

metalogue

Metalogue Multi-Issue Bargaining: Corpus Manual

The METALOGUE Consortium

October 2016

Main Authors: Volha Petukhova
Andrei Malchanau
Youssef Oualil
Dietrich Klakow
Christopher Stevens
Harmen de Weerd
Niels Taatgen



**Project funded by the European Community under the
Seventh Framework Programme for
Research and Technological Development**

Project ref. no.	ICT – 611073
Project title	METALOGUE–Multiperspective Multimodal Dialogue: dialogue system with metacognitive abilities

Document title	Metalogue Multi-Issue Bargaining: Corpus Manual
Number of pages	63 (excluding front matter)
Author(s)	Volha Petukhova, Andrei Malchanau, Youssef Oualil, Dietrich Klakow, Christopher Stevens, Harmen de Weerd, Niels Taatgen
Keywords	dialogue corpus, automatic speech recognition, speaker diarization, speech transcriptions, semantic annotation, dialogue act annotation, ISO 24617-2, negotiation moves

Table of Contents

1. EXECUTIVE SUMMARY	5
2. INTRODUCTION.....	6
3. CORPUS STRUCTURE	6
3.1 DATA	7
3.1.1 Speech.....	7
3.1.2 Transcriptions.....	7
4. DIALOGUE ACT ANNOTATION: BASIC CONCEPTS, DEFINITIONS AND METAMODEL.....	11
4.1 DIALOGUE ACT.....	11
4.2 DIALOGUE PARTICIPANTS	12
4.3 DIALOGUE SEGMENTATION: DIALOGUE UNITS	12
4.4 RELATIONS BETWEEN DIALOGUE UNITS	14
4.5 COMMUNICATIVE FUNCTION QUALIFIERS	15
4.6 DIALOGUE ACT ANNOTATION METAMODEL.....	16
4.7 ISO 24617-2 DIALOGUE ACT ANNOTATION SCHEME	17
4.8 ANNOTATION GUIDELINES.....	20
4.9 PROPOSED EXTENSIONS FOR METALOGUE	20
5. METALOGUE DOMAIN AND TASKS: NEGOTIATION STRUCTURE AND MOVES	21
6. ANNOTATION FORMAT: DIALOGUE ACT MARKUP LANGUAGE (DIAML).....	24
7. ANVIL ANNOTATION TOOL.....	28
7.1 WORKFLOW.....	29
8. CONCLUSIONS AND FUTURE EXTENSIONS	29
9. GLOSSARY	34
10. ANNEXES	38
10.1 ANNEX 1. ISO 24617-2 COMMUNICATIVE FUNCTIONS DEFINITIONS	38
10.1.1 General-purpose functions: information-seeking functions.....	38
10.1.2 General-purpose functions: information-providing functions	39
10.1.3 General-purpose functions: commissives	40
10.1.4 General-purpose functions: directives.....	42
10.1.5 Feedback functions	43
10.1.6 Turn management functions	44

10.1.7 Time management functions	45
10.1.8 Own and partner communication management functions.....	46
10.1.9 Discourse structuring functions	46
10.1.10 Social obligations management functions	47
10.1.11 Contact management functions.....	49
10.1.12 Other functions	49
10.2 ANNEX 2. RHETORICAL (DISCOURSE) RELATIONS	50
10.3 ANNEX 3. NEGOTIATION MOVES	52
10.4 ANNEX 4. SPEECH PRODUCTION DISFLUENCY TYPES	54
10.5 ANNEX 5. GUIDELINES ON USING ANVIL ANNOTATION TOOL.....	55
10.5.1 Example of Metalogue specification for dialogue act annotations using Anvil annotation tool.....	58
10.6 ANNEX 6. TRANSCRIPTION FORMAT SPECIFICATIONS	59
10.6.1 Plain text formats	59
10.6.2 XML-based format: ANVIL and Text Encoding Initiative format (TEI)	60

1 Executive Summary

This manual provides an overview of the semantically annotated Metalogue¹ Multi-Issue Bargaining (MIB) corpus published in LDC Catalog. The manual describes data collection procedures, annotation activities including an overview of basic concepts and their definitions (annotation schemes) with guidelines how to apply them, and used tools.

The manual is organized as follows. The Metalogue annotated corpus structure is presented in Section 3. Section 4 presents linguistic annotation in general and dialogue act annotation in particular, outlines the main Metalogue dialogue act annotation elements and processes, as well as defines basic annotation concepts and summarizes them in a UML-based metamodel. The ISO 24617-2 dialogue act annotation scheme is discussed, as well as its extension and adaptation for the Metalogue purposes. Section 5 specifies Metalogue domain and tasks related to negotiation dialogue structure and negotiation moves. In Section 6 we present the ANVIL annotation tool. Additionally, the manual contains the glossary of terms and definitions as well as Annexes with annotated concepts definitions and examples, annotation guidelines, tools workflows and formats examples.

¹www.metalogue.eu

2 Introduction

The manual describes the semantically annotated dialogue data collection undertaken within the Metalogue project. The project aimed at developing a dialogue system with flexible dialogue management to enable the system's adaptive, reactive, interactive and proactive dialogue behaviour in setting goals, choosing appropriate strategies and monitoring numerous parallel interpretation and management processes. To achieve these goals a negotiation (or more precisely multi-issue bargaining) scenario has been developed as the specific setting and application domain. The dialogue corpus formed the basis for the design of task and interaction models of participants' negotiation behaviour, and subsequently for dialogue system development which is capable of replacing one of the negotiators. The recordings were carried out at the University of Groningen, and transcriptions and annotations were carried out at the Spoken Language System Lab at Saarland University.

The corpus consists of 24 dialogues of a total duration of appr. 2.5 hours comprising about 2.000 speaking turns, 3.650 functional segments and about 5.000 unique tokens. It should be noticed that the corpus contains a significant proportion of non-native English speakers, varying in fluency from nearly-native to challenging-to-transcribe.

The published corpus is delivered with audio recordings, automatic and manual transcriptions, ISO 24617-2 dialogue act annotations and other semantic annotations specific to the Metalogue project. In total, seven types of annotations are performed manually by two trained annotators and corrected by one expert annotator:

1. ISO 24617-2 dialogue act annotations ISO (2012)²
2. Discourse structure acts according to DIT⁺⁺, see <http://dit.uvt.nl/>
3. Contact Management acts according to DIT⁺⁺, see <http://dit.uvt.nl/>
4. Task Management dialogue acts for managing negotiation task introduced in Metalogue, see Petukhova et al. (2016a)
5. Negotiation moves as defined in Petukhova et al. (2016b)
6. Rhetorical relations with ISO 24617-8 discourse relations set Bunt and Prasad (2016)
7. Disfluencies in speech production as defined in Besser (2006)

Annotations were performed using the Anvil tool³. Annotation files are provided in Anvil-specific XML format and ISO 24617-2 ISO Dialogue Act Markup Language (DiAML), see Bunt et al. (2010).

The corpus is well documented, all guidelines, reports, annotation schemes as well as data collection instructional material are provided which makes it possible to easily replicate the carried out experiments.

3 Corpus Structure

The delivered corpus has the folder structure depicted in Figure 7. There are two main folders containing delivered Data and Documentations describing the collection and annotation processes.

²See also http://dit.uvt.nl/#iso_24617-2

³<http://www.anvil-software.org/>

3.1 Data

For dialogue data collection, the specific setting considered involves a multi-issue bargaining scenario in which a representative of a city council and a representative of small business owners negotiate over the implementation of new anti-smoking regulations. The negotiation involves four issues, each with four or five different options (see Figure 2). The task of the negotiators is to negotiate an agreement, which assigns exactly one option to each issue.

