Language Resources and Evaluation manuscript No.
(will be inserted by the editor)

The DialogBank: Dialogues with Interoperable Annotations

Harry Bunt · Volha Petukhova · Andrei Malchanau · Alex Fang · Kars Wijnhoven

the date of receipt and acceptance should be inserted later

Abstract This paper presents the DialogBank, a new language resource consisting of dialogues with gold standard annotations according to the ISO 24617-2 standard. Some of these dialogues have been taken from existing corpora and have been re-annotated, offering the possibility to compare annotations according to different schemes; others have been newly annotated directly according to the standard. The ISO standard annotations in the DialogBank make use of three alternative representation formats, which are shown to be interoperable. The (re-)annotation brought certain deficiencies and limitations of the ISO standard to light, which call for considering possible revisions and extensions, and for exploring the possible integration of dialogue act annotations with other semantic annotations.

1 Introduction

This DialogBank[1] is a new language resource, developed at Tilburg University, which contains dialogues of various kind with gold standard dialogue act annotations according to the ISO 24617-2 standard.[2] This standard builds on previously designed annotation schemes such as DAMSL, DIT^{++}, MRDA, HCRC Map Task, Verbmobil, SWBD-DAMSL, and

Harry Bunt, Kars Wijnhoven
Tilburg Center for Cognition and Communication (TiCC),
Dept. of Communication & Information Sciences, Tilburg University,
Tilburg, The Netherlands
E-mail: harry.bunt@uvt.nl

Volha Petukhova, Andrei Malchanau
Spoken Language Systems, Saarland University,
Saarbrucken, Germany,
E-mail: v.petukhova, a.malchanau@lsv.uni-saarland.de

Alex Fang
Dialogue Systems Group, Department of Linguistics and Translation,
City University of Hong Kong
Hong Kong SAR,
E-mail: alex.fang@cityu.edu.hk

[1] See http://dialogbank.uvt.nl.

[2] ISO 24617-2:2012, Language resource management – Semantic annotation framework – Part 2: Dialogue acts. International Organisation for Standardisation ISO, Geneva. See also Bunt et al. (2010); (2012).

DIT.[3] Most of these schemes have been used to construct annotated corpora, such as the Switchboard, HCRC Map Task, ICSI-MRDA, and DIAMOND corpora.

For nearly all of these annotation schemes, dialogue act annotation consists of segmenting a dialogue into certain grammatical units and marking up each unit with one or more communicative function labels. ISO 24617-2 supports semantically more complete annotation by additionally annotating the following aspects (considered in more detail in Section 2):

1. 'Dimension', or category of semantic content: the annotation scheme supports multidimensional annotation, i.e. multiple communicative functions may be assigned to dialogue segments; different from DAMSL and other multidimensional schemes, an explicitly defined notion of 'dimension' is used that corresponds to a certain category of semantic content. The ISO scheme distinguishes nine dimensions on empirical and theoretical grounds.
2. 'Qualifiers' may be added for expressing that a dialogue act is performed conditionally, with uncertainty, or with a particular sentiment.
3. Dependence relations are defined for expressing semantic relations between dialogue acts, e.g. for indicating which question is answered by a certain answer act (functional dependence relation), or which utterance a feedback act responds to (feedback dependence relation).
4. Rhetorical relations may be annotated to indicate e.g. that one dialogue act explains the performance of another dialogue act.

Most of the dialogues in the DialogBank have been taken from existing corpora and have been re-segmented and re-annotated; some of these also have their original annotations for comparison; this includes dialogues that were previously annotated according to the DIT^{++} annotation scheme, which has been a major source of inspiration for the ISO 24617-2 standard.

The DialogBank presently contains (re-)annotated dialogues from four English-language corpora: HCRC Map Task (Anderson et al., 1991), Switchboard (Jurafsky et al., 1997), TRAINS (Allen et al., 1994) and DBOX (Petukhova et al., 2014); and from four Dutch-language corpora: DIAMOND (Geertzen et al., 2004), Schiphol (Prüst et al., 1984), OVIS (`www.let.rug.nl/vannoord/Ovis`), and the Dutch Map Task corpus (`http://doc.ukdataservice.ac.uk/doc/4632/mrdoc/pdf/4632userguide.pdf`; Caspers, 2000a; 2000b). Dialogues from other corpora, such as the multi-party AMI corpus (`http://groups.inf.ed.ac.uk/ami/corpus/`), the Monroe corpus (Stent, 2000), and the MIB corpus (Petukhova et al., 2016), and in other languages, such as Vietnamese (see Ngo et al., 2018), are planned to be added in the near future.

This paper is organized as follows. Section 2 briefly discusses the use of the ISO 24617-2 standard for the interoperable annotation of dialogue act information. Section 3 discusses the re-annotation (and re-segmentation) of dialogue data from existing corpora, using the pivot XML format of the DiAML markup language defined in the ISO standard. Section 4 introduces two alternative representation formats for ISO 24617-2 annotations, exploiting the distinction of the abstract and concrete syntax made in the definition of DiAML. The interoperability of the three representation formats is shown and their advantages and disadvantages are discussed. Section 5 is concerned with the limitations of the ISO standard that were brought to light during the re-annotation of existing dialogue data and the construction

[3] See Allen & Core (1997); Bunt (2007); Shriberg et al. (2004); Anderson et al. (1991); Alexandersson et al. (1998); Jurafsky et al. (1997); and Bunt (1994; 2000), respectively.

of mappings between different representations. Section 6, finally, contains conclusions from the experiences in building the DialogBank and indicates directions for future work.

2 Interoperable Annotation and the ISO 24617-2 Standard

2.1 Annotations and their Representation

The main motivation for designing annotation standards is to promote the interoperability of annotated corpora. Interoperability of annotations is partly a matter of interchangeable representation formats, such as XML, but more importantly of the underlying concepts. Different annotations can be interpreted across platforms and frameworks only if they encode the same information, or information that can be interpreted through a well-defined mapping. Interoperability at conceptual and semantic levels is of more fundamental importance than interoperability at the level of representation formats, therefore the design of ISO 24617-2 has focused on the identification and specification of empirically and theoretically well-motivated concepts and precise definitions.

ISO 24617-2 includes a comprehensive, application-independent annotation scheme with well-defined concepts and the markup language DiAML (Dialogue Act Markup Language), designed in accordance with the ISO Linguistic Annotation Framework (LAF)[4] and the ISO Principles of Semantic Annotation ('SemAF Principles').[5] LAF makes a fundamental distinction between *annotation* and *representation*: 'annotation' refers to the linguistic information that is added to segments of language data, independent of format; 'representation' refers to the rendering of annotations in a particular format.

Following SemAF Principles, this distinction is implemented in the DiAML definition in the form of an *abstract syntax* that specifies a class of abstract *annotation structures*, which are set-theoretical constructs like pairs and triples, and a *concrete syntax* that specifies a rendering of these annotation structures in a reference format using XML. This reference format is called DiAML-XML. It uses abbreviated XML-expressions, is complete and unambiguous relative to the abstract syntax, i.e. (1) the concrete syntax defines a representation for every structure defined by the abstract syntax; and (2) every expression defined by the concrete syntax represents one and only one structure defined by the abstract syntax. A format with these properties is called *ideal*. Any ideal representation format can be converted through a meaning-preserving mapping to any other ideal format (see Bunt, 2010 for formal definitions and proofs). This is discussed in connection with alternative representations of annotations in the DialogBank in Section 4.

The dialogues in the DialogBank have all been (re-)annotated using the DiAML markup language and the DiAML-XML representation format; additionally, they have also been cast in two alternative tabular representation formats, defined in such a way that they are demonstrably ideal (complete and unambiguous) and more convenient for human readers than XML-based representations.

[4] ISO 24612:2010, Language resource management: Linguistic annotation framework (LAF). International Organisation for Standardisation ISO, Geneva. See also Ide and Romary (2004).

[5] ISO 24617-6:2016, Language resource managemen- St -emantic annotation framework – Part 6: Principles of semantic annotation. International Organisation for Standardisation, Geneva. See also Bunt (2015)

2.2 Main Features of ISO 24617-2 Annotations

As mentioned in the Introduction, ISO 24617-2 annotations differ from most other existing dialogue act annotation schemes in the way they make use of dimensions, qualifiers, and dependence relations and rhetorical relations among dialogue acts. Each of these features is briefly described here.

Dimensions: Utterances in dialogue often have more than one communicative function, as several authors have observed (Allwood, 1992; Bunt, 1994; 2011; Popescu-Belis, 2005; Traum, 2000). The following dialogue fragment illustrates this:

(1) 1. Anne: Henry, can you take us through these slides?
2. Henry: Ehm... sure, just ordering my notes.

In the first utterance, Anne makes a request and assigns the next speaking turn to Henry. In the second utterance, Henry accepts the turn and stalls for time; accepts the request, and explains why he does not fulfill the request right away. The DIT^{++} annotation scheme was designed to optimally support the annotation of multifunctional utterances (Bunt, 2009; 2011). It is based on a well-founded notion of dimension, inspired by the observation that participants in a dialogue perform a range of communicative activities beyond those that relate directly to performing a certain task or activity. They also give and elicit feedback, take turns, stall for time, and demonstrate and monitor attention; moreover, they often perform several of these activities at the same time. The term 'dimension' refers to these various types of communicative activity.

The ISO 24617-2 annotation scheme inherits the following nine dimensions from the DIT^{++} scheme: (1) *Task:* dialogue acts that move the task or activity forward which motivates the dialogue; (2-3) *Feedback*, divided into *Auto-* and *Allo-Feedback*: acts providing or eliciting information about the processing of previous utterances by the current speaker or by the current addressee, respectively; (4) *Turn Management:* activities for obtaining, keeping, releasing, or assigning the right to speak; (5) *Time Management:* acts for managing the use of time in the interaction; (6) *Discourse Structuring:* dialogue acts dealing with topic management, opening and closing (sub-)dialogues, or otherwise structuring the dialogue; (7-8) *Own-* and *Partner Communication Management:* actions by which the sender edits his current contribution or a contribution of another current speaker, respectively; (9) *Social Obligations Management:* dialogue acts for greeting, thanking, apologizing, and other social conventions in communication.

The ISO 24617-2 inventory of communicative functions contains 56 functions, subdivided into general-purpose functions and *dimension-specific functions*. Dimension-specific communicative functions are specific for a particular dimension; for instance *Turn Take* is specific for Turn Management; *Stalling* is specific for Time Management, and *Self-Correction* is specific for Own Communication Management. General-purpose communicative functions, by contrast, can be used in any dimension; for example, *"You misunderstood me"* is an *Inform* in the Allo-Feedback dimension. All types of question, statement, and answer can be used in any dimension, and the same is true for commissive and directive functions, such as *Offer, Suggest,* and *Request*. Table 1 lists the communicative functions defined in ISO 24617-2.

