

Annotating Information Structure in Italian: Characteristics and Cross-Linguistics Applicability of a QUD-Based Approach

Anonymous ACL submission

Abstract

We present a discourse annotation study, in which an annotation method based on *Questions under Discussion (QuD)* is applied to Italian data. The results of our inter-annotator agreement analysis show that the QUD-based approach, originally spelled out for English and German, can successfully be transferred cross-linguistically, supporting good agreement for the annotation of central information structure notions such as *focus* and *non-at-issueness*. Our annotation and inter-annotator agreement study on Italian authentic data confirms the cross-linguistic applicability of the QuD-based approach.

1 Introduction

In this paper, we present a discourse annotation study of Italian data, which uses the annotation scheme and discourse-analytic method, the *QUD-tree* framework, developed in [Reyle and Riester \(2016\)](#), [Riester \(to appear\)](#) and [Riester et al. \(2018\)](#). Its purpose is the cross-linguistic analysis of *information structure* and *discourse structure* of textual data. On the theoretical side, the QUD framework has been applied to a number of different languages, such as German, English and French in [\(Riester et al., 2018\)](#), and various Austronesian languages as discussed in [Riester and Shiohara \(2018\)](#) and [Latrouite and Riester \(2018\)](#). On the applied side, [De Kuthy et al. \(2018\)](#) showed that the QUD based method supports the successful annotation of discourse structure and information structure in German and English spoken language data. Here we want to broaden the crosslinguistic scope of the QUD framework and apply it to another Romance language, Italian. We will explore both the QUD annotation and the information structure annotation including all information structure labels that are part of the annotation scheme proposed in [Riester et al. \(2018\)](#), such

as *focus*, *background*, *contrastive topic*, *nai* and *topic*. Topic is regarded as a notoriously difficult label in agreement studies (cf. [Cook and Bildhauer \(2013\)](#); [Ritz et al. \(2008\)](#)). While the results of our study show that the question-based annotation method supports the successful annotation of discourse structure and of information structure, in particular focus, we will also discuss, using the example of topic, some shortcomings of the QUD based annotation method.

2 The QUD framework

The QUD framework introduced in [Riester et al. \(2018\)](#) presents an explicit method for the reconstruction of QUDs which are usually only discussed as an abstract theoretical term. The center of the QUD framework is a compact representation format for QUD trees, in which the textual assertions (A) represent the terminal nodes of a discourse tree (preserving the linear order of the text from left to right) while (implicit or explicit) QUDs (Q) form the non-terminal nodes. An abstract QUD tree is shown in Figure 1.

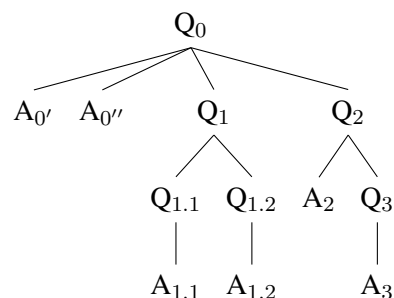


Figure 1: QUD tree

The QUD-tree framework as spelled out in [Riester et al. \(2018\)](#) can be applied to any kind of written or spoken discourse or conversation. It is not language-specific and can, in principle, be

used in order to investigate data from any language. While the exact analysis procedure is described at great length within the guidelines document (Riester et al., 2018), we just briefly introduce some basic principles here.

2.1 Segmentation

Raw texts are segmented into atomic assertions. Apart from orthographic sentence boundaries, segmentation also applies at (1) (information-structurally relevant) coordinations and (2) before (optional) syntactic adjuncts. (Obligatory) sentential arguments (3) are not split off.

