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

Kordula De Kuthy
SFB 833
University of Tübingen
dekuthy@uni-tuebingen.de

Lisa Brunetti
Université de Paris,
LLF, CNRS
lisa.brunetti@
linguist.univ-paris-diderot.fr

Marta Berardi
SFB 833
University of Tübingen
mberardi@sfs.uni-tuebingen.de

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 **?**, **?** and **?**. 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 (**?**), and various Austronesian languages as discussed in **?** and **?**. On the applied side, **?** 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 **?**, such as *focus, background, contrastive topic, nai* and *topic*. Topic is regarded as a notoriously difficult label in agreement studies (cf. **??**). 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 framwork introduced in **?** 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.
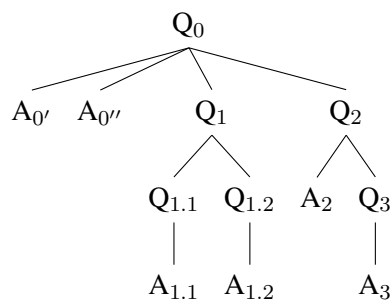

Figure 1: QUD tree

The QUD-tree framework as spelled out in **?** 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 (**?**), we just briefly introduce some basic principles here.

113

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) A_4: Ho appena terminato un romanzo
I-have just finished a novel
'I just finished a novel'
$A_{4'}$: e sono già al lavoro su un nuovo progetto.
and am already at work on a new project
'and I'm already working on a new project.'

(2) $A_{7''}$: Di recente ho ripreso a leggere i
of recently I-have re-started to read the
romanzi di formazione,
novels of coming-of-age
$A_{7'''}$: senza mai tralasciare la narrativa
without ever neglecting the contemproray
contemporanea e i romance.
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
A_{26}:[[ho voluto che]$_{NAI}$ fosse ["forte" ma non
I-have wanted that he-was strong but not
"invincibile"]$_F$].
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 (**????**), cf. **?**:

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) A_{31}: 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
'Even among the early bilinguals who speak two languages, the two languages are almost never completely equivalent,'
Q_{32}: {What about the two languages instead?}
> A_{32}: and [[normalmente]$_{NAI}$ [[ogni lingua]$_T$ [si
and usually each language itself
sviluppa in un contesto specifico]$_F$]$\sim$
develops in a context specific
'and usually each language develops in a specific context.'

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

The resulting tree structure is shown in Figure 2.

$$Q_{25}$$
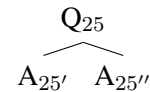$$A_{25'} \quad A_{25''}$$

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. **?**, **?**, **?**, **?** or **?**. The Table in 1 shows the definitions for the information structure categories as introduced in **?**. These are the basis for the labels used in our annotation study.

(7) Q_7: {Cosa ti piace leggere?}
what you like to-read

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

'What do you like to read?'
$> A_7$: [[Di recente]$_{NAI}$ [[ho ripreso a
of recently I-have re-started to
leggere]$_{BG}$ [i romanzi di formazione]$_F$]$\sim$
read the novels of coming-of-age
'I recently started to read the novels of 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_7 in the textual representation marks subordination in the discourse tree, as shown in Figure 2. The focus is *i romanzi di formazione* 'coming-of-age novels', which is labelled [ ]$_F$ and constitutes the answer to the QUD Q_7. The background is linguistic content that is mentioned in this QUD. The question is about what books the interviewee reads or likes to read, so *ho ripreso a leggere* 'I've restarted to read' (labelled [ ]$_{BG}$) is clearly recoverable from the QUD. Focus and background together form the focus domain, labelled [ ]$\sim$. The sentence initial phrase *Di recente* 'recently' is not relevant to answer the QUD Q_7, which would still receive an answer without it, therefore it is labelled [ ]$_{NAI}$.

An example of the label Topic T is given in (8).

(8) Q_{32}: Come puoi riassumere ai tuoi
how can-you summarize to-the your
lettori questo romanzo?
readers this novel
'How can you summarize this novel to your readers?' [...]
$> A_{32}$: [[Senza Etichette]$_T$ [è la storia di Dave]$_F$]$\sim$,
Senza Etichette is the story of Dave

In A_{32}, the clause initial phrase *Senza etichette*, the novel's title, is part of the background (in fact, it is the only background in that utterance) because it is mentioned in Q_{32}. Since it is a referential expression, it is marked [ ]$_T$.

