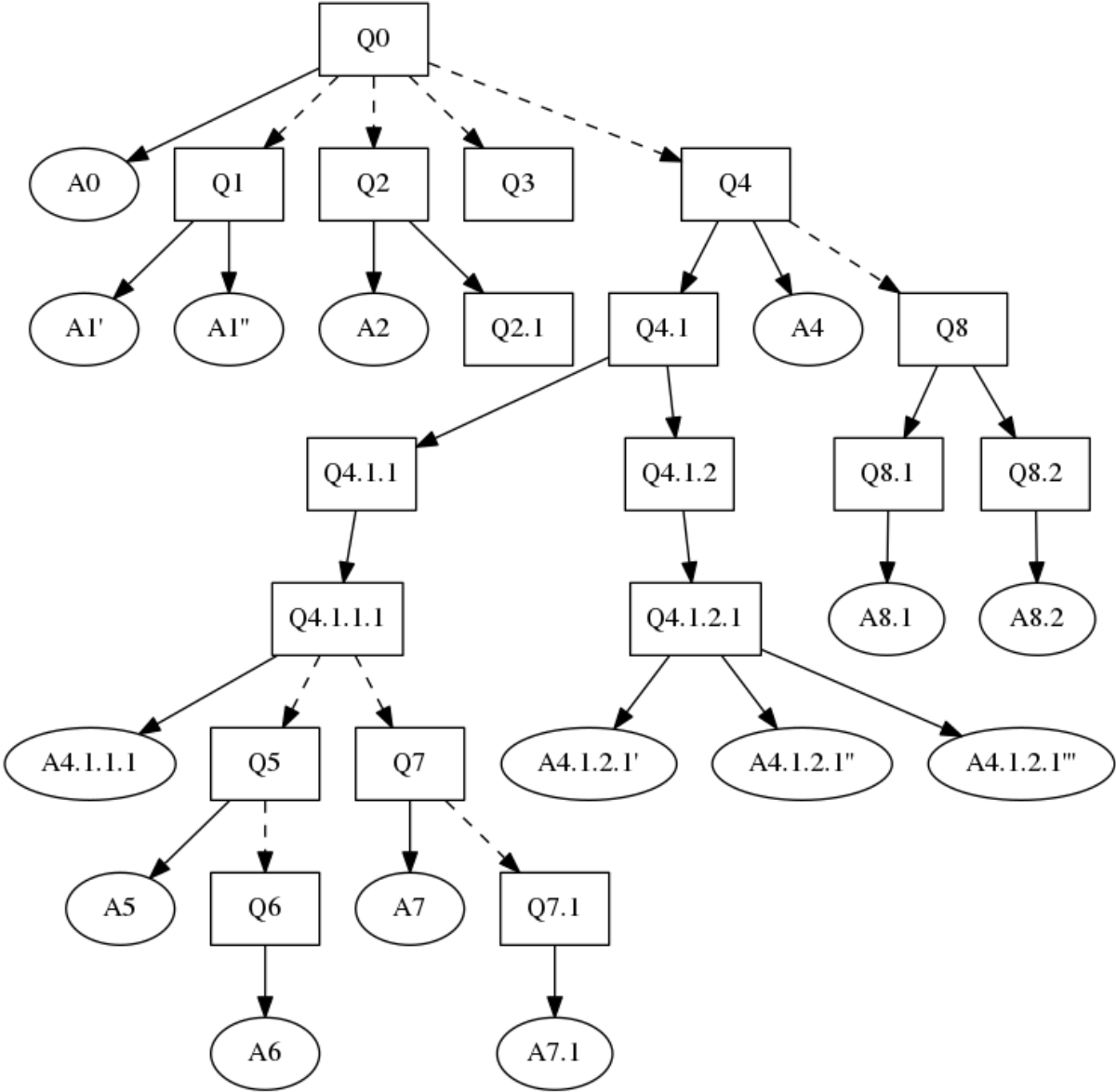

Figure 12: QUD tree for the interview section

References

Stergos Afantenos and Nicholas Asher. Testing SDRT's Right Frontier. In *Proceedings of the 23rd International Conference on Computational Linguistics (COLING)*, pages 1–9, Beijing, 2010.

Scott AnderBois, Adrian Brasoveanu, and Robert Henderson. Crossing the appositive / at-issue meaning boundary. In Nan Li and David Lutz, editors, *Semantics and Linguistic Theory 20*, pages 328–346, Vancouver, 2010.