Each experiment involves a pair of participants that perform a number of separate negotiation scenarios. One of the participant is randomly assigned the role of the city council, the other participant to the role of the business representative. Each participant receives their background story and instructions, as well as their preference profiles for each scenario. For each preference profile, each option was assigned one of nine possible values, which was communicated to the participant through colours. Brighter orange colours indicated increasingly negative options, while brighter blue colours - increasingly more positive options. The use of colour rather than numbers introduces a form of uncertainty in the exact value of a given agreement, which is closer to real-life negotiations. Figure 2 illustrates participants' instructions and preference cards defined for nine different cases. Preference cards with instructions are provided with the corpus in `html` format in the Documentation folder and can be used to replicate the data collection experiments.

Participants negotiated for an agreement with the highest possible value according to their preference information. They were not allowed to accept agreements that had a negative value, and participants were not allowed to show their preference information to each other. No further rules on the negotiation process were imposed.

3.1.1 Speech

In the data collection experiments, the conversational speech was captured with two headset microphones (Sennheiser PC 3 headsets) and saved in FLAC format. 6 unique subjects, undergraduates of age between 19 and 25 participated in these experiments. Speech signal files are of two types: (1) full dialogue session, and (2) segmented speech signal cut per speaker and roughly per turn (process known as *speaker-diarization*). Speaker diarization has been carried out manually using the Audacity free open source software⁴. Recordings are performed in the following setting:

- sample rate: 16000Hz
- sample size: 16-bit
- sample format: linear PCM
- channel: mono

3.1.2 Transcriptions

Participants' speech has been transcribed semi-automatically by (1) running the Automatic Speech Recognizer (ASR) Kaldi Povey (2011) and (2) correcting automatic transcriptions manually. Corrected transcriptions can be used to (re-)train an ASR system.

⁴<http://www.audacityteam.org/>

Metalogue MIB corpus

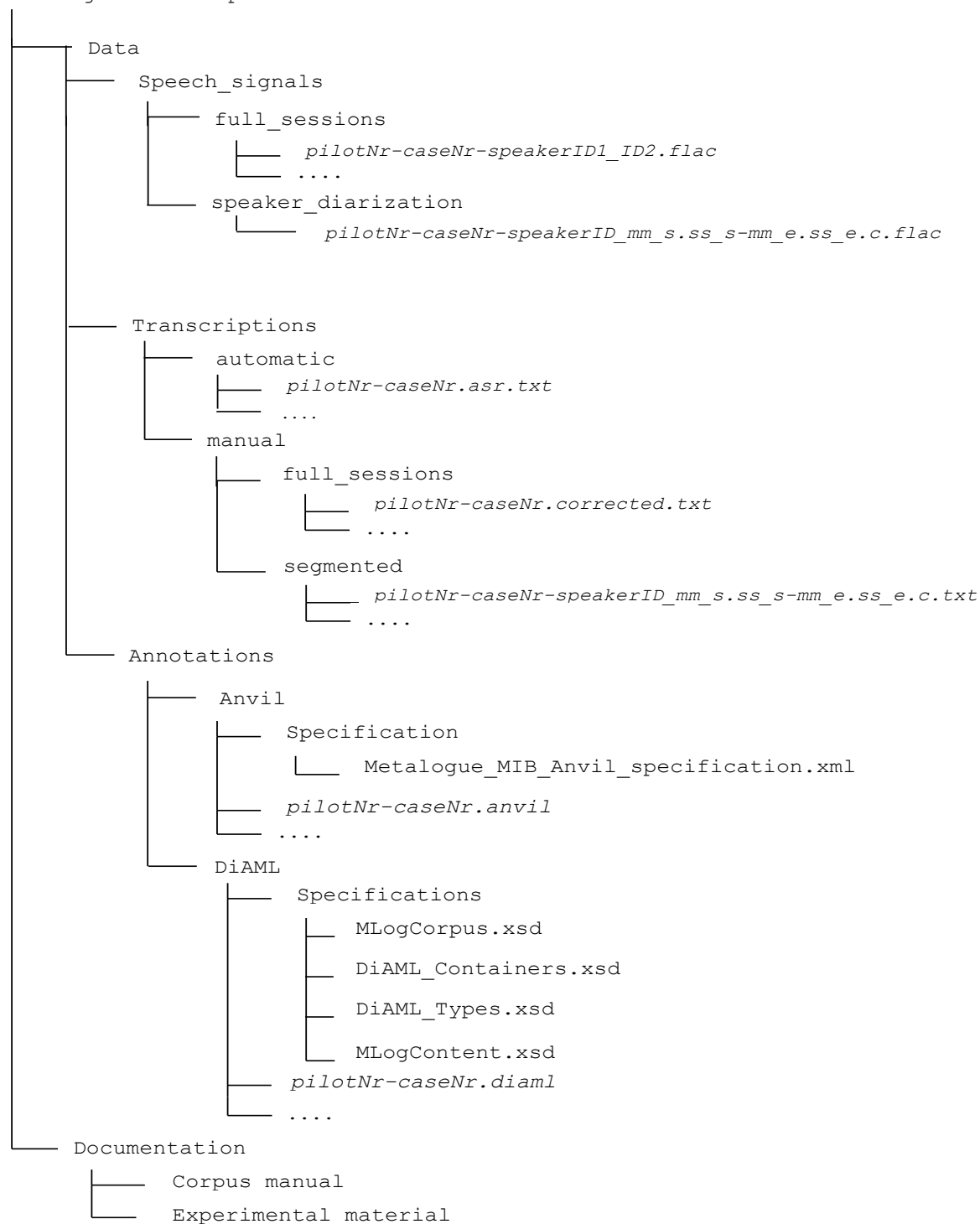


Figure 1: Delivered corpus directory structure.



In this experiment, you will play the role of a professional mediator employed by the government. In this capacity, your task is to negotiate the implementation of an anti-smoking campaign. Although the campaign is nation-wide, cities are allowed to make their own decisions concerning the details of the implementation. In fact, city councils turn out to be quite varied in their preferences. For this reason, you will enter separate negotiations with your trading partner for each city. During these negotiations, the following four aspects need to be decided on:

- Scope of smoking ban;
- Taxation of tobacco products;
- Anti-smoking campaign; and
- Enforcement of tobacco sales restrictions to minors.

Throughout the experiment, you will represent the city councils of various cities, each with different preferences concerning smoking, taxation, campaigns, and financial possibilities available to implement different solutions. You will negotiate with another professional mediator, who will represent local business owners of the same cities, which may also differ in their preferences. In each negotiation, you should try to negotiate a solution that is as close as possible to the preferences of the city that you represent. To help you in this task, you will receive a "preference card" for each city. This card lists the different options for each of the four issues. The background color of each option shows whether the city council views these options as positive or negative according to the gradient bar below.



The options that are in line with the goals of the council (and that you should try to negotiate for) are indicated with a blue background. An orange background, on the other hand, shows that the option is contrary to the goals of the city council. These options would therefore make a proposal less appealing for the city you represent. Finally, options with a white background are neutral options that the council considers to be neither positive nor negative. An agreement with only orange options will be unacceptable for the city council you represent. However, orange options can be compensated by blue options. For example, a bright blue option on taxation combined with a slightly orange option on the anti-smoking campaign results in an acceptable agreement.

Your aim is to negotiate a mutual agreement. If you believe it is not possible to negotiate an agreement with your trading partner, you may decide to stop the negotiation and move on to the next. During the negotiation, it is important to remember that the preference cards you receive are private information. As a professional mediator, you are not allowed to show this information to your trading partner.