In (9), an example of a contrastive topic (CT) is given.

(9) Q_{10}: Se dovessi esprimere tre desideri?
if you-had to-express three wishes
[...]
$> Q_{10.1}$ {What is your first wish?}
$>> A_{10.1}$: quindi [[il primo]$_{CT}$ sarebbe: [la
so the first-one would-be the
libertà e la felicità di mio
freedom and the happiness of my
figlio]$_F$]$\sim$;
son
$> Q_{10.2}$: {What is your second wish?}
$>> A_{10.2}$: [[il secondo]$_{CT}$ è [riuscire a
the second is to-succeed to
emozionare quanti più lettori
touch as-much more readers
possibile]$_F$]$\sim$,
as-possible

The (explicit) question Q_{10} asks the interviewee to tell three wishes. The speaker answers by uttering three different assertions each about one wish. Clearly, *il primo* 'the first (wish)' in $A_{10.1}$ and *il secondo* 'the second (wish)' in $A_{10.2}$ are members of the alternative set mentioned in Q_{10} (*tre desideri* 'three wishes').

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 **?** to annotate two Italian blog interviews.

For the QUD-based annotation we use the tool *TreeAnno* introduced by **?**, which enables the analyst to semi-automatically segment texts, system-

atically enhance them with implicit *Questions under Discussion (QUDs)*, and transform the data into a discourse tree called *QUD tree*, as described in **?**.

4.1 Evaluation setup

Two trained annotators (and also native speakers of Italian) analyzed and annotated two short Italian blog interviews downloaded from the internet [1]. 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 included in the Appendix.

4.2 Method and results

For the comparison of the two annotated documents, we follow the method described in **?**. 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 (**?**) based on our QUD annotations, **?** propose to follow the method described in **?**, 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 sets of units (i.e. text segments) that are 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 possible spans of segments.[2] For our two annotated documents we calculated κ values for the annotation charts derived from our QUD annotations, based on the above described method. For the text *Italian Blog 1,* consisting of 95 segments, we calculated the κ statistics based on 4,256 items (i.e. possible spans of segments), for *Italian Blog 2* with 113 segments based on 6,187 items. The results are shown in Table 2.

---

[2]Generally, for n segments contained in a document, the number of possible text spans is $\frac{n \times (n+1)}{2}$.

Text	Segments	Spans	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 (**?**) 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 (**?**) 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 (**?**), 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 **?**: focus (F), background (BG), non-at-issue material (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 (**?**) 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,
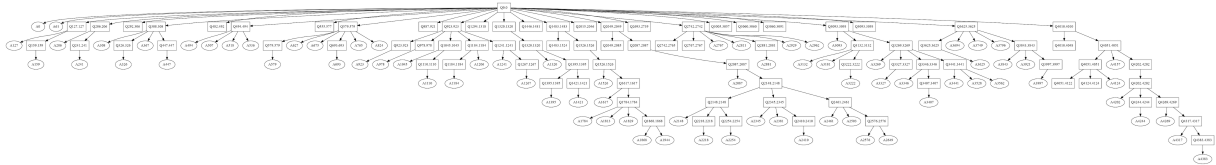
Figure 3: A QUD tree analyses for Italian Blog 1

Figure 4: Annotation in WebAnno

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

following previous work (**???**). In addition to the specifications in **?**, 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.

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 agreements 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 (**?**) 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. **?** 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 **?**, 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 **?** 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.

(10)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

Clitic personal pronouns, such as *le* in A_2 in (11), are also typical candidates for (continuing) topics.

(11) A_1: Abbiamo fatto quattro chiacchiere con Maria Verdiana Rigoglioso per parlare di *Senza Etichette*, il romanzo che ha pubblicato con Libromania.
     'We had a chat with Maria Verdiana Rigoglioso to talk about *Senza Etichette*, the novel she published with Libromania.'

   Q_2: {What did you do with her exactly?}

> A_2:[[Le]$_T$ [abbiamo fatto un po' di domande]$_F$]
     to-her we-have made a little of questions
     'We asked her a few questions'
> Q_3:{What for? }
> A_3:[per [conoscere retroscena e curiosità del
     to know ins-and-outs and trivia of-the
     romanzo]$_F$].
     novel
     'to get to know the background and trivia of the novel.'