- (1) A₄: Ho appena terminato un romanzo
I-have just finished a novel
 A₄: e sono già al lavoro su un nuovo progetto, ...
and am already at work on a new project
- (2) A_{7''}: Di recente ho ripreso a leggere i romanzi di formazione,
of recently I-have re-started to read the novels of coming-of-age
 A_{7'''}: senza mai tralasciare la narrativa contemporanea e i romance.
without ever neglecting the contemporay fiction and the romance'
- (3) A_{25'}: [[Alek]_T [è frutto della mia fantasia]_F],
Alex is result of-the my imagination
 A_{25''}: [[nasce in relazione a Dave]_F],
he-is-born in relation to Dave
 A₂₆: [[ho voluto che]_{NAI} fosse ["forte" ma non "invincibile"]_F].
I-have wanted that he-was strong but not invincible

2.2 QUD principles

The actual identification of a QUD for each assertion is guided by a number of explicit principles adapted from the formal literature on information structure (Rooth, 1992; Schwarzschild, 1999; Büring, 2008, 2016), cf. Riester et al. (2018):

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

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

MAXIMIZE-Q-ANAPHORICITY: Implicit QUDs should contain as much given (or salient) material as possible.

Example (4) shows that from these principles we can derive QUD Q_{32} for assertion A_{32} in the

context of A_{31} , whereas any of the questions in (5), used in place of Q_{32} , would violate at least one of the QUD constraints in the same context.

- (4) A₃₁: Anche tra i bilingui precoci che
even among the early bilinguals who
 parlano due lingue quasi mai le due
speak two languages almost never the two
 lingue sono del tutto equivalenti,
languages are completely equivalent
 Q₃₂: {What about the two languages instead?}
 > A₃₂: and [[normalmente]_{NAI} [[ogni lingua]_T
and usually each language
 [si sviluppa in un contesto specifico]_F]_~
itself develops in a context specific
- (5) a. {What about speaking two languages?}
 (#Q-A-CONGRUENCE)
 b. {What about the specific context?}
 (#Q-GIVENNESS)
 c. {What happens next?}
 (#MAXIMIZE-Q-ANAPHORICITY)

Two or more assertions are defined as parallel if and only if they share some semantically identical content and represent partial answers to the same QUD, see Example (6), where the semantically shared content is *Alek* (omitted in the second assertion).

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

- (6) Q₂₅: {What about the connection with reality in Alek?}
 > A_{25'}: [[Alek]_T [è frutto della mia fantasia]_F],
Alex is the result of-the my imagination
 > A_{25''}: [[nasce in relazione a Dave]_F],
is-born in relation to Dave

The resulting tree structure is shown in Figure 2.

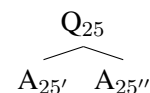


Figure 2: Two coordinated (parallel) assertions.

3 QUDs and information structure

The basis of our annotation approach is an alternative-based definition of information structural categories, in line with e.g. Rooth (1992), Büring (2003), Beaver and Clark (2008), Krifka

(2008) or Roberts (2012). The Table in 1 shows the definitions for the information structure categories as introduced in Riester et al. (2018). These are the basis for the labels used in our annotation study.

| Category (Label) | Definition |
|-----------------------------|--|
| Focus domain (\sim) | Part of an assertion that has the same background as the current QUD and that contains a focus |
| Focus (F) | Constituent that answers the current QUD |
| Background (BG) | Material mentioned in the current QUD |
| Contrastive topic (CT) | Material backgrounded w.r.t. the current QUD and focal w.r.t. a super-question |
| Topic (T) | Distinguished discourse referent identifying what the sentence is about |
| Non-at-issue material (NAI) | Optional material w.r.t. the current QUD |

Table 1: Information structure: Label inventory

(7) Q_7 : {Cosa ti piace leggere?}
what you like to-read
 > A_7 : [[Di recente]_{NAI} [[ho ripreso a
of recently I-have re-started to
leggere]_{BG} [i romanzi di
read the novels of
*formazione]_F] \sim
*coming-of-age**

(7) is an example demonstrating the assignment of information-structure labels in the context of a QUD (in curly brackets). Note that the indentation (>) of A_{15} in the textual representation marks subordination in the discourse tree, as shown in Figure 2.