Qualifiers: Three types of qualifiers are included in ISO 24617-2, namely for indicating a speaker's (un-)certainty, (un-)conditionality, and sentiment, For certainty only two rather coarse-grained qualifiers are defined, *certain* and *uncertain*, and likewise for conditionality:

Table 1 ISO 24617-2 communicative functions

General-Purpose	*Dimension-Specific Communicative Functions*	
Communicative Functions	*Function*	*Dimension*
Inform	AutoPositive	Auto-Feedback
Agreement	AutoNegative	
Disagreement	AlloPositive	Allo-Feedback
Correction	AlloNegative	
Answer	FeedbackElicitation	
Confirm	Staling	Time Management
Disconfirm	Pausing	
Question	Turn Take	Turn Management
Set-Question	Turn Grab	
Propositional Question	Turn Accept	
Choice-Question	Turn Keep	
Check-Question	Turn Give	
Offer	Turn Release	
Address Offer	Self-Correction	Own Communication Man.
Accept Offer	Self-Error	
Decline Offer	Retraction	
Promise	Completion	Partner Communication Man.
Request	Correct Misspeaking	
Address Request	Interaction Structuring	Discourse Structuring
Accept Request	Opening	
Decline Request	Init-Greeting	Social Obligations Man.
Suggest	Return Greeting	
Address Suggest	Init-Self-Introduction	
Accept Suggest	Return Self-Introduction	
Decline Suggest	Apology	
Instruct	Accept Apology	
	Thanking	
	Accept Thanking	
	Init-Goodbye	
	Return Goodbye	

conditional and *unconditional*. For sentiment the coarse-grained values *positive* and *negative* have been considered, and have been used in some dialogue annotations; however, the ISO standard does not specify any particular set of sentiment qualifiers; such values are expected to be provided by ongoing research on sentiment analysis and representation. The different qualifiers are applicable to different classes of dialogue acts. Sentiment qualifiers are applicable to any dialogue act with a general-purpose function (GPF); conditionality qualifiers to dialogue acts with a commissive or directive function (*Promise, Offer, Suggestion, Request*, etc.); and certainty qualifiers are applicable to dialogue acts with an 'information-providing' function' (*Inform, Agreement, Disagreement, Correction, Answer, Confirm, Disconfirm).*

Functional dependence relations are indispensable for the interpretation of dialogue acts that are responsive in nature, such as *Answer, Confirmation, Disagreement, Accept Apology,* and *Decline Offer*. The meaning of these acts depends crucially on the dialogue act that they respond to. Functional dependence relations connect occurrences of such dialogue acts to their 'antecedent' and correspond to links for marking up a segment not only as having the function of an answer, for example, but also indicating which question is answered.

Note that ISO 24617-2 in its present form does not support the marking up of the semantic content of a dialogue act (but a future revision may be extended in this direction; see Bunt et al., 2017); the content information concerning a dialogue act is in its dimension, which

can be viewed as indicating a type of semantic content (e.g., the content of a dialogue act in the Task dimension is task-related information; that in a feedback dimension is processing information; that in the Turn Management dimension is information about the allocation of the speaker role, etc.). Dialogue acts have a formal semantics in terms of updating the information states of dialogue participants (see Bunt, 2014) which interprets DiAML annotations as update operations that apply to a semantic content specification.

Feedback dependence relations play a similar role for interpreting feedback acts as functional dependence relations for responsive dialogue acts; their meaning is partly or entirely determined by the utterance(s) that the feedback is about. This is obvious for 'inarticulate' feedback acts, like *OK"* and *"Yes"*. Feedback acts often refer to the immediately preceding utterance, but can also refer further back and to more than one utterance (Petukhova et al., 2011). The ISO 24617-2 annotation scheme therefore includes links for marking up these dependences; an example can be seen in (8b).

Rhetorical relations have been studied mostly for their occurrence in written texts, where they are crucial for a full understanding of the individual sentences, but they also play a role in spoken dialogue where they occur in two different ways, illustrated in the following examples (where the participants talk about remote controls and their design):

(2) 1. A: I can never find them.
    2. B: That's because they don't have a fixed location.

(3) 1. A: Where would you position the buttons?
    2. A: I think that has some impact on many things

In (2) the dialogue acts expressed by A's and B's utterances are related by a *Cause* relation between their respective semantic contents: the content of the second causes the content of the first; in (3), by contrast, the second dialogue act forms a reason for performing the first, so the causal relation is between the two dialogue acts as a whole, rather than between their semantic contents. The annotation of a rhetorical relation is illustrated in example (8b).

Different from functional and feedback dependences, which are an integral part of dialogue acts with a responsive function and of feedback acts, respectively, rhetorical relations give additional information about the ways in which dialogue acts are semantically or pragmatically related. Similar to the case of sentiment qualifiers, the ISO 24617-2 standard does not specify any particular set of rhetorical relations, but rather expects such a set to be provided by ongoing research in that area (see e.g. Burkhardt et al., 2017). In the mean time, it has become common practice to use a slightly extended version of the DR-Core set of relations; see Section 5.5

2.3 Segmentation

According to ISO 24617-2, dialogue acts are expressed by 'functional segments' of linguistic or other communicative behaviour, defined as *minimal stretches of communicative behaviour that have a communicative function*, 'minimal' in the sense of not including any material that does not contribute to the expression of that function (or to the specification of the semantic content of the dialogue act). Functional segments are mostly shorter than turns, may be discontinuous, may overlap, and may contain parts contributed by different speakers. A segment carrying a feedback function for instance frequently overlaps with a segment that carries a task-related function.

The requirement of being ‘minimal’ has been added in order for communicative functions to be assigned as accurately as possible to those stretches of behaviour which express these functions. The following example illustrates this:

(4) Can you tell me what time the train to *ehm,...* Viareggio leaves?

The speaker interrupts himself while formulating a request for information since he needs a bit of time to provide the name of a destination. The small interrupting segment *ehm,...* does not contribute to the expression of the request, so according to the minimality condition it does not belong to the functional segment that corresponds to the request. The utterance in (4) should thus be analysed as consisting of two functional segments: the discontinuous segment *Can you tell me what time the train to [ ] Viareggio leaves?*, corresponding to a request, and the segment *ehm,...* corresponding to a Stalling act. This can be annotated in DiAML as follows, where ‘fs1’ and ‘fs2’ indicate the two functional segments:

(5)
```
<dialogueAct xml:id="da1" target="#fs1"
  speaker="#s" addressee="#a" dimension="task"
  communicativeFunction="request" conditionality="conditional"/>
<dialogueAct xml:id="da2" target="#fs2"
  speaker="#s" addressee="#a"
  communicativeFunction="stalling" dimension="timeManagement"/>
```

Note that in this example the yes-no question of the form *Can you tell me...* has been interpreted as a conditional request, i.e. as: *“Please tell me, if you can,...”*.

A functional segment is most often a part of what is contributed by the participant who occupies the speaker role, but it may happen that a dialogue act is spread over multiple turns, as in the following example, where the utterances in turns 6, 8, 11, and 13 together form the functional segment that contains B’s answer to the question in turn 5:

(6)

1.	A:	I’ve skied in Colorado, and we usually go to New Mexico because it’s a little cheaper —
2.	B:	Ooh,
3.	A:	— you know
4.	B:	Uh-huh
5.	B:	Where in Colorado?
6.	A:	I’ve been to Telluride, which is on the West side,
7.	B:	Yes
8.	A:	and, uh, Copper
9.	A:	Copper is kind of my favorite up there
10.	B:	Really?
11.	A:	Breckennridge —
12.	B:	Uh-huh
13.	A:	— and Keystone

This example forms a tricky case for segmentation and dialogue act annotation, for although the answer is not complete until turn 13, participant B provides intermediate feedback in the turns 7, 10, and 12, and participant A provides an intermediate assessment of the answer part in turn 8, so these answer parts seem to deserve a dialogue act-like status as well. See also the discussion in Section 5.1.

2.4 ISO 24617-2 Metamodel

The metamodel, displayed in Figure 1, shows the classes of concepts that are used in ISO 24617-2 annotations. It indicates that a dialogue act has one sender, one or more addressees, zero or more other participants (such as bystanders or an audience; see Clark, 1996), one dimension, one communicative function, zero or more functional and feedback dependence relations, possibly one or more qualifiers, and possibly one or more rhetorical relations to other dialogue acts.

According to the metamodel, the following ingredients make up an ISO 24617-2 annotation, where the second column indicates the number of each kind of element:

(7)

ingredient	*number*
a functional segment, specifying a stretch of dialogue that carries one or more communicative functions	1
the dialogue acts expressed by a functional segment, with for each dialogue act:	1-9
the sender	1
the addressee(s);	≥ 1
any other dialogue participants	≥ 0
the dimension and communicative function;	1
functional dependence relations (only for responsive acts)	0 or 1
feedback dependence relations (only for feedback acts);	0 or 1
qualifiers (if any);	0 - 3
rhetorical relations between dialogue acts.	≥ 0

Of these elements, rhetorical relations strictly speaking fall outside the scope of ISO 24617-2, which has only a minimal provision for allowing to specify a rhetorical relation between dialogue acts but does not include any particular set of such relations. Since the establishment of ISO 2517-2 as an international standard, it has become common practice by users of the standard to include annotations of rhetorical relations (see Bunt et al., 2017b), mostly using the relations defined in ISO 24617-8.[6]

3 Re-annotation and Re-segmentation

3.1 Overview

The (re-) annotation of the dialogues that were included in the DialogBank started in some cases from raw primary data or transcriptions, and in some cases from previous annotations, represented in a variety of formats. Some of the dialogues were previously re-annotated using a version of the DIT^{++} annotation scheme (see Bunt, 2009 and Petukhova, 2011), which has been the basis for the ISO 24617-2 annotation scheme. Due to the differences between (a) ISO 24617-2 annotations and those of other schemes, and (b) the representation formats used, dialogues from different corpora require different approaches to their (re-) segmentation, (re-) annotation, and (re-) formatting:

[6] ISO 24617-8:2016, Language resource management - Semantic annotation framework – Part 8: Semantic relations in discourse, Core annotation scheme (DR-Core). International Organisation for Standardisation ISO, Geneva.