Scope of smoking ban	Taxation of tobacco products
All outdoor smoking allowed	No change in tobacco taxes
No smoking in public transportation	5% increase in tobacco taxes
No smoking in public transportation and parks	10% increase in tobacco taxes
No smoking in public transportation, parks, and open air events	15% increase in tobacco taxes
	25% increase in tobacco taxes

Anti-smoking campaign	Enforcement of tobacco sales restrictions to minors
Flyer and billboard campaign in shopping district	Police fines for minors in possession of tobacco products
Anti-smoking posters at all tobacco sales points	Ban on tobacco vending machines
Anti-smoking television advertisements	Police fines for selling tobacco products to minors
Anti-smoking advertisements across all traditional mass media	Identification required for all tobacco purchases
	Government-issued tobacco card for all tobacco purchases

Select a negotiation:

- Case 2
- Case 3
- Case 4
- Case 5
- Case 6
- Case 7
- Case 8
- Case 9

Figure 2: Example of participants instructions and values of issues presented to participants in colours.

File naming scheme

1. Pilot number ranging from pilot 0 (pre-pilot) to pilot 3;
2. Case number ranging from case 1 to 9;
3. Speaker unique ID;
4. Utterances/Segments named using the scheme "mm_s.ss_s-mm_e.ss_e.c.flac", where mm_s.ss_s is the starting time stamp in minutes (mm_s) and seconds (ss_s) whereas mm_e.ss_e is the ending time stamp in minutes (mm_e) and seconds (ss_e). If the recording is longer than one hour then the naming scheme becomes "hh_s.mm_s.ss_s-hh_e.mm_e.ss_e.c.flac".
5. The time stamps should be extracted directly from the recording (time stamps indicated by audacity at the start and end marks).
6. Files may be additionally marked with comments by adding ".c" where c is one of the abbreviations listed below.

Example: pilot1-case6-sp17_08.22-08.30.n.flac is the segment recorded in pilot1, case 6 belonging to the speaker 7, and starts at 8 minutes and 22 seconds and finished at 8 minutes and 30 seconds during the recording session. It contains some noise (n).

Comment Abbreviations

n = contains noise (breathing, clicking, prominent other voices)

h = hesitation

g = contains some non-English (e.g. German) discussion

r = repetition (of a word, for example: during a hesitation)

ls = long silence

The ASR output (.asr.txt) and the manually corrected transcriptions (.corrected.txt) are provided in plain text format using UTF-8 character encoding. Manual transcriptions are of three types: (1) per speaker diarized file (per speaker per turn) ; (2) per dialogue session; and (3) for all sessions.

4 Dialogue act annotation: basic concepts, definitions and metamodel

According to the ISO Linguistic Annotation Framework (ISO 24612:2009) the term ‘annotation’ refers to the linguistic information that is added to segments of language data and/or nonverbal communicative behaviour. In Metalogue, we mostly consider annotations of semantic and pragmatic multimodal phenomena. For this purposes, *dialogue acts* play an important role. *Dialogue act annotation*, the activity of marking up stretches of dialogue with information about the dialogue acts performed, is usually limited to marking up their communicative functions using a given set of such functions (a ‘tag set’).

The manual defines the following core annotation concepts, processes and tools:

1. dialogue basic concepts such as dialogue participants, dialogue units, semantic descriptions, etc., and summarize them in metamodel in form of an UML diagram;
2. dialogue act annotation scheme consisting of tagset of core dialogue acts (ISO 24617-2 and proposed extensions for Metalogue) and annotation guidelines;
3. dialogue act markup language (DiAML);
4. multimodal dialogue act annotation tools.

The manual contains an elaborated terms and definition glossary. In Appendices (1) ISO 24617-2 communicative functions definitions, (2) ISO 24617-8 rhetorical (discourse) relations, (3) negotiation moves definitions, (4) speech production disfluency types, (5) ANVIL workflow and specification can be found, as well as (6) transcription format examples for spoken dialogue data.

4.1 Dialogue Act

The notion of a dialogue act is a key notion in theories of dialogue. Dialogue acts are often used in studies of dialogue phenomena, in describing the interpretation of communicative behaviour of participants in dialogue, and in the design of dialogue systems. Describing communicative behaviour in terms of dialogue acts is a way of characterizing the meaning of the behaviour. The idea of interpreting dialogue behaviour in terms of communicative actions such as statements, questions, promises, requests, and greetings, goes back to speech act theory (Austin, 1962; Searle, 1969), which has been an important source of inspiration for modern dialogue act theory. Informally speaking, a dialogue act is an act of communicative behaviour performed for some purpose, e.g. acts provide information, request the performance of an action, apologise for a misunderstanding, and provide feedback. ISO standard 24617-2 ISO (2012) defines a dialogue act as

- (1) *communicative activity of a participant in dialogue interpreted as having a certain communicative function and semantic content, and possibly also having certain functional dependence relations, rhetorical relations and feedback dependence relations.*

A *communicative function* specifies the way semantic content is to be used by the addressee to update his context model when he understands the corresponding aspect of the meaning of a dialogue utterance.

Semantic content indicates what the behaviour is about: which objects, events, situations, relations, properties, etc. Annotation of semantic content is concerned with annotating different natural language phenomena like events, named entities, semantic roles, semantic relations, etc. In Metalogue, we annotated negotiation moves as main concepts describing semantic content of dialogue acts in negotiations, see Section 4.1. When performing dialogue act annotation, we also identify the *semantic content type*. The type of semantic content can be annotated and captured by a ‘dimension’ tag. In DAMSL (Dialogue Act Markup using Several Layers), a dimension is defined as “*an abstract characterisation of the content of an utterance*” (Allen and Core, 1997). It is noticed that “*in task-oriented dialogues, we can roughly divide utterances into those that address the task in some way, those that address the communication process (Communication Management), and those that do not fall neatly into either category (Other-Level). In addition, we can subdivide the first category into utterances that advance the task (Task) and those that discuss the problem solving process or experimental scenario (Task Management)*” (Allen and Core, 1997). This is a coarse distinction of semantic content types, which may be refined by subdividing Communication Management into feedback, turn management, topic management, and other aspects. The notion of a **dimension** is defined as follows:

- (2) *A dimension is a class of dialogue acts concerned with one particular aspect of communication, corresponding to a particular type of semantic content, which a dialogue act can address independently from other dimensions.*

Dimensions, in the sense introduced here, classify dialogue acts. What is usually called a ‘dialogue act taxonomy’ is in fact a taxonomy of the *communicative functions* of dialogue acts (like Question, Offer, Confirmation, Signal-Understanding, Turn-Grabbing, Greeting, Stalling,...).

4.2 Dialogue participants

According to the definition given above, a dialogue act has at least two participants: (1) an agent whose communicative behaviour is interpreted, usually called the “speaker”, or “sender”; and (2) a participant to whom he is speaking and whose information state he wants to influence, called the “addressee” (also called “hearer” or “recipient”). There may of course be more than one addressee. For natural multimodal dialogue, where some of the dialogue acts are expressed in speech, some in a combination of speech and nonverbal elements, and some purely nonverbally, it is best to use the term “sender” for the agent who performs a dialogue act.