4 Evaluation: Discourse structure

In the present annotation study based on the above described QUD framework, our goal is to show that the discourse annotation in terms of QUDs can be applied reliably to naturally occurring data - in this particular case, Italian data. We conducted an

empirical study, in which annotators followed the QUD guidelines described in Riester et al. (2018) to annotate two Italian blog interviews.

For the QUD-based annotation we use the tool *TreeAnno* introduced by De Kuthy et al. (2018), which enables the analyst to semi-automatically segment texts, systematically enhance them with implicit *Questions under Discussion (QUDs)*, and transform the data into a discourse tree called *QUD tree*, as described in De Kuthy et al. (2018).

4.1 Evaluation setup

Two trained annotators (and also native speakers of Italian) analyzed and annotated two short Italian blog interviews downloaded from the internet¹. The first blog interview consists of 95 text segments, the second one of 113 segments. The QUD discourse tree for Blog 1 resulting from the first annotator is shown in Figure 3, the other three discourse trees are uploaded as separate files as part of the Appendix.

4.2 Method and results

For the comparison of the two annotated documents, we follow the method described in De Kuthy et al. (2018). The basic idea is that for the comparison of two QUD annotations one needs to calculate an inter-annotator agreement score that takes into account, for every segment and every possible span of segments, whether a QUD is present or not. In order to compute a κ statistics (Cohen, 1960) based on our QUD annotations, De Kuthy et al. (2018) propose to follow the method described in Marcu et al. (1999), which was developed for measuring agreement in the labelling of rhetorical structure categories in texts. The method is based on the idea of mapping the hierarchical structure of a discourse tree onto a matrix or chart filled with categorical values. In our case, the values are whether there exists a (Q)uestion spanning the respective segments – start to end – or (n)ot).

A κ statistics can then be computed between two charts that represent two different QUD annotations for the same text; more precisely between the two resulting sets of cells in the upper half of each chart. For our two annotated documents we calculated κ values for the annotation charts derived from our QUD annotations, based on the

¹Blog 1 URL: <http://purl.org/info-struct/Italian-blog-1>, Blog 2 URL: <http://purl.org/info-struct/Italian-blog-2>

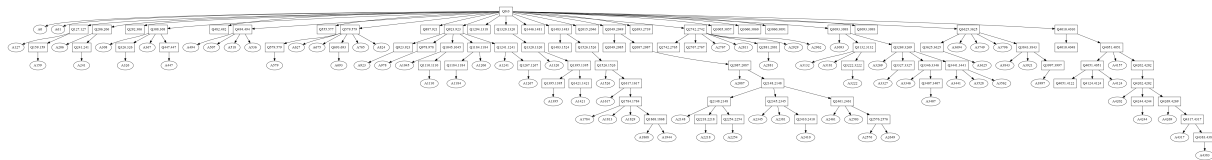


Figure 3: A QUD tree analyses for Italian Blog 1

above described method. For the text *Italian Blog 1*, consisting of 95 segments, we calculated the κ statistics based on 4,256 items, for *Italian Blog 2* with 113 segments based on 5,350 items. The results are shown in Table 2

| Text | Segments | Cells | κ |
|----------------|----------|-------|----------|
| Italian Blog 1 | 95 | 4,256 | .61 |
| Italian Blog 2 | 113 | 6,187 | .51 |

Table 2: Kappa values for QUD-annotated Italian dialogues

The values show moderate agreement between the annotator pairs. For Blog 1, the κ value is at .61, which is substantially higher than what (De Kuthy et al., 2018) report for the QUD annotations of their German and English texts: their κ values are around .5. For our Blog 2, the κ value is at .51, which is thus very similar to the scores reported in (De Kuthy et al., 2018) for texts of similar length. Our two annotated Italian texts are relatively short, only around 100 sentences each, so it is perhaps too early to interpret the results, in particular since this is a rather complex task. However, since the results are comparable to those reported in (De Kuthy et al., 2018), we take this as a further proof that the QUD-based annotation of discourse can successfully be applied cross-linguistically.

5 Evaluation: Information structure

The second major issue we are interested in is to evaluate the reliability of information-structure annotation based on the previous identification of QUDs.