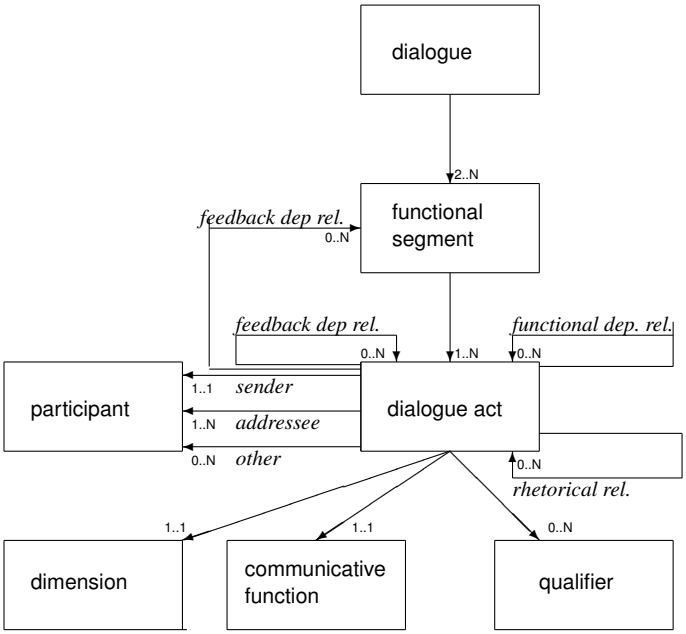

Fig. 1 ISO 24617-2 Metamodel

- The dialogues from the HRCR Map Task and TRAINS corpora have previously been re-annotated according to the DIT^{++} annotation scheme, release 5 (2010; see `http://dit.uvt.nl`) using the ANVIL tool (Kipp, 2001; 2014; Bunt et al., 2012).
  These annotations are represented in XML and contain the type of information that is expected in DiAML, but in a format that differs in several respects from DiAML-XML, and includes various Anvil-specific features. This 'DiAML-Anvil' format required substantial reformatting. The annotations were moreover enriched with the annotation of rhetorical relations, using the DR-Core set of relations, defined in ISO 24617-8 (see Bunt and Prasad, 2016).
- The dialogues from the Switchboard (SWBD-DA) corpus were made available in a 3-column tabular format, with one column containing communicative function labels from the SWBD-DAMSL annotation scheme. Fang et al. (2011; 2012a,b) applied semi-automatic procedures for replacing the SWBD-DAMSL tags by ISO 24617-2 function tags while retaining the SWBD-DA segmentation, showing that 84% of the re-tagging can be done automatically. The resulting 'SWBD-ISO' corpus (see Bunt et al., 2016) forms an interesting resource halfway between the SWBD-DA corpus and an ISO-annotated version of the same; since the ISO standard assumes a more fine-grained way of segmenting dialogue, an annotation that is made directly according to the ISO 24617-2 scheme differs substantially from a SWBD-ISO annotation.
  The Switchboard dialogues in the DialogBank were re-segmented and re-annotated with ISO 24617-2 tags, adding besides communicative function labels also tags for dimension, qualifiers, functional and feedback dependence relations, and DR-Core rhetorical relations.

Origin	Lang	Original and previous representations	Original annotation	Previous annotation
HCRC Map Task (3)	EN	NITE XML DiAML-Anvil	HCRC Map Task communicative functions;	DIT^{++}5.0
Switchboard (4)	EN	3-column tabular	SWBD-DAMSL communicative functions	ISO 24617-2 comm. functions
TRAINS (3)	EN	DiAML-Anvil	DAMSL communicative functions	DIT^{++}4.0
DBOX (5)	EN	DiAML-XML	full ISO 24617-2 annotation	–
DIAMOND (3)	NL	13-column tabular	DIT^{++}3.0 communicative functions and dimensions	DIT^{++}3.0
Dutch Map Task (2)	NL	plain text transcript	no dialogue act annotation	–
OVIS (3)	NL	plain text transcript	DIT^{++} communicative functions and dimensions	DIT^{++}3.0
Schiphol Airport (2)	NL	plain text transcript	DIT^{++} communicative functions and dimensions	DIT^{++}3.0

Table 2 Types of data in the DialogBank corpus.

- The dialogues of the DBOX corpus were annotated using the ANVIL tool, according to the ISO 24617-2 annotation scheme with a few minor extensions, justified by domain-specific requirements.
- The dialogues in the DIAMOND corpus were annotated with the communicative functions and dimensions of DIT^{++} release 3 (2007), using the DitAT annotation tool (Geertzen, 2007). This tool produces representations in a 13-column tabular form, with one column for each of the ten DIT^{++} dimensions (see Section 2.5).
- The dialogues in the Dutch Map Task corpus were collected with the primary aim to study the phonology and phonetics of intonation in dialogue (see Caspers, 2000a; 2000b). The dialogues were made available with orthographic transcription. These dialogues were annotated according to ISO 24617-2 from scratch in a tabular representation format for DiAML annotations which is defined below, see section 3.3.
- The dialogues from the OVIS and Schiphol corpora were annotated with the communicative functions and dimensions of DIT^{++} release 3 (2007) and produced with the ANVIL tool, resulting in representations in DiAML-Anvil format. They were re-annotated from scratch.

Table 1 summarizes the previous and original annotations and representations of the material in the DialogBank. At the time of writing, a total of 25 annotated dialogues were included; this number is steadily increasing. The next two subsections describe the representation of DiAML annotations in XML and their representation in the tabular formats that have been defined to facilitate the inclusion of re-annotated dialogue material in the DialogBank.

3.2 DiAML-XML

The representation of annotations in DiAML-XML makes use of two XML elements, one to represent individual dialogue acts and one to represent a rhetorical relation between dialogue acts. A <dialogueAct> element has attributes whose values represent the following components, corresponding with the components listed in (7):

1. the speaker;

2. the addressee(s);
3. any other participants (possibly none);
4. the communicative function;
5. the dimension;
6. qualifiers (if any); and
7. functional and feedback dependence relations.

Example (8b) shows the use of these XML elements in the representation of the annotation of the dialogue fragment[7] in (8a), which contains a rhetorical relation (Elaboration) between the dialogue acts in utterances 1 and 3, and a feedback dependence between the dialogue acts in utterances 3 and 4.

(8) a. 1. G: go south and you'll pass some cliffs on your right
2. F: uhm...
3. G: and some adobe huts on your left
4. F: oh okay

b.
```
<diaml xmlns="http://www.iso.org/diaml">
<dialogueAct xml:id="da1" target="#fs1" sender="#g" addressee="#f"
    dimension="task" communicativeFunction="instruct" />
<dialogueAct xml:id="da2" target="#fs2" sender="#f" addressee="#f"
    dimension="turnManagement" communicativeFunction="turnTake" />
<dialogueAct xml:id="da3" target="#fs2" sender="#f" addressee="#g"
    dimension="timeManagement" communicativeFunction="stalling" />
<dialogueAct xml:id="da4" target="#fs3" sender="#g" addressee="#f"
    dimension="task" communicativeFunction="inform" />
<rhetoricalLink dact="#da4" rhetoAntecedent="#da1"
    rhetoRel="elaboration" />
<dialogueAct xml:id="da5" target="#fs4" sender="#f" addressee="#g"
    dimension="autoFeedback" communicativeFunction="autoPositive"
    feedbackDependence="#da1 #da4"/>
</diaml>
```

It may be noted here that some of the previous annotations of dialogues in the DialogBank were made using versions of the DIT^{++} annotation scheme and the DiAML-XML representation format, even though DiAML was designed for the ISO 24617-2 annotation scheme rather than for DIT^{++}. This combination is possible due to the fact that the two annotation schemes are structurally identical; the only difference between them is that DIT^{++} contains one more dimension (the Contact Management dimension), and has more fine-grained sets of dimension-specific communicative functions for some of the dimensions (notably for feedback and discourse structuring). For the use of DiAML-XML this simply means that some additional values are defined for the attributes 'dimension' and 'communicativeFunction', as shown in (8b).

An important point to note is that DiAML-XML is in fact a compact way of using XML for representing annotation structures. For example, the annotation shown in (8b) can be regarded as an abbreviation of the XML-representation in (9), in which 'fs' stands for 'feature structure' and 'f' for feature (following ISO standard 24610 for the representation of feature structures)[8].

[7] From the HCRC Map Task corpus, Anderson et al. (1991).

[8] ISO 24610:2006, Language resource management: Feature structures. International Organisation for Standardisation ISO, Geneva; see also Lee et al. (2004).

(9)
```
<fs xml:id="da1">
  <f name="target"><value="#fs1"/></f>
  <f name="sender"><value="#g"/></f>
  <f name="addressee"><value="#f"/></f>
  <f name="dimension"><value="task"/></f>
  <f name="communicativeFunction"><value="instruct"/></f>
</fs>
<fs xml:id="da2">
  <f name="target"><value="#fs2"/></f>
  <f name="sender"><value="#f"/></f>
  <f name="addressee"><value="#g"/></f>
  <f name="dimension"><value="turnManagement"/></f>
  <f name="communicativeFunction"><value="turnTake/></f>
</fs>
<fs xml:id="da3">
  <f name="target"><value="#fs2"/></f>
  <f name="sender"><value="#f"/></f>
  <f name="addressee"><value="#g"/></f>
  <f name="dimension"><value="timeManagement"/></f>
  <f name="communicativeFunction"><value="stalling/></f>
</fs>
<fs xml:id="da4">
  <f name="target"><value="#fs3"/></f>
  <f name="sender"><value="#g"/></f>
  <f name="addressee"><value="#f"/></f>
  <f name="dimension"><value="task"/></f>
  <f name="communicativeFunction"><value="inform"/></f>
</fs>
<fs xml:id="rL1">
  <f name="dact"><value="#da4"/></f>
  <f name="rhetoAntecedent"><value="#da1"/></f>
  <f name="rhetoRel"><elaboration"="#f"/></f>
</fs>
<fs xml:id="da5">
  <f name="target"><value="#fs4"/></f>
  <f name="sender"><value="#f"/></f>
  <f name="addressee"><value="#g"/></f>
  <f name="dimension"><value="autoFeedback"/></f>
  <f name="communicativeFunction"><value="autoPositive/></f>
</fs>
```

Compared to the full XML representation in (9), the relative compactness of the DiAML-XML representation in (8b) is an obvious advantage. The equivalence of the compact form with the full XML form is important for the possibility to combine dialogue act annotations with annotations of other semantic or pragmatic information. This is discussed in Section 5.3.

id	function	transcript
sw0105-0001-A001-01	qw	A.1 utt1: Jimmy, {D so } how do you get most of your news? /
sw0105-0002-B002-01	sd	B.1 utt1: {D Well, [I kind of, + {F uh, } I] watch the {F uh, } national news every day, for one /
sw0105-0003-B002-02	sd	B.2 utt1 I also read one or two papers a day /
sw0105-0004-B002-03	sd	B.3 utt1: {C and } [I'm a, + I'm pretty much a] news junkie /
sw0105-0005-B002-04	sd	B.4 utt1: {C and } I tune in to CNN a lot. /
sw0105-0006-A003-01	ba	A.3 utt1: {F Oh, } wow. /

Table 3 Annotation of Switchboard (SWBD-DA) dialogue fragment.