Besides sender and addressee(s), there may be various types of side-participants who witness a dialogue without participating in it. The presence of side-participants may influence the communicative behaviour of the participants, if these are aware of their presence, as in a television interview or a talk show. Clark (1996) distinguishes between ‘side-participants’, ‘bystanders’, and ‘overhearers’, depending on the role that they play in the communicative situation. For Metalogue tutoring interactions please consider the class diagram presented in D3.2

4.3 Dialogue segmentation: dialogue units

A dialogue act being a unit in the semantic description of communicative behaviour in dialogue, the question arises what stretches of such behaviour are considered as corresponding to dialogue acts. Spoken dialogues are traditionally segmented into *turns*. The notion of a *turn unit* has been introduced and defined as:

- (3) *stretch of communicative behaviour produced by one participant which includes the use of speech, and is bounded by periods where that participant is not speaking.*

According to this definition, a turn unit is produced by a speaker who may, in addition to speaking, also produce nonverbal communicative behaviour (such as gestures and facial expressions), and turn units produced by different speakers may overlap.

Turn units consist of more fine-grained units called *utterances*⁵. Utterances are linguistically defined contiguous stretches of (linguistic) behaviour. Levinson (1983) writes: “An utterance is the issuance of a sentence, a sentence-analogue, or sentence-fragment, in an actual context”. For example:⁶

- (4) A1: First of all just to kind of make sure that we all know each other
A2: I'm Laura and I'm the project manager

The speaker in A1 introduces the next topic for discussion in a meeting, and in A2 she introduces herself (and the role she will play in the dialogue). A1 and A2 constitute two utterances, together making up a turn unit produced by speaker A.

Segmenting a dialogue into utterances has the advantage of more fine-grained units being annotated, allowing more precise annotation; however, the notion of an utterance as a smaller unit inside a turn does not have a clear definition, and the detection of utterance boundaries is a highly nontrivial task.

The stretches of behaviour that are relevant for interpretation as dialogue acts often coincide with utterances, but they may be discontinuous, may overlap, and may even contain parts of more than one turn. They therefore do not always correspond to utterances, which is why we have introduced the notion of a **functional segment** as a minimal stretch of communicative behaviour that has a communicative function (and possibly more than one).⁷ Thus, the units of dialogue that our analysis will be concerned with are functional segments that are potentially overlapping, discontinuous and spread over different turns (see Geertzen et al. (2007).

The most natural way to take this into account in dialogue act annotation is to assign communicative functions to all those segments of behaviour that express a dialogue act, allowing these segments to overlap and to be discontinuous and to spread over multiple turns. For example, consider the 3-way segmentation of S's utterance in the following dialogue fragment, where the functional segments in each dimension are indicated in boldface:

- U: What time is the first train to the airport on Sunday morning please?
S: The first train to the airport on Sunday morning is let me see... at 5:45.
(5) TA **The first train to the airport on Sunday morning is at 5:45** ; let me see...
FB **The first train to the airport on Sunday morning** ; is let me see... at 5:45
Ti **.... let me see...** ; The first train to the airport on Sunday morning is at 5:45

In the example (5) the second turn is segmented in three dimensions: (1) Task/Activity (TA); (2) Feedback (FB); and (3) Time Management (Ti). In the TA dimension, the turn is segmented

⁵In the literature the term “utterance” is sometimes used to designate everything contributed in a single turn, in the sense of what we call a turn unit, see e.g. Allwood (1992), who uses the term “grammatical unit” for what we call “utterance”).

⁶From the AMI meeting corpus - ES2002a.

⁷These stretches are ‘minimal’ in sense of not being unnecessarily long. The rule here is: do not add material which does not contribute to the communicative function.

into the discontinuous functional segment *The first train to the airport on Sunday morning is at 5:45*, which has the function of an answer in this dimension, and the intervening stretch ... *let me see...*, which does not have a communicative function in this dimension. In the Time Management dimension the same segmentation applies, but now it's only the segment ... *let me see...* which has a communicative function (Stalling). Finally, in the Feedback dimension the turn is segmented into the functional segment *The first train to the airport on Sunday morning*, which provides positive feedback on understanding the preceding question, and the contiguous stretch *is ... let me see... at 5:45*, which is not a functional segment.

4.4 Relations between dialogue units

In order to analyse what happens in dialogue it is insufficient to only consider its segments in isolation. It is uncontroversial that discourse modelling requires the consideration of relations between semantically or pragmatically relevant units, but the nature, the purpose and the definitions of units in discourse and their relations are the subject of much controversy (see e.g. Hovy, 1990). To the rhetorical relations identified in monologue (e.g explanation, justification, cause,...), dialogue adds relations such as those between a question and an answer, and between an utterance and feedback about its understanding.

All responsive dialogue acts have a 'functional antecedent', being the dialogue acts that they respond to; those of type A have a semantic content that is co-determined by that of their functional antecedent. This is a relation between two dialogue acts or between a dialogue act and a group of dialogue acts, as in (6).

- (6) A1: Can you tell me what time is the first flight in the morning to Munich?
 B1: On what day do you want to travel?
 A2: Tomorrow.
 B2: Tomorrow morning
 B3: The first flight that I have is at 7:45.

The dialogue act in B3 is functionally related to the group consisting of the question in A1 and the answer (to B1) in A2, which together are equivalent to a more complete question which B3 answers.

If the meaning of a responsive dialogue act depends on the meaning of a previous dialogue act (or dialogue acts) due to its communicative function, then this dependence is called a *functional dependence relation* (Bunt et al., 2010). More explicitly:

- (7) *A functional dependence relation exists between a dialogue act DA_1 and one or more previous dialogue acts $\{DA_2, \dots, DA_N\}$ iff the meaning of DA_1 depends on the meaning of $\{DA_2, \dots, DA_N\}$ due to the responsive character of DA_1 .*

Responsive dialogue acts of another type provide or elicit information about the (perceived) success in processing a segment of communicative behaviour earlier in the dialogue. Such a relation is called a *feedback dependence relation*. This type of relation has been defined in ISO standard 24617-2 as follows:

- (8) *A feedback dependence relation is a relation between a feedback act and the stretch of communicative behaviour whose processing the act provides or elicits information about.*

Feedback acts refer explicitly or implicitly to the stretch of dialogue that they provide or elicit information about. This stretch of dialogue forms part of its semantic content.

Rhetorical relations have been proposed as an explanation for the construction of coherence in discourse or at least as crucial modelling tools for capturing this coherence, see e.g. Hobbs (1985a); Mann and Thompson (1988); Sanders et al. (1992); Asher and Lascarides (2003). The idea is that two text segments or sentences in written discourse, or two segments or utterances in dialogue, are linked together by means of certain relations, for which various terms have been used such as ‘rhetorical relations’, ‘coherence relations’, or ‘discourse relations’, such as *Explanation*, *Justification*, *Cause*, etc., see Annex 2 for the Metalogue rhetorical relations set.

4.5 Communicative function qualifiers

Participants in a dialogue do not just exchange information by simple statements, direct questions and clear-cut answers. They may be less straightforward in expressing their communicative intentions, formulating a question indirectly or accepting a request conditionally. They often indicate their attitude toward their communicative partners, toward what they are saying, or toward things that they intend to do. They emphasize, express doubts, criticize, show interest, and so on. All this can be signalled in various ways, e.g. by using verbal indicators like modals, by intonation and by utilizing body language and facial expressions. Approaches to the analysis, annotation, or computational modelling of dialogue behaviour struggle with these phenomena. This is especially true for attempts to annotate spoken and multimodal dialogue with information about the communicative actions (‘dialogue acts’) that the participants perform. For example:

- (9) 1. A: Would you like to have some coffee?
 2. B: Only if you have it ready.
 3. B: Coffee could be nice, but what time is it now?

The response 2 in (9) can be characterized as *conditional acceptance of offer* and response 3 as *uncertain acceptance of offer*.