What about cases where the topic is neither a dislocated expression, nor a clitic? Our annotation method should be able to single out such cases, but this is not always true. The example above nicely illustrates a case where our annotators disagreed about labelling a given referential expression as topic: the PP *del romanzo* in A_3, which is already introduced in the previous sentence, A_1. One annotator chose to nevertheless include it in the focus and label A_3 as an all-focus assertion. The other annotator, while annotating a similar QUD, chose to label the PP as a topic. Indeed, strictly speaking, this given PP should then also be part of the QUD ("What for, with respect to the novel?").

(12) Q_3: What for with respect to the novel?
     > A_3:[per [conoscere retroscena e curiosità]$_F$
           to know ins-and-outs and trivia
           [del romanzo]$_T$] .
           of-the novel

It may be observed that the PP *del romanzo* is embedded inside the verb's direct object NP. Our assumption is that informational categories are defined and identified solely by pragmatic means, in particular by the QUD-related properties given in Table 1. Despite such an assumption, we may suppose that it was the syntactically embedded position of *del romanzo* that led one annotator to consider it as part of the focus, or more precisely, the fact that the focus (*retroscena e curiosità*) did not form a constituent on its own without the PP *del romanzo*. The relationship between the given-new structure and the syntactic structure has not been discussed by **?**, but it is something that might be worth addressing in the future. Of course, if the syntactic position of the topic must be invoked to complete the picture and arrive at its identification, then we expect different levels of complexity in the task of annotating aboutness topics depending on the language.

In other cases, the topic was well identified by both annotators, such as *le due lingue* in (13).

(13) A$_1$: [Spesso]$_{NAI}$ [si pensa]$_{NAI}$ [che sia
often   one thinks   that is
bilingue solo [chi è   stato esposto a due
bilingual only who has been exposed to two
lingue   fin   dalla tenera infanzia]$_F$
languages since earliest   infancy
'It is often thought that only those who have been exposed to two languages since childhood are bilingual.'

Q$_{1.1}$: {One thinks that bilinguals are those who do what, with such two languages?}

> A$_{1.1}$: e   [[parla]$_F$ [le due lingue]$_T$ [in modo
and speaks   the two languages in way
perfetto e   equivalente]$_F$].
perfect and equivalent
'and speak the two languages in a perfect and equivalent way.'

In this example syntax does not help to identify the topic status of the direct object *le due lingue*. Such expression is mentioned in A$_1$ as part of the focus, but instead of being promoted to topic in the subsequent utterance by some syntactic device for topic shift (such as left dislocation, cf. **?**), it is left in situ. One reason for the speaker's choice may be the fact that the topic expression is inside a free relative, a construction that seems to be incompatible with dislocations, as the unacceptability of examples below shows:

(14) a. ??Chi l'italiano, lo conosce sa   bene dove
who the italian it knows   knows well where
sta l'errore.
is   the mistake

b. ??Ho   dato un bel voto a chi   il primo
I-have given a good note to whom the first
esercizio, lo ha   fatto bene.
exercice   it they-have done well

Since *due lingue* is mentioned in the previous sentence, the context tells us that this expression is clearly background. Since it's a referential expression, it has all that is required to be identified as topic. Note that a clitic pronoun might have been acceptable here (see example (15)), but this option is not chosen by the speaker/writer.

(15) A$_{1.1}$: e   [[le]$_T$ [parla]$_F$ [in modo perfetto e
and them speaks   in way   perfect and
equivalente]$_F$].
equivalent

The mechanism of identifying parallel structures (multiple answers to the same question) is a strategy that our annotation tool provides to help recognizing 'hidden' topics.

(16) A$_{53}$:   I   genitori dovrebbero lasciare spazio
the parents   should   leave   space

al   bambino o bambina che
to-the boy   or girl   which
c'è   in loro
there is in them
'Parents should leave room for the child in them'

Q$_{54}$: {To do what?}

> A$_{54}$: [[per giocare con i   figli]$_F$] ,
to   play   with their children

> Q$_{55}$: {Parents should experience languages in what way?}

>> A$_{55'}$: [dovrebbero [soprattutto]$_{NAI}$ vivere [le
they-should above-all   live   the
lingue]$_T$ [come esperienza]$_F$]
languages as   experience
'they should above all live languages as an experience'

>> A$_{55''}$: e   [[non come performance da
and not   as   performance to
misurare]$_F$] .
measure
'and not as a performance to be measured.'