3.3 Representations in Tabular Form

As mentioned above, some of the dialogues in the DialogBank were previously annotated using tabular formats, where typically the rows correspond to the segmentation; one column contains the transcribed speech; and the other column(s) contain the annotation. This is illustrated in Table 2 for a fragment of a Switchboard dialogue, originally annotated according to the SWBD-DAMSL scheme, which uses a 3-column representation format, and in Table 3 for the annotation in a 13-column format of a fragment of a dialogue from the TRAINS corpus using the DIT^{++} 4.0 annotation scheme and produced with the DitAT annotation tool (Geertzen, 2007).

id	sp	transcript	Task	Auto-Feed-back	Allo-Feed-back	Turn Man.	Time Man.	Discourse Structur-ing	Contact Man.	OCM	PCM	SOM
1	S	hello				Turn Take			Contact indication			Initial Greeting
2	S	can I help you						Offer				
3	U	uhm,				Turn Take	Stalling					
4	U	yes hello, maybe		Eval. positive				Accept Offer				Return Greeting
5	U	I'd like to take a tanker ..	Inform					Topic intro.				

Table 4 Representation in tabular form of DIT^{++} 4.0 annotations produced with the DitAT tool for a fragment of a TRAINS dialogue. Abbreviations used: sp = speaker, OCM = Own Communication management, PCM = Partner Communication Management, SOM = Social Obligations Management.

The formats used in Tables 3 and 4 look rather different, and certainly very different from the XML format used in (8), yet they all contain essentially the same information. For example, the numbers in the first column in Table 3 as well as in Table 4 can be interpreted as identifiers of functional segments; the strings in the third column in both cases as containing the textual transcriptions of these segments; the second column in Table 4 as indicating the speaker of a segment (and, by implication for a two-person dialogue, the addressee), which in Table 3 is part of the information in the third column; and the cells in other columns as representing the dialogue act annotations. The row numbered 4 in Table 4 thus corresponds to the following XML expression (with dialogue act identifiers added):

(10)
```
<dialogueAct xml:id="da1" target="#fs4" sender="#u" addressee="#s"
  dimension="autoFeedback" communicativeFunction="evalPositive" />
<dialogueAct xml:id="da2" target="#fs1" sender="#u" addressee="#s"
  dimension="discourseStructuring"
  communicativeFunction="acceptOffer" certainty="uncertain"/>
```

```
<dialogueAct xml:id="da3" target="#fs1" sender="#u"
  addressee="#s" dimension="socialObligationsManagement"
  communicativeFunction="returnGreeting" />
```

The two tabular formats shown here are less expressive than the DiAML-XML format[9] in that, firstly, the information assigned to dialogue segments is limited to communicative functions only (Table 3) or to communicative functions and dimensions (Table 4); and secondly only contiguous, non-overlapping functional segments can be represented. The former limitation can be overcome by extending the information about a dialogue act in a cell of the table by adding qualifiers, dependences, and rhetorical relations. To overcome the latter limitation, and make the tabular representations compatible with the stand-off requirement of ISO annotation standards, we will describe some further adjustments in the next section. The resulting adaptations of the formats illustrated in Tables 5 and 6 are called DiAML-TabSW and DiAML-MultiTab, respectively, and will be shown to be ideal – complete and unambiguous.

4 Interoperability of Representations

4.1 Abstract Syntax and Alternative Representations

The introduction in the ISO standard of an abstract syntax, besides a concrete representation format, was to allow precise determination of the interoperability of alternative representations. Figure 2 displays the relations between an abstract syntax, one or more alternative ideal (complete and unambiguous) representation formats, and the semantics of a markup language.

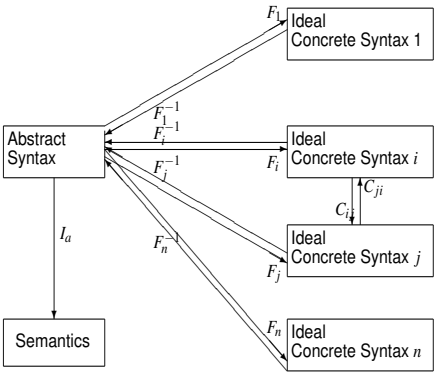

Fig. 2 Abstract and concrete syntax, and semantics

Since the DiAML-XML format is defined in such a way that it is ideal, a function F_{XML} can be defined that maps annotation structures as defined by the abstract syntax to

[9] Switchboard dialogues are also available in the XML-based NXT format.

an XML expression; due to the unambiguity, this function has an inverse F^{-1}_{XML} which maps any DiAML-XML expression to the annotation structure that it represents.

A tabular representation can be formally defined as a matrix of which each row is an n-tuple of elements that correspond to the contents of its cells. The encoding functions $F_{MultiTab}$ and F_{TabSW} are defined below in terms of this formalization, Their existence proves the completeness of the modified tabular formats MultiTab and TabSW. Similarly, defining their inverses $F^{-1}_{MultiTab}$ and F^{-1}_{TabSW}, shows their unambiguity. As a result, the composition of functions such as

(11) $C_{MultiTab \rightarrow XML} = F_{XML}\ o\ F^{-1}_{MultiTab}$

defines a conversion from annotations, represented in the DiAML-MultiTab format to representations in the DiAML- XML format. The inter-convertibility of the three formats is exploited in the DialogBank by allowing the user to view the annotations in the form that is most convenient to him or her, as well as by converting the tabular formats to the XML format for automatic processing, if desired.

4.2 Abstract Syntax

The abstract syntax of DiAML reflects the conceptual analysis of dialogue acts that underlies the ISO 24617-2 annotation scheme. On this analysis, a dialogue act is characterized by seven elements, corresponding to the components listed in (7):

1. the sender; every dialogue act has exactly one sender who is 'responsible' for the act, even though more than one speaker may contribute; see example (12):

   (12) 1. A: and then should I specify the uhm, uhm,
   2. B: budget code, you should specify the budget code, that's 5611

   In this example, A is struggling to formulate a question and B helps by providing the term that A was looking for. The first part of B's utterance is a dialogue act with the communicative function *Completion*, in the Partner Communication Management dimension. The functional segment *and then should I specify the budget code*, made up of parts of what A and B say, expresses a question for which A is 'responsible' and is considered the sender. The second part of B's utterance *"you should specify the budget code"* is an answer to that question (the third part is an elaboration of that answer).
2. one or more addressees; in a two-person dialogue the addressee is just the one who is not the sender; in multiparty dialogues, such as those of the AMI corpus, all the participants who are not the sender are addressees, unless the speaker picks out one of them (in which case the other participants form the 'other participants').
3. zero or more other participants (if any), such as a bystander or an audience;
4. the communicative function;
5. the dimension;
6. zero or more functional dependence relations or feedback dependence relations;
7. zero or more qualifiers of certainty, conditionality, and/or sentiment.

Whether a dialogue act has a dependence relation to another dialogue act is determined by its communicative function and dimension. A functional dependence means that the semantic content of a dialogue act is co-determined by the semantic content of a previous dialogue act, due to having a communicative function of a responsive character. This is for example the case for answers, whose meaning is partly determined by the question that is being

answered, but also for the acceptance or rejection of offers, suggestions, requests, and the acceptance of apologies and thankings.

The semantic content of a feedback act (in the Auto-Feedback or in the Allo-Feedback dimension) is partly determined by what the feedback is about. Feedback utterances like *"OK"*, *"Yes"*, and *"Really?"* illustrate this. While positive feedback acts are typically about the processing of previous dialogue acts, negative feedback acts are often about a problem in understanding something, and may thus refer to a segment of speech rather than to its interpretation as a dialogue act. ISO 24617-2 therefore allows feedback dependence relations to have both dialogue acts and dialogue segments as antecedents.

Since responsive dialogue acts and feedback acts are semantically incomplete without the specification of functional and feedback dependences, these are part of the structures that are used to annotate such acts.

Different from functional and feedback dependence relations, rhetorical relations are not part of the meaning of a dialogue act, but add information to the way two or more semantically complete dialogue acts are related; they are therefore not part of a structure that describes a dialogue act, but they occur in link structures that relate dialogue acts, as illustrated in (8) on page 7.

An abstract syntax consists in general of: (a) a specification of the elements from which annotation structures are built up, called a 'conceptual inventory', and (b) a specification of the possible ways of constructing annotation structures using these elements. The DiAML abstract syntax is defined by the following specification:

Specification 1. DiAML abstract syntax.
a. Conceptual inventory
The DiAML conceptual inventory consists of six sets:

1. A set of dimensions, notably the nine dimensions listed in Section 2.2.
2. A set of communicative functions, namely the 56 functions listed in Table 1; the set is partitioned into 'general-purpose' functions, which can be used in any dimension, and for each dimension except *Task* a set of 'dimension-specific' functions (no task-specific communicative functions are defined, since the annotation scheme is designed to be application-independent). A subset *RSP* of the set of communicative functions is specified as the 'responsive' communicative functions.
3. A set of qualifiers that can be associated with dialogue acts, partitioned into subsets for certainty, conditionality, and sentiment.
4. A set of rhetorical relations that can hold between dialogue acts (or their semantic contents).
5. A set of dialogue participants, including possible side-participants or audiences, besides actively participating speakers and addressees.
6. A set of functional segments of primary data.

The sets of functional segments and dialogue participants are specific for a particular annotation task; the other concepts are task-independent.

b. Annotation structures
An annotation structure is a set

(13) $\{\varepsilon_1, \ldots, \varepsilon_k, L_1, \ldots, L_m\}$

consisting of the entity structures $\{\varepsilon_1, \ldots, \varepsilon_k,\}$, with $k \geq 1$, and the link structures $\{L_1, \ldots, L_m\}$ (with $m \geq 0$). Entity structures contain semantic information about a functional segment; link structures describe semantic relations between functional segments.

An entity structure in DiAML is a pair

(14) $\varepsilon = \langle m, \alpha \rangle$

consisting of a functional segment m (a 'markable') and the characterization of a dialogue act α, which is a 7-tuple as in (15), where S is the sender of the dialogue act; A is a non-empty set of addressees; H is a (possibly empty) set of other dialogue participants; d is a dimension; f is a communicative function; Q is a (possibly empty) set of qualifiers, and Δ is a (possibly empty) set of other dialogue acts that the dialogue act in focus depends on.