A qualifier is an additional element in the description of dialogue acts. Semantically, qualifiers describe and provide more accurate definitional meaning for another element. Communicative function qualifiers do not change but specify more precisely the way the act’s semantic content changes the addressee’s information state, e.g. by expressing the strength or weakness of certain assumptions and beliefs, or the physical and emotional abilities and state of a dialogue participant. In other words, qualifiers provide a more detailed description of the speaker’s intention and is formally defined in ISO 24627-2 as:

- (10) *A predicate that can be associated with a communicative function.*

A study of these phenomena by Petukhova & Bunt (2009c) indicates that, for dealing with the most frequent cases, the following qualifiers are sufficient:

- **certainty:** the qualifier ‘uncertain’ can be used with information-providing functions, in order to indicate that the speaker is uncertain about the correctness of the information that he provides.
- **conditionality:** the qualifier ‘conditional’ can be used with action-discussion functions, which have in common that the speaker assumes that the action under discussion can be performed by the participant whose action is discussed (the speaker, in the case of

commissives; the addressee, in the case of directives). The ‘conditional’ qualifier indicates that this assumption is dropped.

- sentiment: these qualifiers indicate that the speaker has a certain emotional attitude to the semantic content that the communicative function is applied to, or towards the addressee.

4.6 Dialogue act annotation metamodel

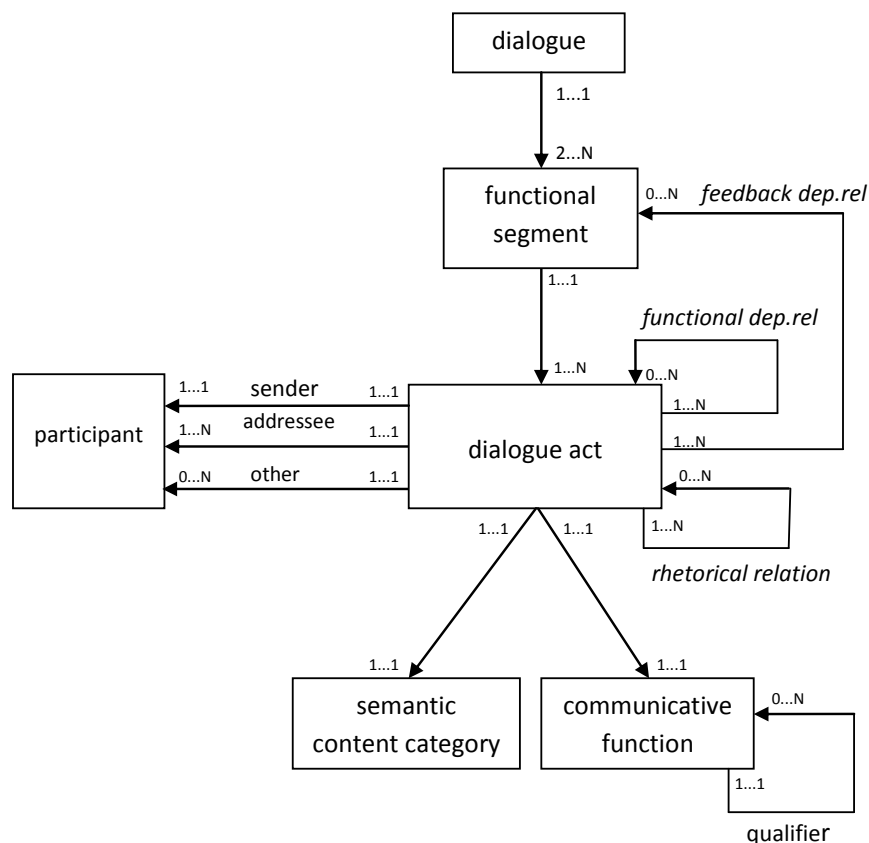


Figure 3: ISO 24617-2 metamodel for dialogue act annotation.

To sum up, in the characterization of the notion of a dialogue act and its realization, as given so far, the following key elements occur, which will form the backbone of the proposed metamodel⁸ for dialogue act annotation (see Figure 3):

- sender (or ‘speaker’)
- addressee(s)
- participants in other roles (such as overhearers)
- functional segment

⁸The term ‘metamodel’ is often used to describe a very general model that tries to capture the most basic notions underlying several alternative models, see e.g. Bunt and Romary (2004).

- dialogue act
- communicative function
- communicative function qualifier
- semantic content type
- functional dependence relation
- feedback dependence relation
- rhetorical relations between dialogue acts

The metamodel shows a representation of the fundamental upper-level concepts that are involved in dialogue act annotation. A dialogue consists of two or more functional segments, as indicated by 2..N at the head of the arrow relating them. Each functional segment is related to one or more dialogue acts, reflecting the possible multifunctionality of functional segments. Each dialogue act has exactly one sender, one or more addressees, and possibly other participants (0..N). It has a semantic content of a certain type, and one communicative function, which may have any number of function qualifiers; and is possibly related to other dialogue acts through functional dependence and rhetorical relations, and to functional segments through feedback dependence relations.

4.7 ISO 24617-2 dialogue act annotation scheme

During the 1980s and 1990s a number of dialogue act annotation schemes have been developed, such as those of the TRAINS project in the US (Allen et al., 1994), the HCRC MapTask studies in the UK (Carletta et al., 1996), and the Verbmobil project in Germany (Alexandersson et al., 1998). These schemes were all designed for a specific purpose and a specific application domain. In the 1990s a general-purpose scheme for multidimensional dialogue act annotation was designed called DAMSL (Allen and Core, 1997). Several variations and extensions of the DAMSL scheme have been constructed for special purposes, such as Switchboard-DAMSL (Jurafsky et al., 1997), COCONUT (Di Eugenio et al., 1998) and MRDA (Dhillon et al., 2004). The DIT⁺⁺ scheme (Bunt, 2006 and 2009) combines the multidimensional DIT scheme developed earlier (Bunt, 1994) with concepts from DAMSL and various other schemes, and provides precise definitions for its communicative functions and dimensions.

In September 2012, the ISO standard 24617-2 “Semantic annotation framework, Part 2: Dialogue acts” has been developed where a comprehensive annotation scheme and markup language DiaML were designed. The ISO 24617-2 standard annotation scheme is a comprehensive, application-independent scheme whose concepts are empirically and theoretically well-motivated, and may be exploited for constructing annotated dialogue corpora. ISO 24617-2 is a highly multidimensional scheme supporting multifunctionality, since it offers the possibility to assign multiple dialogue act tags to one dialogue segment. The ISO 24617-2 taxonomy of communicative functions distinguishes 9 dimensions: addressing information about a certain (*Task*); the processing of utterances by the speaker (*Auto-feedback*) or by the addressee (*Allo-feedback*); the management of difficulties in the speaker’s contributions (*Own-Communication Management*) or that of the addressee (*Partner Communication Management*); the speaker’s need for time to continue the dialogue (*Time Management*); the allocation of the speaker role (*Turn Management*); the structuring of the dialogue (*Dialogue Structuring*); and the management of social obligations (*Social Obligations Management*). Figure 4 depicts 9 core dimensions