Nicholas Asher. Troubles on the right frontier. In Anton Benz and Peter Kühnlein, editors, *Constraints in Discourse*, volume 172 of *Pragmatics & Beyond New Series*, pages 29–52. John Benjamins, Amsterdam, 2008.

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

Stefan Baumann and Arndt Riester. Referential and Lexical Givenness: Semantic, prosodic and cognitive aspects. In Gorka Elordieta and Pilar Prieto, editors, *Prosody and Meaning*, volume 25 of *Interface Explorations*, pages 119–162. Mouton de Gruyter, Berlin, 2012.

David Beaver. Presupposition accommodation: A plea for common sense. In Larry Moss, Jonathan Ginzburg, and Maarten De Rijke, editors, *Logic, Language and Computation*, volume 2, pages 21–44. CSLI, Stanford, 1999.

David Beaver and Brady Clark. *Sense and Sensitivity. How Focus Determines Meaning*. Wiley-Blackwell, Oxford, 2008.

David Beaver and Henk Zeevat. Accommodation. In G. Ramchand and C. Reiss, editors, *The Oxford Handbook of Linguistic Interfaces*, pages 503–538. Oxford University Press, 2007.

Daniel Büring. On D-Trees, Beans, and B-Accents. *Linguistics & Philosophy*, 26(5):511–545, 2003.

Daniel Büring. What's New (and What's Given) in the Theory of Focus? In *Proceedings of the 34th Annual Meeting of the Berkeley Linguistics Society*, pages 403–424, 2008.

Daniel Büring. *Intonation and Meaning*. Oxford University Press, 2016.

Lauri Carlson. *Dialogue Games: An Approach to Discourse Analysis*. Reidel, Dordrecht, 1983.

Kerstin Eckart and Markus Gärtner. Creating silver standard annotations for a corpus of non-standard data. In Stefanie Dipper, Friedrich Neubarth, and Heike Zinsmeister, editors, *Proceedings of the 13th Conference on Natural Language Processing (KONVENS)*, volume 16 of *Bochumer Linguistische Arbeitsberichte*, pages 90–96, Bochum, Germany, 2016.

Martina Faller. *Semantics and Pragmatics of Evidentials in Cuzco Quechua*. PhD thesis, Stanford University, 2002.

Jonathan Ginzburg. An update semantics for dialogue. In Harry Bunt, editor, *Proceedings of the 1st International Workshop on Computational Semantics (IWCS)*, Tilburg, NL, 1994.

Jonathan Ginzburg. Dynamics and the semantics of dialogue. In Jerry Seligman and Dag Westerståhl, editors, *Logic, Language, and Computation*, pages 221–237. CSLI, Stanford, CA, USA, 1996.

Jeroen Groenendijk and Martin Stokhof. *Studies in the Semantics of Questions and the Pragmatics of Answers*. PhD thesis, Universiteit van Amsterdam, 1984.

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

Carlos Gussenhoven. On the limits of focus projection in English. In Peter Bosch and Rob van der Sandt, editors, *Focus. Linguistic, Cognitive and Computational Perspectives*, pages 43–55. Cambridge University Press, 1999.

Eva Hajičová, Barbara H. Partee, and Petr Sgall. *Topic-Focus Articulation, Tripartite Structures, and Semantic Content*. Springer, Dordrecht, 1998.

Charles Hamblin. Questions in Montague English. *Foundation of Language*, 10:41–53, 1973.

Jerry Hobbs. On the coherence and structure of discourse. Technical Report CSLI-85-37, Stanford University, 1985.

Julie Hunter and Márta Abrusán. Rhetorical relations and QUDs. In *New Frontiers in Artificial Intelligence: JSAI-isAI Workshops LENLS, JURISIN, KCSD, LLLL Revised Selected Papers*. Springer, to appear.

Sabine Iatridou. *Topics in Conditionals*. PhD thesis, Massachusetts Institute of Technology, Cambridge, MA, 1991.

Wolfgang Klein and Christiane von Stutterheim. Quaestio und referentielle Bewegung in Erzählungen. *Linguistische Berichte*, 109:163–183, 1987.

Manfred Krifka. A compositional semantics for multiple focus constructions. In Joachim Jacobs, editor, *Informationsstruktur und Grammatik*, pages 17–53. Westdeutscher Verlag, Opladen, 1992.

Manfred Krifka. Basic notions of information structure. *Interdisciplinary Studies on Information Structure*, 6, 2007. Potsdam: Universitätsverlag.

Knud Lambrecht. *Information Structure and Sentence Form*. Cambridge University Press, 1994.