5.1 Evaluation setup

For the evaluation of the information structural annotation, the same two Italian blog texts were annotated by the same two trained annotators, who still followed the guidelines of Riester et al. (2018). We aimed at annotating all five categories that are mentioned in Riester et al. (2018): focus (F), background (BG), non-at-issue material

| | | |
|----|--|------------|
| 6 | Da dove vieni? | |
| 7 | Vengo da Palermo, una città che profuma di mare e cultura, | BG, F |
| 8 | ma ciò che sono viene anche da tanti posti in cui ho lasciato una parte di me, | F, BG, F |
| 9 | come Milano, Londra e Nottingham. | F |
| 10 | Mai senza? | |
| 11 | Mio figlio, il Caffè, il mio quaderno, un po' di follia. | F |
| 12 | Cosa ti piace leggere? | |
| 13 | Sono una divortrice di libri quasi onnivora. | F, BG, F |
| 14 | Escludendo gli horror, leggo un po' di tutto. | F, BG, F |
| 15 | Amo i classici, | F |
| 16 | tengo sul comodino <i>Orgoglio e Pregiudizio</i> e <i>I Sonetti di Shakespeare</i> . | F |
| 17 | Di recente ho ripreso a leggere i romanzi di formazione, | NAI, BG, F |
| 18 | senza mai tralasciare la narrativa contemporanea e i romance. | F |

Figure 4: Annotation in WebAnno

(NAI), contrastive topic (CT) and topic (T). Focus domain labels (\sim) were not annotated, since each text segment (assertion) already corresponds to one focus domain. The annotators based their annotations on the previously performed QUD analysis in the TreeAnno tool. As an annotation tool for the token-based information-structure annotation, WebAnno (Yimam et al., 2013) was chosen. Figure 4 shows a screenshot of the information-structure annotation of the beginning of *Blog 1*.

5.2 Method and results

As agreement measure for the evaluation of the information structure annotation, we calculated κ values on the annotated texts based on tokens, following previous work (Ritz et al., 2008; Calhoun et al., 2010; De Kuthy et al., 2018). In addition to the specifications in Riester et al. (2018), in particular the QUD-to-information-structure mapping from Table 1, we defined a number of heuristic (but potentially debatable) rules in order to prevent disagreement due to theoretically unclear issues, such as:

- Discourse connectors (*but, and, although, because, therefore* etc.) at the beginning of discourse segments are not annotated.

- Punctuation: Quotation marks around an expression, commas within and at the right edge of an expression are part of the markable. Periods, colons, semicolons, exclamation marks are not.

Results are shown in Table 3, divided into scores for all labels taken together, and individual scores for each of the four labels.

| Text | Label | Tokens | κ |
|----------------|-------|--------|----------|
| Italian Blog 1 | all | 847 | .70 |
| | F | | .72 |
| | BG | | .21 |
| | CT | | .85 |
| | NAI | | .53 |
| | T | | .45 |
| Italian Blog 2 | all | 1243 | .58 |
| | F | | .51 |
| | BG | | .1 |
| | CT | | .1 |
| | NAI | | .62 |
| | T | | .35 |

Table 3: Kappa for information structure annotation

The results are rather heterogeneous in both texts but overall they show that the QUD-based method does contribute to a successful annotation of information structure in Italian for a range of labels. For the first text Blog 1, the overall agreement score for all annotated categories taken together is at .7, which shows substantial agreement, the score for focus annotation alone being at .72. The agreement scores for the second blog are overall lower, but with .58 for the overall agreement and .51 with agreement for focus they are still at a relatively high level and still comparable to the scores that (De Kuthy et al., 2018) report for the annotation of German and English data (which are at around .65). The category NAI, the classification of non-at-issue material, also received reasonable agreement scores at .53 in Blog 1 and .62 for Blog 2. The agreement scores for the other three categories, BG and CT, differ a lot between the two texts. In Blog 1, the score for contrastive topic is very high with .85, in Blog 2 the score .1 shows that there was hardly any agreement between the two annotators. This might be due to the fact that there were only very few cases for which the label CT was used. In Blog 1, the label CT was used for 9 and 12 tokens in the two annotations, in Blog 2 it was assigned to 13 and 14 tokens (out of