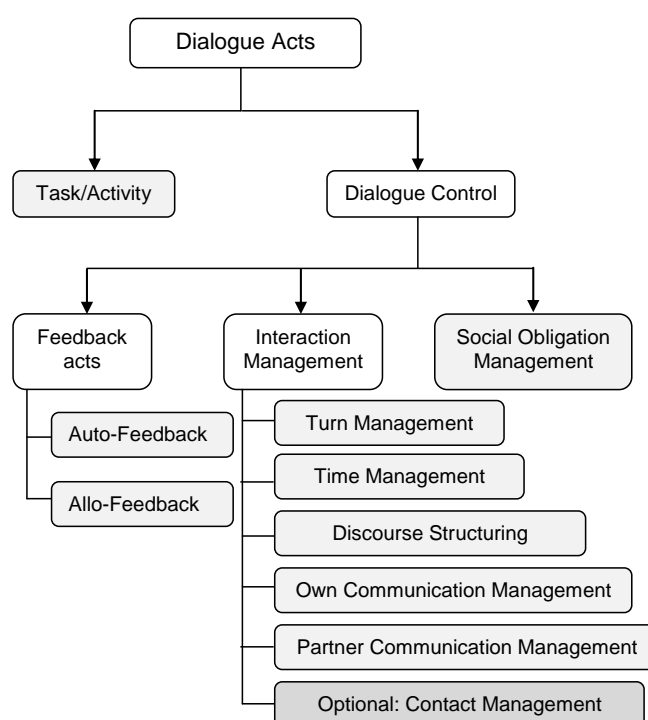


Figure 4: ISO 24617-2 9+1-dimensional hierarchy.

defined in ISO 24617-2 dialogue act annotation scheme and 1 optional dimension dealing with *Contact Management*.

A 'dimension' in the context of dialogue act analysis being a type of communicative activity, corresponding to one of the multiple aspects of interacting that dialogue participants monitor and manage.

There are communicative functions that are not specifically related to any dimension, e.g. one can ask a question about any type of semantic content, provide an answer about any type of content, or request the performance of any type of action (such as '*Please close the door*' or '*Could you please repeat that*'). Question, Answer, Request, Offer, Inform, and many other functions have this property that they can be applied to a wide range of semantic content types. Given a set of dimensions, the dialogue act that results from applying the function to a particular content can be classified depending on the type of its content. Because they can be used to address any dimension, these communicative functions are called *general-purpose communicative functions* (GP functions). General-purpose functions are of two types: *information transfer* functions and *action discussion* functions. Information transfer functions are used to obtain (*information-seeking* functions) or to provide information (*information providing* functions). Action discussion functions have a semantic content consisting of an action, and possibly also a description of a manner or frequency of performing the action and are concerned either with the speaker's commitment to perform a certain action (commissives) or his wish that the addressee performs an action (directives). Figure 5 presents the GP functions hierarchy defined in ISO 24617-2.

In contrast with the general-purpose communicative functions, other functions can only be used to address a specific dimension, such as Turn Keep and Turn Release which are specific for the dimension of Turn Management; and Stalling and Pause for the dimension of Time Management. Table 1 lists some examples of *dimension-specific communicative functions* (DS functions) in

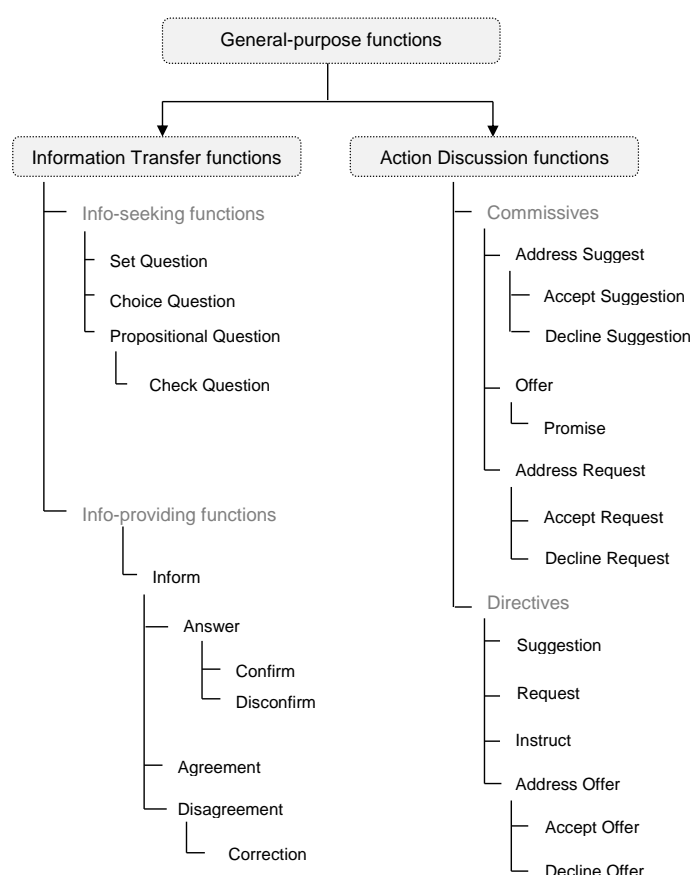


Figure 5: ISO 24617-2 general-purpose functions.

some of the dimensions distinguished in the ISO 24617-2 scheme.

Dimension	Dimension-specific comm, functions	Typical expressions
Auto-Feedback	Negative Positive	<i>Huh?</i> <i>OK.</i>
Allo-Feedback	Negative EvaluationElicitation	<i>THIS Thursday.</i> <i>OK?</i>
Turn Management	TurnKeeping TurnGrabbing TurnGiving	final intonational rise hold gesture with hand <i>Yes.</i>
Time Management	Stalling Pausing	slowing down speech; fillers <i>Just a minute</i>
Contact Management	ContactChecking	<i>Hello?</i>
Discourse Structure Management	TopicShiftAnnouncement	<i>Something else.</i>
Social Obligations Management	Apology Greeting	<i>I'm sorry.</i> <i>Hello.</i>
	Thanking	<i>Good morning</i> <i>Thanks.</i>

Table 1: Examples of dimension-specific communicative functions and their expression for some of the dimensions distinguished in the ISO 24617-2 dialogue act annotation scheme.

Figure 6 presents all dimension specific communicative functions defined in ISO 24617-2. In total, there are 32 dimension-specific and 25 general-purpose communicative functions. All ISO 24617-2 definitions can be found in Annex 1 as specified for ISOcat Data Category Registry (see <http://www.isocat.org/>).

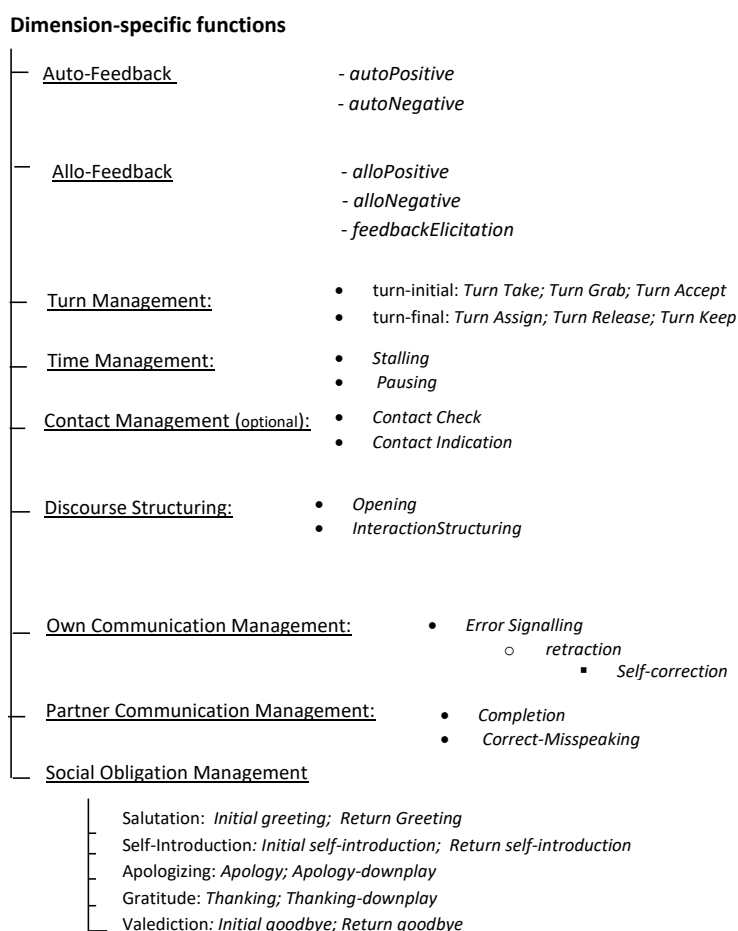


Figure 6: ISO 24617-2 dimension-specific functions.