Alex Lascarides and Nicholas Asher. Segmented Discourse Representation Theory: Dynamic semantics with discourse structure. In Harry Bunt and Reinhard Muskens, editors, *Computing Meaning. Volume 3*, pages 87–124. Springer, Dordrecht, NL, 2008.

William Mann and Sandra Thompson. Rhetorical Structure Theory: Towards a functional theory of text organization. *Text*, 8:243–281, 1988.

Sarah Murray. *Evidentiality and the Structure of Speech Acts*. PhD thesis, Rutgers University, 2010.

Edgar Onea. *Potential Questions at the Semantics-Pragmatics Interface*. Brill, Leiden, 2016.

Livia Polanyi. A formal model of the structure of discourse. *Journal of Pragmatics*, 12(5-6): 601–638, 1988.

Christopher Potts. *The Logic of Conventional Implicatures*. Oxford University Press, 2005.

Tanya Reinhart. Pragmatics and linguistics: An analysis of sentence topics in pragmatics and philosophy I. *Philosophica*, 27(1):53–94, 1981.

Uwe Reyle and Arndt Riester. Joint information structure and discourse structure analysis in an Underspecified DRT framework. In Julie Hunter, Mandy Simons, and Matthew Stone, editors, *Proceedings of the 20th Workshop on the Semantics and Pragmatics of Dialogue (JerSem)*, pages 15–24, New Brunswick, NJ, USA, 2016.

Arndt Riester and Stefan Baumann. Focus triggers and focus types from a corpus perspective. *Dialogue & Discourse*, 4(2):215–248, 2013.

Arndt Riester and Stefan Baumann. *The RefLex Scheme – Annotation Guidelines*, volume 14. SinSpeC. Working Papers of the SFB 732, University of Stuttgart, 2017.

Arndt Riester and Jörn Piontek. Anarchy in the NP: When new nouns get deaccented and given nouns don't. *Lingua*, 165(B):230–253, 2015.

Arndt Riester, Lisa Brunetti, and Kordula De Kuthy. Annotation guidelines for Questions under Discussion and information structure. under revision. manuscript.

Craige Roberts. Focus, the flow of information, and universal grammar. In Peter Culicover and Louise McNally, editors, *The Limits of Syntax*, volume 29 of *Syntax and Semantics*, pages 109–160. Academic Press, San Diego, 1998.

Craige Roberts. Information structure: Towards an integrated formal theory of pragmatics. *Semantics and Pragmatics*, 5(6):1–69, 2012. 1996 version in OSU Working Papers in Linguistics 49.

Mats Rooth. A theory of focus interpretation. *Natural Language Semantics*, 1(1):75–116, 1992.

Roger Schwarzschild. GIVENness, AvoidF, and other constraints on the placement of accent. *Natural Language Semantics*, 7(2):141–177, 1999.

Mandy Simons, Judith Tonhauser, David Beaver, and Craige Roberts. What projects and why. In *Proceedings of SALT 20*, pages 309–327, Vancouver, 2010.

Kristen Syrett and Todor Koev. Experimental evidence for the truth conditional contribution and shifting information status of appositives. *Journal of Semantics*, 0:1–53, 2014. doi: 10.1093/jos/ffu007.

Maite Taboada and William Mann. Rhetorical Structure Theory: Looking back and moving ahead. *Discourse Studies*, 8:423–459, 2006.

Judith Tonhauser. Diagnosing (not-)at-issue content. In *Procceddings of Semantics of Underrepresented Languages of the Americas (SULA)*, pages 239–254, Amherst, 2012. University of Massachusetts.

Judith Tonhauser, David Beaver, Craige Roberts, and Mandy Simons. Toward a taxonomy of projective content. *Language*, 89(1):66–109, 2013.

Jan van Kuppevelt. Discourse structure, topicality and questioning. *Journal of Linguistics*, 31: 109–149, 1995.

Leah Velleman and David Beaver. Question-based models of information structure. In Caroline Féry and Shinichiro Ishihara, editors, *The Oxford Handbook of Information Structure*, pages 86–107. Oxford University Press, 2016.

Christiane von Stutterheim and Wolfgang Klein. Referential movement in descriptive and narrative discourse. In Rainer Dietrich and Carl Graumann, editors, *Language Processing in Social Context*, pages 39–76. North Holland, Amsterdam, 1989.

Bonnie Webber. Structure and ostension in the interpretation of discourse deixis. *Natural Language and Cognitive Processes*, 6(2):107–135, 1991.