1243 tokens). The case is similar with respect to background: in the two annotated documents, the label BG was only assigned for around 40 tokens in Blog 1 and 30 tokens in Blog 2. This means that, if the annotators disagreed in only one token when assigning the label CT or BG, this had a much greater impact on the agreement scores for these labels than in the cases of disagreement for assigning focus labels. The category topic (T) received relatively low agreement scores at .45 and .35, but still at a level which other studies report for categories like focus (cf. Ritz et al. 2008 report a κ of .44 for focus). In the following section we will qualitatively evaluate why the annotation scheme seems to better support the successful annotation of a category like focus, whereas there seems to be much more disagreement when annotating topic.

6 Qualitative Evaluation: The Case of Topics

In the question-based definitions of our information structure labels, the focus corresponds to those parts of an assertion that answer the current QUD. Especially in case of overt questions, but with implicit QUDs, the annotators agree on focus.

The definition of topic in the QUD framework, however, is the only one that does not take the current QUD into account. As remarked by Riester et al. (2018), while potentially all referential expressions inside the background could be labelled as topic, one might argue that not all referential expressions inside the background are actually aboutness topics. But unfortunately, the QUD method is not meant to single out the best topic candidate. And Riester et al. (2018) do not provide any rules that help to distinguish between better and worse topic candidates. The only cue that is given through the current QUD is that all focal expressions are excluded as topic candidates.

A typical topic expression in Italian would be a clitic left or right dislocated phrase (see *quel libro* below), but no dislocation was present in our data, probably due to the fact that a blog interview is less interactive than an spoken conversation, and these construction are typically used in interaction.

- (8) a. *Quel libro, l'ho dato a Giorgio.*
that book it I-have given to Giorgio
 b. *L'ho dato a Giorgio, quel libro.*
it I-have given to Giorgio that book

500 Clitic personal pronouns, such as *le* in A₂ in (9),
501 are also typical candidates for (continuing) top-
502 ics.

- 503 (9) A₁: Abbiamo fatto quattro chiacchiere con Maria Ver-
504 diana Rigoglioso per parlare di *Senza Etichette*, il
505 romanzo che ha pubblicato con Libromania.
506 'We had a chat with Maria Verdiana Rigoglioso to
507 talk about *Senza Etichette*, the novel she published
508 with Libromania.'
509 Q₂: {What did you do with her exactly?}
510 > A₂: [[Le]_T [abbiamo fatto un po' di domande]_F]
511 *to-her we-have made a little of questions*
512 > Q₃: {What for? }
513 > A₃: [per [conoscere retroscena e curiosità del
514 *to know ins-and-outs and trivia of-the*
515 *romanzo]_F].
516 *romanzo]_F].
517 *romanzo]_F].
518 *romanzo]_F].
519 *romanzo]_F].
520 *romanzo]_F].
521 *romanzo]_F].
522 *romanzo]_F].
523 *romanzo]_F].
524 *romanzo]_F].
525 *romanzo]_F].
526 *romanzo]_F].
527 *romanzo]_F].
528 *romanzo]_F].
529 *romanzo]_F].
530 *romanzo]_F].
531 *romanzo]_F].
532 *romanzo]_F].
533 *romanzo]_F].*******************

534 What about cases where the topic is neither a dis-
535 located expression, nor a clitic? Our annotation
536 method should be able to single out such cases, but
537 this is not always true. The example above nicely
538 illustrates a case where our annotators disagreed
539 about labelling a given referential expression as
540 topic: the PP *del romanzo* in A₃, which is already
541 introduced in the previous sentence, A₁. One an-
542 notator chose to nevertheless include it in the focus
543 and label A₃ as an all-focus assertion. The other
544 annotator, while annotating a similar QUD, chose
545 to label the PP as a topic. Indeed, strictly speak-
546 ing, this given PP should then also be part of the
547 QUD ("What for, with respect to the novel?").