4.8 Annotation guidelines

Another important part of an annotation scheme is *annotation guidelines*, which provide general principles and concrete instructions for how the tags should be used. They serve two main purposes: (1) to support the decision-making process of human annotators; and (2) to provide recommendations for possible extensions, modifications, or restrictions of the scheme as the need arises for particular applications.

Annotation guidelines for ISO 24617-2 can be found at <http://dit.uvt.nl/is-guida.html>.

4.9 Proposed extensions for Metalogue

In Metalogue, we used the above described ISO 24617-2 dialogue act annotation scheme. The ISO 24617-2 annotation schema, however, cannot be expected to be ideal for every kind of dialogue analysis, for every task domain, for every kind of dialogue, and for every annotation purpose. Nevertheless, general principles underlying the design of the schema and the DiAML annotation language are useful for accommodating extensions, modifications, or restrictions of the schema and the annotation language, as the need arises for particular applications. In Metalogue, we followed the main design principles and guidelines for schema extension and

restriction formulated in ISO 24617-2 standard in Section 12.

For Metalogue purposes, we considered 2 additional dimensions (11 in total): *Contact Management*, which is non-core optional in ISO24617-2, since, for example, in Metalogue negotiation sessions when the system represented by an avatar is involved as a participant managing the contact is an important aspect, and *Task Management*, for dialogue acts addressing the negotiation process and procedure. This includes utterances that involve coordinating the activities of the two speakers (e.g., “Are you keeping track of the time?”, “Let’s go issue by issue”, etc.), asking for help on the procedures (e.g., “Do I need to state the problem?”) or asking about the status of the process (e.g., “Are we done with the agenda?”). It is important to distinguish between utterances that concern the task management when addressing the task procedures and discourse structuring when addressing management of the interactive/dialogue behaviour and procedures. Examples of the later are utterances like “To sum up”, “Let’s move to an next topic”, etc. For Discourse structuring, we (re-)introduced 4 additional dimension-specific functions that are not included in ISO 26417-2, however, defined in DIT⁺⁺ Bunt (1999):

- *Topic introduction*, where the speaker wants to introduce the topic mentioned in the semantic content;
- *Topic shift announcement*, where the speaker wants to change the topic;
- *Topic shift*, where the speaker wants to shift the topic to the one mentioned in the semantic content;
- *Closing*, where the speaker indicates that he/she plans to end the current dialogue shortly

The underlying task of dialogue participants in Metalogue is concerned with negotiation scenarios. Therefore, an additional set of negotiation moves is defined in Annex 3 and modelled later as dialogue act content, see Sections 5 and 6.

In Metalogue we also extended the set of rhetorical relations, which is left unspecified in ISO 24617-2. The following set is considered based on empirical evidence from collected data that is mainly based on PDTB discourse relations Prasad et al. (2008), which on its turn form the basis for recently proposed and accepted ISO 24617-8 - Language Resource Management - Semantic Annotation Framework - Part 8: Semantic relations in discourse, Core annotation scheme (ISO DR-core) Bunt and Prasad (2016), and extensions adopted from Hovy and Maier (1995), see Annex 2 for definitions.

5 Metalogue domain and tasks: negotiation structure and moves

Metalogue dialogue domain is multi-issue bargaining training. Analysing the collected negotiation data we noticed that the ISO 24617-2 dialogue acts inventory is not sufficient to interpret and model negotiation interactions, and requires some extensions. Consider the following example:

- (11) P1: What’s your opinion on scope of smoking ban?
 P2: I think there shouldn’t be smoking in public transportation and parks

Analyses according to ISO dialogue acts standard will result in assigning to P1 *Set Question* tag and to P2 *Set Answer*. Dialogue context model will be updated accordingly. For negotiation analysis, P1 is rather *Offer Elicitation* act and P2 is an *Offer*. When analysing negotiations

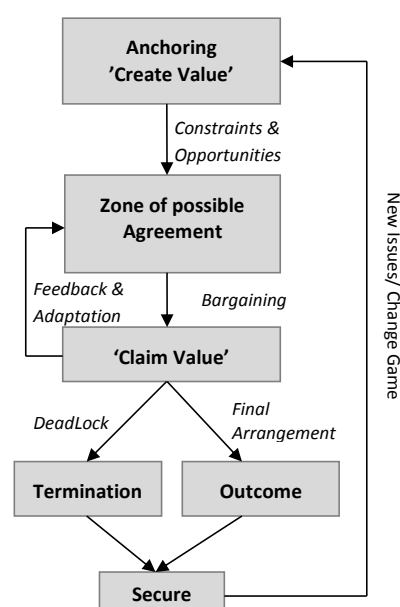


Figure 7: Negotiation phases associated with negotiation structure or certain negotiation strategy.

ID	Speaker	Start-End time	Utterance	Dialogue Aact tag	Negotiation Move	Relation
u1	p1	00.00-00.16	in this city I would suggest all outdoor smoking allowed	Task;Suggest	OfferValue	
u2	p2	00.16-00.17	uh-uhu	AutoPositive [u1]		
u3	p1	00.17-00.25	no changes in tobacco taxes and			
u4	p1	00.25-00.30	then anti-smoking television advertisement	Task;Suggest	OfferValue	list[u1]
u5	p2	00.30-00.31	uh-uhu	AutoPositive [u3,u4]	OfferValue	list[u1,u3]
u6	p1	00.31-00.33	so what do you think	Task;SetQuestion	ElicitOfferValue	
u7	p2	00.32-00.33	uhm	TurnTake;Stal		
u8	p2	00.33-00.34	yeah	AutoPositive		
u9	p2	00.34-00.36	that's bit difficult for me	Task;SetAnswer[u6]	AddressOfferValue[u6]	
u10	p2	00.36-00.42	because that really doesn't meet our goals	Task;Inform		justify [u9]
u11	p2	00.41-00.49	but we can sure look if we can find a solution maybe	Task;Suggest		contrast [u10]
u12	p2	00.49-00.56	maybe i start with the worst points for me	Task;Suggest		
u13	p1	00.55-00.56	okay	Task;AcceptSuggest[u12]		
u14	p2	00.55-01.01	it's the scope of the smoking ban	DiscourseStructuring;TopicShift		
u15	p1	01.01-01.02	uh-uhu	DiscourseStructuring; Agreement [u14]		
u16	p2	01.00-01.12	as i said i think only to allow outdoor smoking is not enough	Task;Inform	DeclineOfferValue [u1]	
u17	p2	01.13-01.28	i think it would be fine if we stop smoking in public transportation	Task;Inform	BargainDown	
u18	p1	01.36-01.37	okay i would go for that point	Task;Agreement[u17]	AcceptOfferValue[u17]	

Table 2: Example of multi-level negotiation dialogue annotation.

offers, counter-offers, commitments, concessions, etc. are commonly used (see Watkins (2003), Afantenos et al. (2012), Hindriks et al. (2007)). Both types of information are necessary to model negotiation dialogues successfully. This information allows the system to interpret partners' and to generate adequate communicative behaviour and to interpret partner's negotiation actions and strategies, and to take correct decisions in negotiation.