(15) $\alpha = \langle S, A, H, d, f, Q, \Delta \rangle$

A link structure in DiAML is a triple (16), consisting of an entity structure ε that corresponds to a dialogue act, a non-empty set E of entity structures that correspond to rhetorically related dialogue acts, and the rhetorical relation ρ that relates the dialogue acts in ε and E.

(16) $\langle \varepsilon, E, \rho \rangle$

4.3 DiAML Representations

4.3.1 Anchoring Annotations in Primary Data

DiAML relies on a three-level architecture:

(1) a primary source, which may correspond to a speech recording, a video clip, a textual transcription, or a low-level annotation thereof;
(2) the marking of functional segments in the primary source;
(3) the dialogue act information associated with the functional segments.

Annotation in DiAML is concerned with level (3) and follows the stand-off annotation approach: annotations refer to segments of the primary data specified at level (2), and the primary data are kept separate. The 3-level architecture is clearly visible in DiAML-XML representations, such as (8), where functional segments appear as the values of the 'target' attribute, which are assumed to be given as markables; Figure 3 shows how these markables can be defined at level 2 in a TEI-compliant way.

To make the tabular representation formats shown in Tables 2 and 3 fit into this 3-level architecture, these formats were modified as described in the next two subsections.

4.3.2 The DiAML-TabSW format

DiAML-TabSW is an ideal format for representing the annotation structures defined by the DiAML abstract syntax, inspired by the 3-column format shown in Table 2. Note that the third column in Table 2 represents three things: (1) the speaker, (2) the slash units into which a turn may be subdivided ('B.1 utt1', 'B.2 utt2', etc.), which are in fact already identified by the codings in the first column (which additionally identify the dialogue and the dialogue turn of which a slash unit forms part), and (3) a transcript of what the speaker said (with in-line markups, mostly related to disfluencies). As a first step towards a 'clean' and formally interpretable representation format, these three ingredients were separated by removing marks like 'utt1' and replacing the contents of the first column by functional segment identifiers, and introducing a separate column to represent the speaker. The replacing functional segment identifiers are in fact references to specifications of stretches of the primary

```
<?xml version="1.0" encoding="UTF-8"?>
<TEI xmlns="http://www.tei-c.org/ns/1.0">
  <body / >
    <div><head>The dialogue turns, segmented into words (TEI-compliant)</head>
      <u>
       <w xml:id="w1">right</w>
       <w xml:id="w2">go</w>
       <w xml:id="w3">south</w>
       <w xml:id="w4">and</w>
       <w xml:id="w5">you'll</w>
       <w xml:id="w6">pass</w>
       <w xml:id="w7">some</w>
       <w xml:id="w8">cliffs</w>
       <w xml:id="w9">on</w>
       <w xml:id="w10">your</w>
       <w xml:id="w11">right</w>
      </u>
    </div>
      <div><head>Identification of functional segments</head>
      <spanGrp xml:id="ves1" type="functionalVerbalSegment">
        <span xml:id="ts1" type="textStretch" from="w1" to="w1"/>
      </spanGrp>
      <fs type="functionalSegment" xml:id="fs1"/>
      <f name="verbalComponent" fVal="#ves1"/ >
      <fs/ >
      <spanGrp xml:id="ves2" type="functionalVerbalSegment">
        <span xml:id="ts2" type="textStretch" from="w2" to="w11"/>
      </spanGrp>
      <fs type="functionalSegment" xml:id="fs2">
        <f name="verbalComponent" fVal="#ves2"/>
      </fs>
    </div>
  </body>
</TEI>
```

Fig. 3 TEI-compliant segmentation of primary data.

data in a separate file, for instance as a sequence of word tokens or as a stretch of speech with a given start- and end point. This file corresponds to level (2) in the 3-level architecture, and forms an implementation of stand-off annotation in tabular form. It remedies the limitations of the Table 2 representation of being unable to deal with discontinuous or overlapping functional segments. For example, the discontinuous functional segment sw0105.fs3 in Table 4 is specified in the file sw0105-fs as consisting of the word tokens w12, w13, w14, and w16 (*I, kind, of, I*).

For the sake of readability, the text of a functional segment is represented in an extra column (the fifth column in Table 5); the transcripts of speaker turns were retained as in Table 2, allowing one to see immediately where a functional segment occurs in an utterance. The textual information in the two rightmost columns in Table 5 is strictly speaking redundant, and plays no role in the semantic interpretation of DiAML annotations, but only serves to make the annotation representations more readable.[10] Note that the second column in Table 5 contains the DiAML representation of (1) dimension; (2) communicative function; (3) qualifiers (if any); (4) dependences (if any); and (5) rhetorical relations (if any) - for all the dialogue acts expressed by the functional segment of that row, in the form (17),

[10] See ISO 24617-6 (Principles of semantic annotation, or Bunt (2015) for the use of elements in a concrete representation that have no correspondence to elements in the underlying abstract syntax and semantics.

Markables	*DA-ID*	*Dialogue acts*	*Se*	*FS text*	*Turn transcript*
sw0105-fs.1	da1	Ta:setQuestion	A	Jimmy, so how do you get most of your news?	Jimmy, {D so } how do you get most of your news? /
			B		{D Well, } [I kind of, + {F uh, }I] watch the, national news every day for one. / I also read one or two papers a day / {C and } [I'm a, + I'm pretty much a] news junkie / {C and } I tune in to CNN a lot. /
sw0105-fs.2	da2 da3	TiM:stalling TuM:turnTake	B:	Well,	
sw0105-fs.3	da4	OCM: selfCorrection	B	I kind of, I	
sw0105-fs.4	da5	TiM:stalling	B	uh	
sw0105-fs.5	da6	Ta:answer (da1)	B	I watch the national news every day, for one	
sw0105-fs.6	da7	TiM:stalling	B	uh	
sw0105-fs.7	da8	Ta:answer (da1) {Expansion: xpder da7}	B	I also read on or two papers a day	
sw0105-fs.8	da9	TuM:turnKeep	B	and	
sw0105-fs.9	da10	Ta:inform	B	I'm pretty much a news junkie	
sw0105-fs.10	da11	OCM: selfCorrection	B	I'm a, I'm pretty much a	
sw0105-fs.11	da12	TuM:turnKeep	B	and	
sw0105-fs.12	da13	Ta:answer (da1) {Expansion: xpder da6, da8}	B	I tune in to CNN a lot	

Table 5 ISO 24617-2 annotation of Switchboard SWMD-DA dialogue fragment in Table 3, represented in DiAML-TabSW format. Abbreviations used: Ta = Task, TiM = Time Management, TuM = Turn Management, OCM = Own Communication Management, Se = Sender, FS = functional segment, xpder = expander.

where an asterisk designates zero or more elements of the same type (and 'CF' stands for 'communicative function').

(17) Dimension:CF (dependence:antecedent*) [qualifier]* {rhetorel:antecedent*}

4.3.3 The DiAML-MultiTab Format

The tabular representation format produced by the DitAT tool for DIT annotations, shown in Table 4, was likewise modified in order to be fully ISO-compliant. The identifiers of functional segments in the leftmost column in Table 3 were replaced by references to the functional segment specifications. The 'transcript' column in Table 4 was split into a column containing functional segment texts and one containing turn transcripts.

The Contact Management column in Table 4 was deleted, since this dimension has not been adopted in ISO 24617-2. Like in DiAML-TabSW, the contents of the cells in the dimension columns were enriched to contain complete dialogue act information according to ISO 24617-2, including dependences, qualifiers, and rhetorical relations. The resulting DiAML-MultiTab format is shown in Table 6.

FS	*Se*	*FS text*	*Turn transcription*	Task	Auto-Feedback	Turn Man.	Time Management	Discourse Structuring	SOM
TR3-f1	S	hello	hello, can I help you						da1:Init, Greeting
TR3-fs2	S	can I help you						da2:Offer	
TR3-fs3	U u	uhm	uhm, yes hello, maybe, I'd like to take a tanker from Corning and bring it to Elmira			 da3:Turn	 da4:Stalling		
TR3-fs4	U	yes hello			da5:Auto-Positive (da1)	Take			da6: Ret. Greeting (da1)
TR3-fs5	U	yes maybe						da7: Accept Offer (da2) [uncertain]	
TR3-fs6	U	I'd like to take a tanker from Corning and bring it to Elmira		da8: Inform					
TR3-fs7	S	allright	allright		da9:Auto-Positive (da8)				
TR3-fs8	U	and from Elmira I'd like to load orange juice into the tanker	and from Elmira I'd like to load orange juice into the tanker	da10: Inform {Expansion da8}					
TR3-fs9	S	mm-hm			da11:Auto-Positive (da10)				
TR3-fs10	U	I'd like then to take the anker back to Corning	I'd like then to take the the tanker back to Corning	da12: Inform {Expansion da10}					
TR3-fs11	U	the the					da13:Stalling		

Table 6 ISO 24617-2 annotation of TRAINS dialogue fragment from Table 4 slightly extended and epresented in DiAML-MultiTab format. Abbreviations used: FS = functional segment, Se = Sender, SOM = Social Obligations Management.

4.4 Encodings and Mappings

For encoding the annotation structures of the abstract syntax in a tabular format, note first that functional and feedback dependence relations give rise to nested structures in the abstract syntax. For example, an answer by participant A to a question by participant B about the task domain takes the form of an entity structure with the following schematic form, in which the question that this answer depends on is embedded in the entity structure (for simplicity omitting empty sets of 'other participants' and qualifiers):

$$\langle m, \langle A, B, Task, answer, \langle m, \langle B, A, Task, question \rangle \rangle \rangle \rangle$$

'Flat' representation of annotation structures is made possible by introducing identifiers for dialogue act representations and using these to refer from one dialogue act to another, as for example in Table 5 for dialogue act da6.

A second point to consider is the representation of link structures, corresponding to rhetorical relations between dialogue acts. Having the general form (16), link structures contain two or more entity structures and thus inherit the nested character of the latter. To obtain flat tabular representations of the information in link structures, rhetorical relations are most conveniently represented as a property of the second argument of a relation, using dialogue act identifiers to represent the rhetorically related acts, since this corresponds to the point in te dialogue where the existence of a rhetorical relation typically becomes apparent; see e.g. dialogue act da8 in Table 5.