- 548 (10) Q₃: What for with respect to the novel?
549 > A₃: [per [conoscere retroscena e curiosità]_F
550 *to know ins-and-outs and trivia*
551 [del romanzo]_T].
552 *of-the novel*
553 *of-the novel*

554 It may be observed that the PP *del romanzo* is
555 embedded inside the verb's direct object NP. Al-
556 though informational categories are in principle
557 independent from the syntactic structure, we may
558 suppose that it was the syntactically embedded po-
559 sition of this expression that led one annotator to
560 consider it as part of the focus (which is the first
561 part of the NP *retroscena e curiosità del romanzo*).
562 The relationship between the given-new structure
563 and the syntactic structure has not been discussed
564 by Riester et al. (2018), but it is something that
565 might be worth addressing in the future.

566 In other cases, the topic was well identified by
567 both annotators, such as *le due lingue* in (11).

- 568 (11) A₁: [Spesso]_{NAI} [si pensa]_{NAI} [che sia
569 *often one thinks that is*

550 bilingue solo [chi è stato esposto a due
551 *bilingual only who has been exposed to two*
552 *lingue fin dalla tenera infanzia]_F]
553 *languages since earliest infancy**

- 554 Q_{1.1}: {One thinks that bilinguals are those
555 who do what, with such two lan-
556 guages?}
557 > A_{1.1}: e [[parla]_F [le due lingue]_T [in modo
558 *and speaks the two languages in way*
559 perfetto e equivalente]_F].
560 *perfect and equivalent*

561 In this example syntax does not help to identify
562 the topic status of the direct object *le due lingue*.
563 Such expression is mentioned in A₁ as part of the
564 focus, but instead of being promoted to topic in the
565 subsequent utterance by some syntactic device for
566 topic shift (such as left dislocation, cf. Lambrecht
567 1994), it is left in situ. One reason for the speaker's
568 choice may be the fact that the topic expression
569 is inside a free relative, a construction that seems
570 to be incompatible with dislocations, as the unac-
571 ceptability of examples below shows:

- 572 (12)a. ??Chi l'italiano, lo conosce sa bene
573 *who the italian it knows knows*
574 dove sta l'errore.
575 *well where is the mistake*
576 b. ??Ho dato un bel voto a chi il primo
577 *I-have given a good note to whom the first*
578 esercizio, lo ha fatto bene.
579 *exercise it they-have done well*

580 Since *due lingue* is mentioned in the previous
581 sentence, the context tells us that this expression
582 is clearly background. Since it's a referential ex-
583 pression, it has all that is required to be identified
584 as topic. Note that a clitic pronoun might have
585 been acceptable here (see example (13)), but this
586 option is not chosen by the speaker/writer.

- 587 (13) > A_{1.1}: e [[le]_T [parla]_F [in modo perfetto e
588 *and them speaks in way perfect and*
589 equivalente]_F].
590 *equivalent*

591 The mechanism of identifying parallel struc-
592 tures (multiple answers to the same question) is
593 a strategy that our annotation tool provides to help
594 recognizing 'hidden' topics.

- 595 (14) A₅₃: I genitori dovrebbero lasciare spazio
596 *the parents should leave space*
597 al bambino o bambina che
598 *to-the boy or girl which*
599 c'è in loro
600 *there is in them*
601 Q₅₄: {To do what?}

- 600 > A₅₄: [[per giocare con i figli]_F],
601 to play with their children
602 > Q₅₅: {Parents should experience lan-
603 guages in what way?}
604 >> A_{55'}: [dovrebbero [soprattutto]_{NAI} vivere
605 they-should above-all live
606 [le lingue]_T [come esperienza]_F]
607 the languages as experience
608 >> A_{55''}: e [[non come performance da
609 and not as performance to
610 misurare]_F].
611 measure

611 Clearly, the fact that *le lingue* (which again oc-
612 cupies a canonical post-verbal position in A_{55'}) is
613 elided in A_{55''}, shows that it represents shared ma-
614 terial between A_{55'} and A_{55''}, and therefore is part
615 of the background.