Clearly, the fact that *le lingue* (which again occupies a canonical post-verbal position in A$_{55'}$) is elided in A$_{55''}$, shows that it represents shared material between A$_{55'}$ and A$_{55''}$, and therefore is part of the background.

Cases of topic shift were easily recognized by the two annotators. One example is given below in (17). The referent *la mamma che parla la lingua minoritaria per crescere i suoi bambini bilingui* is introduced in the overt question Q$_{24.1}$ and then it continues as topic in the answer A$_{24.1}$. Then the topic changes and becomes *i bambini* in A$_{25}$. In A$_{26}$, the topic changes back to *la mamma madrelingua*.

(17) Q$_{24.1}$:   La mamma che parla la lingua
the mother   who speaks the minority
minoritaria per crescere i   suoi
language   to raise   the her
bambini bilingui, cosa fa?
bilingual children, what she-does
'The mother who speaks the minority language to raise her bilingual children what does she do?'

> A$_{24.1}$: [[Parla   la propria lingua   ai
she-speaks the her-own language to-the
figli.]$_F$]
children'
'She speaks her own language to her children.'

> Q$_{25}$: {What do the children do?}

>> A$_{25}$: Solo che [molto spesso]$_{NAI}$ [[i
only that very   often   the
bambini]$_T$ [pur capendola
children   even understanding-her
perfettamente]$_{NAI}$ [non parlano
perfectly   not speak
attivamente la   sua lingua]$_F$]
actively   the her language
'It's just that very often children, even

though they understand her perfectly, don't actively speak her language.'

>> Q$_{26}$: {What can the mother do then?}

>>> A$_{26}$: Ecco quindi che [[la mamma
there then that the mother
madrelingua]$_T$ può [cominciare ad usare
mothertongue can start to use
la creatività]$_F$]
the creativity
'This is where mother-tongue mother can begin to use her creativity.'

The fact that the topic is a preverbal subject also helped the annotators to recognize it. As discussed in (**?**), preverbal subjects are typical sentence topics, and our two annotators agreed more often when the topic was in that position. The so-called hidden topics were more challenging.

And even if an expression was correctly included within the background, the two annotators still had to decide for every referential item that was part of the background whether to label it as a topic or not. Not surprisingly, they sometimes agreed, as in (13), and they sometimes picked different elements. Since there are several characteristics of the text and the preceding discourse that have to be taken into account for the identification of possible topics, we hypothesise that this category will probably always be annotated with less accuracy than the other information structure categories such as focus or non-at-issue material.