On this approach, the contents of a cell in the dimension-related columns in a DiAML-MultiTab representation have the form of a string with a structure as in (17) except for the absence of a dimension specification, and preceded by a dialogue act identifier:

(18) dai: CF (dependence:antecedent*) [qualifier]* {rhetorel:antecedent*}

This form can be viewed as a string representation of a pair consisting of an index and a 4-tuple containing (1) a communicative function, (2) a set of dialogue acts with a dependence relation, (3) a set of qualifiers, and (4) zero, one or more sets of rhetorically related dialogue acts with a rhetorical relation, i.e. a structure of the following form:

(19) $\langle i, \langle f_i, \Delta_i, Q_i, R_i \rangle \rangle$

(in which R_i in turn is a set of pairs consisting of a rhetorical relation and a set of rhetorically related dialogue acts, i.e. R_i has the general form $\{\langle R_{i1}, E_{i1} \rangle, ... \langle R_{in}, E_{in} \rangle\}$).

The encoding function $F_{MultiTab}$ that represents annotation structures in the form of DiAML-MultiTab tables can be defined formally in terms of formal operations that turn entity structures and link structures into structures like (19), which are straightforward to represent as strings of the form (18).

The following specification defines a procedure that accomplishes this.

Specification 2. Encoding DiAML annotation structures in DiAML-MultiTab.
For a given annotation structure $A = \{\varepsilon_1, \ldots, \varepsilon_k, L_1, \ldots, L_m\}$:

1. Step 1: introduction of identifiers for entity structures.
   Note that an entity structure in DiAML is a pair $\varepsilon = \langle m, \alpha \rangle$ consisting of a functional segment m (a 'markable') and the semantic information α that characterizes a dialogue act.
   First, sort the entity structures $\varepsilon_1, \ldots, \varepsilon_k$ in A according to their markables, and sort entity structures with the same markable by their dimension, with Task = 1, Auto-Feedback =

2, Allo-Feedback = 3, Turn Management = 4, Time Management = 5, Own Communication Management = 6, Partner Communication Management = 7, Discourse Structuring = 8, Social Obligation Management = 9. Assign to each entity structure ε in A the index $J(\varepsilon)$ corresponding to its position in the resulting ordering.[11]
Output of this step is a sequence $E_A = \{\langle \varepsilon_1, i\rangle, \ldots, \langle \varepsilon_k, n\rangle\}$ of indexed entity structures.

2. Step 2: flattening of entity structures and link structures.
In all the indexed entity structures in E_A replace any embedded entity structures that they may contain by their index.
More precisely, if $\langle \varepsilon_i, j\rangle \in E_A$, with $\varepsilon_i = \langle m_i, \langle S, A, H, f_i, d_i, \Delta_i, Q_i\rangle\rangle$ and $\Delta_i = \{\varepsilon_{i1}, ..., \varepsilon_{ik}\}$ then replace Δ_i by the set of indices $\{J(\varepsilon_{i1}), ..., J(\varepsilon_{ik})\}$.
Likewise, in every link structure $L_i = \langle \varepsilon_i, E_i, \rho_i\rangle \in A$ replace the entity structures in E_i by their index.
3. Step 3: extraction from indexed entity structures of dialogue act information to be represented in cells in a DiAML-MultiTab matrix.
This is a transformation T_e applied to entity indexed structures that reorders the elements in the nested tuples in such a way that the resulting structures contain pairs $\langle i, \langle f_i, \Delta_i, Q_i\rangle\rangle$ which have the desired form (18) except for the inclusion of link structure information. Step 3 will take care of this. The transformation in the present step is:
$T_e(\langle\langle m, \langle S, A, H, d, f, Q, \Delta\rangle\rangle, i\rangle) = \langle m, \langle S, A, H, \langle d, \langle i, \langle f, \Delta, Q\rangle\rangle\rangle\rangle\rangle$
4. Step 4: restructuring the information in link structures to an entity-like form.
A link structure $L = \langle \varepsilon_1, \{\varepsilon_2, ..., \varepsilon_k\}, \rho\rangle$, with $\varepsilon_1 = \langle m_1, \langle S_1, A_1, H_1, d_1, f_1, Q_1, \Delta_1\rangle\rangle$ is transformed as follows:
$T_L(L) = \langle m_1, \langle S_1, A_1, H_1, \langle d_1, \langle i_1, \langle f_1, \Delta_1, Q_1, \langle \rho, \{i_2, .., i_k\}\rangle\rangle\rangle\rangle\rangle\rangle$
where $i_2, .., i_k$ are the indices of the structures $\{\langle \varepsilon_2, i_2\rangle, ..., \langle \varepsilon_k, i_k\rangle\}$, built in step 1.
The structures built in this step are copies of structures built in step 3, extended with information from rhetorical links. The next step merges the structures with and without rhetorical link information, thereby removing duplicate information.
5. Step 5: merge of structures built in the previous two steps with and without rhetorical link information.
This merge operation succeeds only if both arguments are identical except that one of them has additional rhetorical link information specified. It is formally defined as:
$\langle m, \langle S, A, H, \langle d, \langle i, \langle f, \Delta, Q\rangle\rangle\rangle\rangle\rangle \cup \langle m, \langle S, A, H, \langle d, \langle i, \langle f, \Delta, Q, \langle \rho, \{i_1, .., i_k\}\rangle\rangle\rangle\rangle\rangle\rangle =$
$\langle m, \langle S, A, H, \langle d, \langle i, \langle f, \Delta, Q, \langle \rho, \{i_1, .., i_k\}\rangle\rangle\rangle\rangle\rangle\rangle$.
6. Step 6: combination of structures constructed so far that have the same markable and the same sender, and will be represented in the same row in DiAML-MultiTab.
This operation is defined as:
$\langle m, \langle S, A, H, \langle d_1, \langle i, \alpha\rangle\rangle\rangle\rangle + \langle m, \langle S, A, H, \langle d_2, \langle j, \beta\rangle\rangle\rangle\rangle =$
$\langle m, \langle S, A, H, \{\langle d_1, \langle i, \alpha\rangle\rangle, \langle d_2, \langle j, \beta\}\rangle\rangle\rangle$
The steps 1-6 produce a set of structures of the form $\langle m, \langle S, A, H, \{\langle d_1, a_1\rangle, ...\langle d_k, a_k\rangle\}\rangle\rangle$, which correspond to rows in a DiAML-XML matrix, where a_j is a tuple $\langle i, \langle f, \Delta, Q, R_{rh}\rangle\rangle$ that corresponds to the content of a cell in of one of the nine dimension-related columns in the matrix (where Δ and Q may be empty sets and are conveniently suppressed in representations). A collection of structures, ordered by their first component (the markables), forms a DiAML-MultiTab matrix.

[11] According to the metamodel shown in Fig. 7, a functional segment has at most one communicative function in a given dimension, therefore this is guaranteed to produce a unique index.

What remains to be done in order to arrive at DiAML-MultiTab representations is to represent the cell contents by strings and order the columns by their dimension ordering (see Step 1). Moreover, for convenience, column headers are introduced and two columns are added that represent the textual content of functional segments and turns, but this is formally redundant and not part of the $F_{MultiTab}$ function.

7. Step 7: representation of cell contents in string form.
   This is a rather trivial final step, where sequenceses of the form $\langle i, \langle f_i, \Delta_i, Q_i, R_i \rangle \rangle$ are replaced by strings of the form:
   "da$_i$:"$\oplus F_{MultiTab}(f_i) \oplus$"("$\oplus F_{MultiTab}(\Delta_i) \oplus$")"$\oplus$"["$\oplus F_{MultiTab}(Q_i) \oplus$"]"$\oplus \{F_{MultiTab}(R_i)\}$
   in which $\oplus$ designates string catenation; $F_{MultiTab}(f_i)$ is the DiAML-MultiTab name of the comunicative function f_i, etc.

The DiAML-TabSW representation format differs from the DiAML-MultiTab format only in that all dialogue acts are represented in a single column with an indication of their dimension.The DiAML-TabSW encoding of annotation structures (the function F_{TabSW}) is therefore defined in a similar way. The definition of the encoding functions $F_{MultiTab}$ and F_{TabSW} demonstrates the completeness of the two formats; their unambiguity can likewise be shown by specifying the reverse functions. Together with the corresponding functions for the DiAML-XML format, this demonstrates the interoperability of the three representation formats.

4.5 Advantages of Alternative Representation Formats

The DiAML-XML representation format was originally motivated by its compactness, relative to full-out standard XML. When developing the DialogBank, starting from previously annotated dialogues, adjustments had to be made in segmentation, annotation, and representation format. The DiAML-TabSW and DiAML-MultiTab formats were defined for simplifying the reformatting of SWBD-DA dialogues and DIAMOND dialogues, and allowing easy comparison between original and new or revised annotations. For example, comparing Table 3 and Table 5 one immediatly sees the more fine-grained segmentation used in ISO 24617-2 annotation then in Switchboard-DAMSL annotation, and the more detailed characterization of dialogue acts with dimension information, functional and feeedback dependences, and rhetorical relations. Likewise, comparing Table 4 and Table 6, one directly notices again the more fine-grained use of functional segments and the richer information about dialogue acts.

Moreover, besides making adjustments it was also necessary to check and correct annotations in many cases, in order to obtain a collection of 'gold standard' annotated dialogues. Detection and correction of errors and detection of omissions was found to be much easier for annotations represented in one of the tabular formats than for those represented in the XML-based format. User-based evaluation has shown the usability of both tabular DiAML formats, for trained as well as for untrained annotators (Wijnnhoven, 2016).

The interoperability of the three DiAML representation formats has been exploited by implementing conversions between any two of the three DiAML representation formats, using their common underlying abstract syntax as an interlingua, in a Python program that works both on MS Windows and Apple platforms (Wijnhoven, 2016), and is available from the DialogBank.[12]) This allows users to view (and to produce) ISO 24617-2 annotations in

[12] See `https://dialogbank.uvt.nl/representation-formats/`.

the representation format that they find most convenient, given the annotation and viewing tools at hand.

5 ISO 24617-2 Limitations and Extensions

In building the DialogBank two limitations were discovered of the ISO 24617-2 annotation scheme. One of these is due to an oversight in designing the scheme; the other is due to the scope that was chosen for the standard.

5.1 Annotating Feedback Dependence Relations

Feedback acts are about the processing of something that was said before. The nature of this 'something' depends on the kind of feedback. Feedback by means of expressions like *"OK"*, *"Uh-huh"*, or *"Really?"* is about one or more previous dialogue acts, while feedback by means of *"Tuesday?"* or *"What did you say?"* is about a previous utterance segment, rather than about a dialogue act. The ISO 24617-2 annotation scheme therefore allows both dialogue acts and functional segments as antecedents for feedback dependence relations.