616 Cases of topic shift were easily recognized by
617 the two annotators. One example is given be-
618 low in (15). The referent *la mamma che parla*
619 *la lingua minoritaria per crescere i suoi bambini*
620 *bilingui* is introduced in the overt question Q_{24.1}
621 and then it continues as topic in the answer A_{24.1}.
622 Then the topic changes and becomes *i bambini* in
623 A₂₅. In A₂₆, the topic changes back to *la mamma*
624 *madrelingua*.

- 625 (15)Q_{24.1}: La mamma che parla la lingua
626 the mother who speaks the minority
627 minoritaria per crescere i suoi
628 language to raise the her
629 bambini bilingui, cosa fa?
630 bilingual children, what she-does
631 > A_{24.1}: [[Parla la propria lingua ai
632 she-speaks the her-own language to-the
633 figli.]_F]
634 children'
635 > Q₂₅: {What do the children do?}
636 >> A₂₅: Solo che [molto spesso]_{NAI} [[i
637 only that very often the
638 bambini]_T [pur capendola
639 children even understanding-her
640 perfettamente]_{NAI} [non parlano
641 perfectly not speak
642 attivamente la sua lingua]_F]
643 actively the her language
644 >> Q₂₆: {What can the mother do then?}
645 >>> A₂₆: Ecco quindi che [[la mamma
646 there then that the mother
647 madrelingua]_T può [cominciare ad
648 mothertongue can start to
649 usare la creatività]_F]
650 use the creativity

647 The fact that the topic is a preverbal subject also
648 helped the annotators to recognize it. As discussed
649 in (Brunetti, 2009), preverbal subjects are typi-

650 cal sentence topics, and our two annotators agreed
651 more often when the topic was in that position.
652 The so-called hidden topics were more challeng-
653 ing.

654 And even if an expression was correctly in-
655 cluded within the background, the two annotators
656 still had to decide for every referential item that
657 was part of the background whether to label it as
658 a topic or not. Not surprisingly, they sometimes
659 agreed, as in (11), and they sometimes picked dif-
660 ferent elements. Since there are several character-
661 istics of the text and the preceding discourse that
662 have to be taken into account for the identification
663 of possible topics, we hypothesise that this cate-
664 gory will probably always be annotated with less
665 accuracy than the other information structure cat-
666 egories such as focus or non-at-issue material.

667 7 Conclusion

668 We have presented a novel method for the annota-
669 tion of information structure which achieves good
670 inter-annotator scores. In particular the agree-
671 ment scores for focus are much higher than the
672 results reported in other similar annotation stud-
673 ies on naturally occurring data (cf. Ritz et al.
674 2008). The method is based on the reconstruction
675 of QUDs, from which the annotation of IS cate-
676 gories is then derived. The results of our inter-
677 annotator agreement analysis show that the QUD-
678 based approach, originally spelled out for English
679 and German, can successfully be transferred cross-
680 linguistically, supporting good agreement for the
681 annotation of central information structure notions
682 such as *focus* and non-at-issueness, with (*con-*
683 *trastive*) *topic* and *background* showing lower lev-
684 els of agreement for some texts due to underrep-
685 resentation of those information structural cate-
686 gories in some of the data analyzed. Thanks to the
687 QUD-based method, attention was drawn to some
688 interesting aspects of Italian information structure,
689 and in particular of Italian topics. Some difficul-
690 ties of topic identification were shown to be re-
691 duced by the adopted annotation procedure.

692 References

- 693 David Beaver and Brady Clark. 2008. *Sense and Sen-*
694 *sitivity. How Focus Determines Meaning*. Wiley-
695 Blackwell, Oxford.
696
697 Lisa Brunetti. 2009. On the semantic and contextual
698 factors that determine topic selection in italian and
699 spanish. *The linguistic review*, 26(2-3):261–289.