7 Conclusion

We have presented a novel method for the annotation of information structure which achieves good inter-annotator scores. In particular the agreement scores for focus are much higher than the results reported in other similar annotation studies on naturally occurring data (cf. **?**). The method is based on the reconstruction of QUDs, from which the annotation of IS categories is then derived. 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, with *(contrastive) topic* and *background* showing lower levels of agreement for some texts due to underrepresentation of those information structural categories in some of the data analysed. Thanks to the QUD-based method, attention was drawn to some interesting aspects of Italian information structure, and in particular of Italian topics. Some difficulties of topic identification were shown to be reduced by the adopted annotation procedure. We believe that the discussion of the problems occurring with the labelling of topics in Italian not only contributes to the analysis of topics in Romance languages, but also helps to refine the QUD annotation procedure in general, so that future annotators are more aware of problematic cases which will hopefully lead to even more reliable annotations.
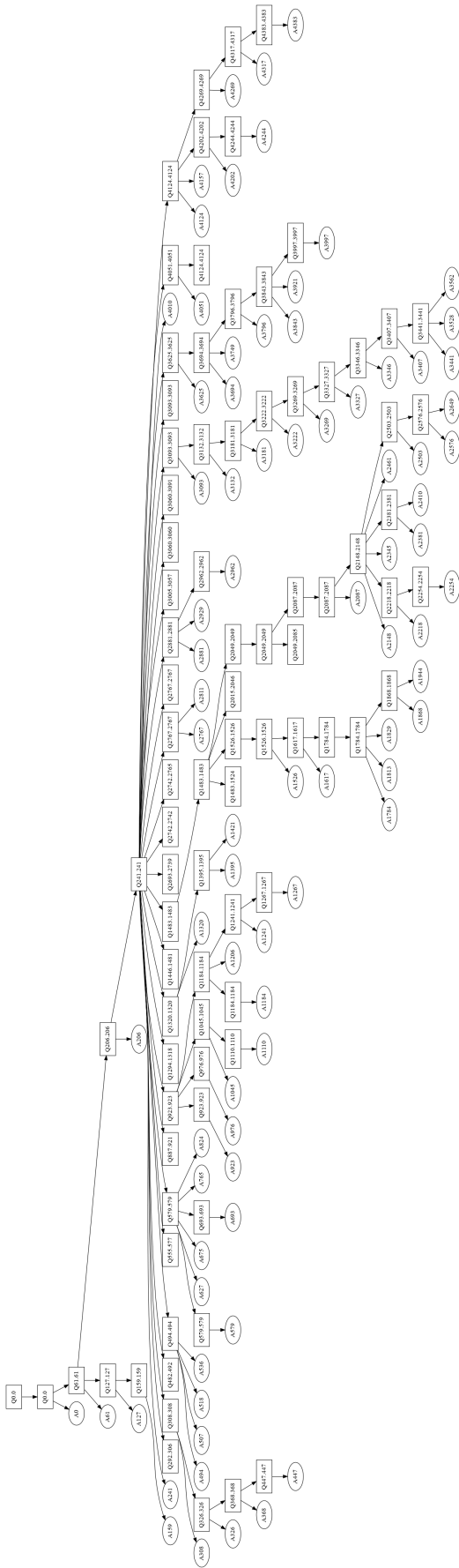
A Appendices

Figure 5: A QUD tree analyses for Italian Blog 1 (Second Annotator)

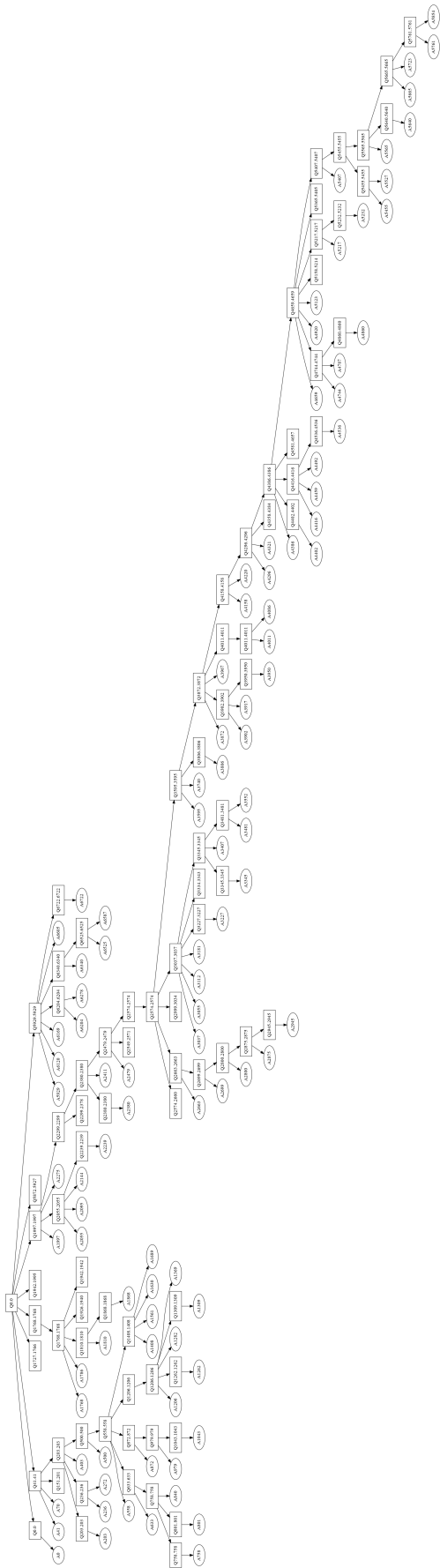Figure 6: A QUD tree analyses for Italian Blog 2 (First Annotator)

Figure 7: A QUD tree analyses for Italian Blog 2 (Second Annotator)