The ISO scheme is in fact not quite correct at this point, since segment-related feedback is not necessarily about a *functional* segment; it may be about *any* previous segment, functional or not, such as a single word or a sequence of words within a functional segment. In the latter case the ISO scheme only allows annotating a feedback dependence relation to the functional segment containing the expression that the feedback act refers to. The DBOX dialogues in the DialogBank deviate in this respect from the ISO standard, since for feedback dependences non-functional segments were introduced. In a future revision, the ISO standard should include such a possibility to allow more accurate markup of feedback dependences.

5.2 Annotating Rhetorical Relations

ISO 24617-2 does not require the marking up of rhetorical relations, such as Cause, Contrast, or Elaboration, and does not specify any particular set of relations that could be used; it only specifies *how* a rhetorical relation between two dialogue acts can be marked up.

In 2016 ISO standard 24617-8 has been established for the annotation of rhetorical relations in discourse. This standard, also called "DR-Core", defines a set of 18 'core' relations that are shared by many annotation schemes for rhetorical relations, and this set has been used for marking up rhetorical relations between dialogue acts in most of the dialogues in the DialogBank. However, when annotating rhetorical relations between dialogue acts two problems were noted: (1) the lack of a possibility to mark up argument roles, and (2) the impossibility to distinguish between a rhetorical relation that links two dialogue acts and one that links the semantic content of two dialogue acts (or mixed cases). These problems are discussed in the rest of this section.

5.2.1 Argument roles

Rhetorical relations are commonly assumed to have two arguments, for example, a Cause relation has two arguments, one called 'Reason' and one called 'Result'. In order to allow annotions to have a well-defined semantics, DR-Core requires the argument roles to be

marked up, as in (20), where the event of John pushing Jim is marked up as being a reason for the event of Jim falling on the ground.

(20) John pushed Tim. He fell on the ground.

```
<drArg xml:id="a1" target="#s1" type="event"/>
<dRel xml:id="r1" rel="cause"/>
<drArg xml:id="a2" target="#s2" type="event"/>
<drLink rel="r1" reason="#a1" result="#a2"/>
```

ISO 24617-2, by contrast, provides just a single slot for specifying a rhetorical relation, and has no provisions for marking up argument roles, as illustrated in (21), where the 'rhetoricalLink' element indicates the occurrence of a causal relation between the *Inform* act expressed by *"he has the flu"* and the answer *"He didn't come in"*, but this does not allow the possibility to indicate that the information in the *Inform* act is the reason in the causal relation, rather than the result.

(21) A: Have you seen Pete today?
B: He didn't come in; he has the flu.

```
<dialogueAct xml:id="da1" target="#fs1" sender="#a" addressee="#b"
  dimension="task" communicativeFunction="propositionalQuestion" />
<dialogueAct xml:id="da2" target="#fs2" sender="#b" addressee="#a"
  dimension="task" communicativeFunction="answer"
  functionalDependence="#da1"/>
<dialogueAct xml:id="da3" target="#fs3" sender="#b" addressee="#a"
  dimension="task" communicativeFunction="inform" />
<rhetoricalLink dact="#da3" rhetoAntecedent="#da2" rhetoRel="cause"/>
```

In some of the annotations in the DialogBank this limitation has been addressed by marking up a relation plus an argument role in strings of the form 'Cause:Reason', which, from a semantic point of view, is not an adequate solution since the underlying abstract syntax and semantics only include rhetorical relations, no argument roles.

5.2.2 Semantic and Pragmatic Rhetorical Relations

Another limitation of the annotation of rhetorical relations in ISO 24617-2 is that it is not possible to distinguish between so-called 'semantic' and 'pragmatic' interpretations of such relations. Example (22) illustrates this distinction:

(22) a. A: Have you seen Pete today?
B: He didn't come in. He has the flu.
b. A: Have you seen Pete today?
B: He didn't come in. He sent me a message saying that he has the flu.

B's utterances in (22a) are causally related in the sense that the semantic content of the second utterance expresses the reason why the content of the first utterance is true. In (22b), by contrast, there is a 'pragmatic' causal relation in the sense that the second utterance expresses the reason why *B says that Pete is not in* - in this case B's utterance is causally related to the dialogue act of informing A that Pete is not in today, rather than to the content of this dialogue act.

In the DR-Core annotation scheme this distinction is represented by indicating the types of the arguments, where 'dialogue act' is one of the possible types, and the type of the

semantic content of a dialogue act (e.g. event or state) is another. This is illustrated in example (23), which shows the annotation of the examples in (22) represented in the markup language of DR-Core, DRelML (Discourse Relations Markup Language).

(23) a. <drArg xml:id="a1" target="#fs2" type="event"/>
<dRel xml:id="r1" rel="cause"/>
<drArg xml:id="a2" target="#fs3" type="state"/>
<drLink rel="r1" result="#a1" reason="#a2"/>

b. <drArg xml:id="a1" target="#fs2" type="dialogueAct"/>
<dRel xml:id="r1" rel="cause"/>
<drArg xml:id="a2" target="#fs3" type="event"/>
<drLink rel="r1" result="#a1" reason="#a2"/>

In both (23a) and (23b) an implicit Cause relation is marked up between the arguments expressed by the markables fs2 (*"Pete did not come in today"*) and fs3 (*"He has the flu".*; *"He sent me a message saying that he has the flu"*, respectively), but in the former case the first argument is the event of Pete not coming in which is caused by the second argument, while in the latter case it is the dialogue act of B informing A that Pete did not come in which is caused by the second argument. This distinction cannot be expressed in DiAML. In DRelML, on the other hand, no information about the arguments of a rhetorical relation can be represented other that their semantic types. For marking up rhetorical relations between dialogue acts it would thus seem attractive to somehow combine DiAML and DRelML. This is discussed in the next subsection.

5.3 Combinations of Annotation Schemes

It was noted in Section 3.2 that DiAML-XML is in fact a compact way of using XML, as illustrated by (8b) and (9).

Likewise, a DRelML annotation of a rhetorical relation like the one in (24a) is a compact form of the full XML expression in (24b):

(24) a. He didn't come in. He has the flu.
<drArg xml:id="e1" target="#s1" type="event"/>
<drArg xml:id="e2" target="#s2" type="event"/>
<drLink rel="cause" reason="#e2" result="#e1"/>

b. <fs xml:id="e1">
<f name="target"><value="#s1"/></f>
<f name="type"><value="event"/></f>
</fs>
<fs xml:id="e2">
<f name="target"><value="#s2"/></f>
<f name="type"><value="event"/></f>
</fs>
<fs xml:id="r1">
<f name="rel"><value="cause"/></f>
<f name="reason"><value="#e2"/></f>
<f name="result"><value="#e1"/></f>
</fs>

Since the concatenation of two XML-expressions is again a legitimate XML-expression, we may combine the relevant bits of a DiAML annotation of dialogue acts and a DRelML annotation of rhetorical relations. Applied to B's utterances in the example (22b) this would lead to the representation shown in (25b) and in compact form in (25c).

(25) a. (A: Have you seen Pete today?)
B: He didn't come in. He sent me a message saying that he has the flu.

b.
```
<fs xml:id="da2">
  <f name="target"><value="#s2"/></f>
  <f name="sender"><value="#b"/></f>
  <f name="addressee"><value="#a"/></f>
  <f name="dimension"><value="task"/></f>
  <f name="communicativeFunction"><value="answer/></f>
  <f name="functionalDependence"><value="#da1/></f>
</fs>
<fs xml:id="da3">
  <f name="target"><value="#s3"/></f>
  <f name="sender"><value="#b"/></f>
  <f name="addressee"><value="#a"/></f>
  <f name="dimension"><value="task"/></f>
  <f name="communicativeFunction"><value="inform/></f>
<fs xml:id="e2">
  <f name="target"><value="#s2"/></f>
  <f name="type"><value="event"/></f>
</fs>
<fs xml:id="e3">
  <f name="target"><value="#s3"/></f>
  <f name="type"><value="event"/></f>
</fs>
<fs xml:id="r1">
  <f name="rel"><value="cause"/></f>
  <f name="reason"><value="#e3"/></f>
  <f name="result"><value="#e2"/></f>
</fs>
```

c.
```
<dialogueAct xml:id="da2" target="#s2" sender="#b" addressee="#a"
  dimension="task" communicativeFunction="answer"
  functionalDependence="#da1">
 <dialogueAct xml:id="da3" target="#s3" sender="#b" addressee="#a"
  dimension="task" communicativeFunction="inform"/>
 <drArg xml:id="e3" target="#s3" type="event" />
 <drLink rel="cause" reason="#e3" result="#da2" />
```

This representation, either in full or in compact form, is not quite satisfactory, since it contains two different annotations of the same segment, namely the segment s3: one that views the segment as a dialogue act and one that views it as an event. Both views are justifiable, but it is strange to have no relation between the two views; this makes the semantic interpretation of such annotations problematic.

The missing relation here is that of semantic content: the event view represented in the <drArg> element concerns the semantic content of the dialogue act da3, so the way to resolve this would be to introduce an XML attribute 'semanticContent' for the <dialogueAct> element whose value is the event in question.

Introducing information about the semantic content of dialogue acts, not just in the XML-based representations but also in the underlying abstract syntax and semantics, would

open up interesting possibilities of combining dialogue act annotation with semantic information addressed by other annotation schemes, in particular by ISO 24617-1 ('ISO-TimeML') and ISO 24617-7 (ISOspace) for the annotation of events and their temporal and spatial properties, by ISO 24617-4 (Semantic Roles) for adding information about the participants in an event, and in the future also for adding information about quantification over events and their participants (see Bunt et al., 2018).

6 Conclusions and Future Work

The DialogBank had its first public release in December 2015. It contains at the time of writing annotated dialogues with the properties shown in Table 1. Material from English-language dialogue corpora (HCRC Map Task, Switchboard, TRAINS) and from Dutch-language corpora (DIAMOND, OVIS, Schiphol, Dutch Map Task) were re-segmented and re-annotated according to ISO 24617-2. To facilitate comparisons between original and final segmentation and annotation, as well as in support of the detection and correction of errors and omissions, two tabular representation formats were defined that were shown to be ideal (complete and unambiguous) and hence interoperable with the reference DiAML-XML format of the ISO 24617-2 standard. The interoperability was exploited by implementing conversions between the three representation formats, allowing users of the DialogBank to view or manipulate or otherwise utilise the annotated dialogues in their preferred format.