- 700 Daniel Büring. 2003. On D-trees, beans, and B-accents. *Linguistics and Philosophy*, 26(5):511–545. 750
- 701 751
- 702 752
- 703 Daniel Büring. 2008. 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. 753
- 704 754
- 705 755
- 706 756
- 707 Daniel Büring. 2016. *Intonation and Meaning*. Oxford University Press. 757
- 708 758
- 709 Sasha Calhoun, Jean Carletta, Jason Brenier, Neil Mayo, Dan Jurafsky, Mark Steedman, and David Beaver. 2010. [The NXT-format Switchboard corpus: A rich resource for investigating the syntax, semantics, pragmatics and prosody of dialogue](#). *Language Resources and Evaluation*, 44:387–419. 759
- 710 760
- 711 761
- 712 762
- 713 763
- 714 764
- 715 765
- 716 766
- 717 767
- 718 768
- 719 769
- 720 770
- 721 771
- 722 772
- 723 773
- 724 774
- 725 775
- 726 776
- 727 777
- 728 778
- 729 779
- 730 780
- 731 781
- 732 782
- 733 783
- 734 784
- 735 785
- 736 786
- 737 787
- 738 788
- 739 789
- 740 790
- 741 791
- 742 792
- 743 793
- 744 794
- 745 795
- 746 796
- 747 797
- 748 798
- 749 799
- Arndt Riester, Lisa Brunetti, and Kordula De Kuthy. 2018. Annotation guidelines for Questions under Discussion and information structure. In Evangelia Adamou, Katharina Haude, and Martine Vanhove, editors, *Information Structure in Lesser-Described Languages: Studies in Syntax and Prosody*. Benjamins, Amsterdam.
- Arndt Riester and Asako Shiohara. 2018. Information structure in sumbawa: A qud analysis. In Asako Shiohara Sonja Riesberg and Atsuko Utsumi, editors, *Perspectives on Information Structure in Austronesian Languages*, pages 285–311. Language Science Press, Berlin.
- Julia Ritz, Stefanie Dipper, and Michael Götze. 2008. Annotation of information structure: An evaluation across different types of texts. In *Proceedings of the Sixth International Conference on Language Resources and Evaluation (LREC)*, pages 2137–2142, Marrakesh.
- Craige Roberts. 2012. [Information structure in discourse: Towards an integrated formal theory of pragmatics](#). *Semantics and Pragmatics*, 5(6):1–69.
- Mats Rooth. 1992. A theory of focus interpretation. *Natural Language Semantics*, 1(1):75–116.
- Roger Schwarzschild. 1999. GIVENness, AvoidF, and other constraints on the placement of accent. *Natural Language Semantics*, 7(2):141–177.
- Seid Muhie Yimam, Iryna Gurevych, Richard Eckart de Castilho, and Chris Biemann. 2013. [WebAnno: A flexible, web-based and visually supported system for distributed annotations](#). In *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics: System Demonstrations*, pages 1–6, Sofia.
- Anja Latrouite and Arndt Riester. 2018. The role of information structure for morphosyntactic choices in tagalog. In Asako Shiohara Sonja Riesberg and Atsuko Utsumi, editors, *Perspectives on Information Structure in Austronesian Languages*, pages 247–284. Language Science Press, Berlin.
- Daniel Marcu, Estibaliz Amorrortu, and Magdalena Romera. 1999. Experiments in constructing a corpus of discourse trees. In *Proceedings of the ACL Workshop on Standards and Tools for Discourse Tagging*, pages 48–57, College Park, MD, USA.
- Uwe Reyle and Arndt Riester. 2016. Joint information structure and discourse structure analysis in an Underspecified DRT framework. In *Proceedings of the 20th Workshop on the Semantics and Pragmatics of Dialogue (JerSem)*, pages 15–24, New Brunswick, NJ, USA.
- Arndt Riester. to appear. Constructing QUD trees. In Malte Zimmermann, Klaus von Heusinger, and Edgar Onea, editors, *Questions in Discourse. Vol. 2: Pragmatics*. Brill, Leiden.