Building the DialogBank made us aware of an incorrectness in ISO 24617-2 for annotating the 'antecedent' of feedback acts that refer to a non-functional stretch of dialogue - an incorrectness that also precludes an accurate annotation of speech editing acts (acts in the Own Communication Management or in the Partner Communication Management dimension). This issue should be taken into account when a revision of the ISO 24617-2 standard is considered (see Bunt et al., 2017a).

Another lesson learned from building the DialogBank concerns the annotation of rhetorical relations between dialogue acts. In ISO 24617-2 this is just an option; it is not obligatory. But from a semantic point of view, the rhetorical relations that link the dialogue acts in a spoken conversation are extremely important, so it would be interesting to integrate dialogue act annotation with the annotation of rhetorical relations, which is the subject matter of ISO standard 24617-8. Integration with the annotation of other semantic information seems the next, interesting step to obtain more valuable annotations, which allow for example the extraction of more useful information from annotated interactive discourse.

Such developments are expected to lead to enriched annotated material in the DialogBank, which will in the near future also be extended with more annotated dialogues from various other corpora and involving other languages besides English and Dutch.

References

Alexandersson J, Buschbeck-Wolf B, Fujinami T, Kipp M, Koch S, Maier E, Reithinger N, Siegel BSM (1998) Dialogue acts in VERBMOBIL-2 (second edition). Verbmobil Report 226. DFKI, Saarbrücken.

Allen J, Core M (1997) DAMSL: Dialogue Act Markup in Several Layers (Draft 2.1). Technical Report. University of Rochester, Rochester, NY, URL `http://www.cs.rochester.edu/research/cisd/resources/damsl/RevisedManual/`

Allen J, Schubert L, Ferguson G, Heeman P, Hwang C, Kato T, Light M, Martin N, Miller B, Poesio M, Traum D (1994) The TRAINS project: A case study in defining a conversational planning agent. Technical Report 532. Department of Computer Science, University of Rochester, Rochester, NY.

Allwood J (1992) On dialogue cohesion. Gothenburg University, Department of Linguistics

Anderson A, Bader M, Bard E, Boyle E, Doherty G, Garrod S, Isard S, Kowtko J, McAllister J, Miller J, Sotillo F, Thompson H, Weinert R (1991) The HCRC Map Task Corpus. Language and Speech 34:351–366.

Bunt H (1994) Context and dialogue control. Think Quarterly 3(1):19–31

Bunt H (2000) Dialogue pragmatics and context specification. In: Bunt H, Black W (eds) Abduction, Belief and Context in Dialogue. Studies in Computational Pragmatics, John Benjamins, Amsterdam, pp 81–150.

Bunt H (2009) The DIT^{++} taxonomy for functional for dialogue markup. In: Heylen D, Pelachaud C, Catizone R, Traum D (eds) Proceedings of EDAML-AAMAS Workshop "Towards a Standard Markup Language for Embodied Dialogue Acts", Budapest, pp 36–48

Bunt H (2010) A methodology for designing semantic annotation languages exploring semantic-syntactic ISO-morphisms. In: Proceedings of the Second International Conference on Global Interoperability for Language Resources (ICGL 2010), Hong Kong: City University, pp 29–46

Bunt H (2011) Multifunctionality in dialogue. Computer, Speech and Language 25:222–245

Bunt H (2015) On the Principles of Semantic Annotation. In: Proceedings 11th Joint ACL-ISO Workshop on Interoperable Semantic Annotation (ISA-11), London, pp 1–13

Bunt H, Prasad R (2016) ISO DR-Core: Core concepts for the annotation of discourse relations. In: Proceedings 12th Joint ACL-ISO Workshop on Interoperable Semantic Annotation (ISA-12), Portoroz, Slovenia, pp 45–54

Bunt H, Alexandersson J, Carletta J, Choe JW, Fang A, Hasida K, Lee K, Petukhova V, Popescu-Belis A, Romary L, Soria C, Traum D (2010) Towards and ISO standard for dialogue act annotation. In: Proceedings 7th International Conference on Language Resources and Evaluation (LREC 2010), Malta, ELRA, Paris

Bunt H, Alexandersson J, Choe JW, Fang A, Hasida K, Petukhova V, Popescu-Belis A, Traum D (2012) ISO 24617-2: A semantically-based standard for dialogue annotation. In: Proceedings 8th International Conference on Language Resources and Evaluation (LREC 2012), ELRA, Istanbul

Bunt H, Petukhova V, Fang A (2017a) Revisiting the ISO Standard for Dialogue Act Annotation. In: Proceedings 13th Joint ISO - ACL Workshop on Interoperable Semantic Annotation (ISA-13), Montpellier, France, September 2017, pp 37–50

Bunt H, Petukhova V, Traum D, Alexandersson J (2017b) Dialogue Act Annotation with the ISO 24617-2 Standard. In: Dahl D (ed) Multimodal Interaction with W3C Standards, Springer, Berlin, pp 109–135

Bunt H, Pustejovsky J, Lee K (2018) Towards an ISO Standard for the Annotation of Quantification. In: Proceedings of the 11th International Conference on Language Resources and Evaluation (LREC 2018), Miyazaki, Japan

Burkhardt F, Pelachaud C, Schuller B, Zovato E (2017) EmotionML. In: Dahl D (ed) Multimodal Interaction with W3C Standards, Springer, Berlin, pp 65–80

Carletta J, Isard S, Kowtko J, Doherty-Sneddon G (1996) HCRC dialogue structure coding manual. University of Edinburgh, technical Report HCRC/TR-82

Caspers J (2000ba) Melodic characteristics of backchannels in Dutch Map Task dialogues. In: Proceedings Sixth International Conference on Spoken Language Processing, ICSLP

2000 and INTERSPEECH 2000, Beijing, pp 611–614.
Caspers J (2000bb) Pitch accents, boundary tones and turn-taking in Dutch Map Task dialogues. In: Proceedings Sixth International Conference on Spoken Language Processing, ICSLP 2000 and INTERSPEECH 2000, Beijing, pp 565–568.
Geertzen J (2007) DitAT: A flexible tool to support web-based dialogue annotation. In: Proceedings 7th International Conference on Computational Semantics (IWCS-7), Tilburg, pp 320–323
Geertzen J, Girard Y, Morante R (2004) The DIAMOND project. In: Proceedings 8th Workshop on the Semantics and Pragmatics of Dialogue (CATALOG 2004), Barcelona
ISO (2006) ISO 24610:2006, Language Resource Management: Feature structures. International Standard. International Organisation for Standardisation ISO, Geneva.
ISO (2009) 24612:2009, Language Resource Management - Linguistic Annotation Framework (LAF). International Organisation for Standardisation ISO, Geneva
ISO (2012a) ISO 24617-1: 2012, Language Resource Management - Semantic Annotation Framework (SemAF) - Part 1: Time and events. International Organisation for Standardisation ISO, Geneva
ISO (2012b) ISO 24617-2:2012, Language Resource Management - Semantic Annotation Framework (SemAF) - Part 2: Dialogue acts. Geneva: International Organisation for Standardisation ISO
ISO (2014a) ISO 24617-4: 2014, Language Resource Management - Semantic Annotation Framework (SemAF) - Part 4: Semantic roles. Geneva: International Organisation for Standardisation ISO
ISO (2014b) ISO 24617-7: 2014, Language Resource Management - Semantic Annotation Framework (SemAF) - Part 7: Spatial information. International Organisation for Standardisation ISO, Geneva
ISO (2015) ISO 24617-6:2015, Language Resource Management - Semantic Annotation Framework (SemAF) - Part 6: Principles of semantic annotation. International Organisation for Standardisation ISO, Geneva
ISO (2016) ISO 24617-8:2016, Language Resource Management - Semantic Annotation Framework (SemAF) - Part 8: Semantic relations in discourse, Core annotation scheme (DR-Core). International Organisation for Standardisation ISO, Geneva
Jurafsky D, Shriberg E, Biasca D (1997) Switchboard SWBD-DAMSL Shallow-Discourse-Function Annotation: Coders Manual, Draft 1.3. University of Colorado
Kipp M (2001) ANVIL - a generic annotation tool for multimodal dialogue. In: Proceedings 7^{th} European Conference on Speech Communication and Technology (Eurospeech), Aalborg, pp 1367–1370
Kipp M (2014) ANVIL: The video annotation research tool. In: Durand J, Gut U, Kristoffersen G (eds) Handbook of Corpus Phonology, Oxford University Press, pp 420–436
Ngo TL, Pham KL, Takeda H (2018) A Vietnamese dialogue act corpus based on the ISO 24617-2 Standard. In: Proceedings 11th Intyernational Conference on Language Resources and Evaluation(LREC 2018), Miyazaki, Japan, ELRA, Paris
Petukhova V, Prévot L, Bunt H (2011) Multi-level discourse relations between dialogue units. In: Proceedings 6th Joint ACL-ISO workshop on Interoperable Semantic Annotation (ISA-6), Oxford, pp 18–28
Petukhova V, Gropp M, Klakow D, Eigner G, Topf M, Srb S, Moticek P, Potard B, Dines J, Deroo O, Egeter R, Meinz U, Liersch S, Schmidt A (2014) The DBOX corpus collection of spoken human-human and human-machine dialogues. In: Proceedings 9^{th} International Conference on Language Resources and Evaluation (LREC 2014)Reykjavik, Iceland, ELRA, Paris

Petukhova V, Stephens C, De Weerd H, Taatgen N, Cnossen F, Malchanau A (2016) Modelling multi-issue bargaining dialogues: Data collection, annotation design and corpus. In: Proceedings 9th International Conference on Language Resources and Evaluation (LREC 2016), Portoroz, Slovenia, pp 3133–3140

Popescu-Belis A (2005) Dialogue Acts: One or More Dimensions? ISSCO Working Paper 62. ISSCO, Geneva, URL `http://www.issco.unige.ch/publicaitons/working-papers/papers/apb-issco-wp62b.pdf`

Prüst H, Minnen G, Beun RJ (1984) Transcriptie dialooogesperiment juni/juli 1984, IPO Rapport 481. Institute for Perception Research, Eindhoven University of Technology

Shriberg E, Dhillon R, Bhagat S, Ang J, Carvey H (2004) The ICSI meeting recorder dialogue act (MRDA) corpus. In: Proceedings 5th SIGDIAL Workshop on Discourse and Dialogue (SIGDIAL 2004, pp 97–100.

Stent A (2000) The Monroe Corpus, Technical Report 728 and Technical Note 99-2. Computer Science Department, University of Rochester, March 2000.

Traum D (2000) 20 questions on dialogue act taxonomies. Journal of Semantics 17(1):7–30.

Wijnhoven K (2016) Annotation Representations and the Construction of the DialogBank. MA Thesis, Tilburg University.