Bargaining structure may shape strategies that negotiators follow and may influence the overall outcome. Negotiation starts with the **anchoring** phase, in which participants bring up early offers and counter-offers establishing jointly possible values contributing to the **Zone of possi-**

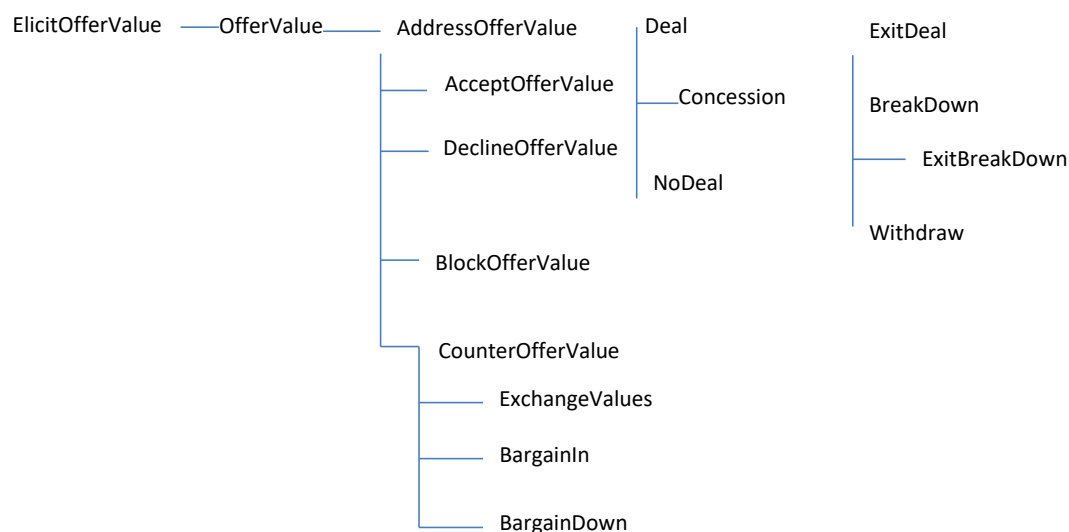


Figure 8: Negotiation moves defined in Metalogue.

ble Agreement (or bargaining range). The Zone of possible Agreement (ZOPA) describes the intellectual zone in negotiations between parties where an agreement can be reached. Within this zone, an agreement is possible. Outside of the zone, no amount of negotiation will yield an agreement.

The actual bargaining occurs in 'Claim Value' phase, potentially leading to (1) adaptation of the originally established ZOPA, (2) Negotiation Outcome, or (3) Negotiation Termination. Negotiation moves observed here are *BargainIn*, *BargainDown*, *BlockOffer*, *Withdraw*, and *Concession(-s)*. **Negotiation Outcome** is the phase associated with all 'walk-away' positions for each partner. This phase is mainly concerned with stating (*partial*) *Agreement(-s)*, *Disagreement(-s)* on a certain value. Negotiations might be terminated. **Termination** is the phases associated with *deadlock* situations in which two or more competing actions are each waiting for the other to finish, and thus neither ever does. No other actions are further possible and interaction stops without any result (either positive or negative) can be reached. Specific acts that can be observed here are *Breakdown(-s)* and *Withdraw(-s)*. Negotiators can move to **Secure** (LockIn) the outcome reached so far and either go to another issue or new negotiation round, where previous *BreakDown(-s)* may be cancelled, e.g. *ExitBreakDown*. Secure phase is concerned with summing up, restating reached negotiation or termination outcomes. Participants take decisions to move with another issue, or continue or re-start the discussion later. Figure 7 depicts the observed negotiation structure.

Analysing the collected data we observed and defined 17 negotiation moves (see Figure 8 for hierarchy and Annex 3 for definitions). Negotiation moves with their arguments (value slots discussed) are modelled as semantic content of task-related dialogue acts and specify the negotiation semantics which can be extended with other entities and relations, e.g. modality, named entities, semantic roles, etc., see next Section 6 for the representation.

6 Annotation format: Dialogue Act Markup Language (DiAML)

The representation of annotations in the ISO Dialogue Act Markup Language (DiAML) relies on a three-level architecture:

1. a primary source, which may correspond to a speech recording, orthographic (or phonetic) transcription or any low-level annotation thereof, e.g. a tokenisation, or to visual movement tracking data (e.g. Kinect) encoded as vectors of movement features;
2. the marking of functional segments from the primary source;
3. the actual dialogue act annotation associated with a functional segment.

The representation of a dialogue act annotation makes use of the XML element `<dialogueAct>`. This element has the following attributes:

- `@target`, whose value is a functional segment identified at the second level;
- `@sender`, `@addressee`, `@otherParticipant`;
- `@communicativeFunction`, `@dimension`;
- `@certainty`, `@conditionality`, and `@sentiment` qualifiers;
- `@functionalDependence` and `@feedbackDependence`, which have `<dialogueAct>` elements and `<functionalSegments>` as values.

Additionally, rhetorical relations among dialogue acts are represented by means of `<rhetoLink>` elements.

`DiAML_Types.xsd` scheme specifies DiAML types such as dialogue act, dimensions, communicative functions, rhetorical links, certainty, conditionality and sentiment qualifiers. These types are defined in `diaml` namespace. In `DiAML_Containers.xsd` elements such as primary data tokens and sounds, and functional segments are specified without a namespace. This allows to use `DiAML_Containers` as chameleon type of schema.

Semantic content of a dialogue act is defined outside DiAML and is represented as `<NegotiationSemantics>` elements, see `MLogContent.xsd`. `MLogContent.xsd` specifies dialogue acts as elements based on `Types` and `Containers`. DiAML containers and elements from `MLogContent.xsd` are defined inside `metalogue` namespace.

In `Metalogue`, shallow negotiation semantics is defined in terms of `<NegotiationMove>` with attributes defined for different types of such moves discussed above and presented in Annex 3. For example:

```
<dialogueAct xml:id="dap1TSK38" sender="#p1" addressee="#p2"
  dimension="task" communicativeFunction="inform" target="#fsp1TSKCV38">
  <NegotiationSemantics>
    <NegotiationMove type="counterOfferValue"/>
  </NegotiationSemantics>
  <rhetoricalLink rhetoAntecedent="#dap2TSK37" rhetoRel="substitution"/>
</dialogueAct>
```

Additionally, in `Metalogue` `<NegotiationSemantics>` elements contain information on the annotated `<Repairs>` types, see Annex 4. Further, dependent on an annotation goals, approach, granularity and type of semantic processing, negotiation semantics specification is extended with `<Arg>` elements for possible negotiated issues and values, `<Operators>` for logical operators between arguments and optional `<Modalities>` elements to describe speaker's attitudes and

perception of the events probability. Semantics can be further extended to represent predicate-argument structures, named entities, semantic roles or other relations, etc.

A concrete example of the use of DiAML in (12). P2's utterance is segmented into two overlapping functional segments: one in the Auto-Feedback dimension and one in the Task dimension, with value 'answer' qualified as 'uncertain'. Annotations may be attached directly to primary data like stretches of speech, defined by temporal begin and end points, but more often they will be attached to structures at other levels of analysis, such as the output of a tokenizer. TEI-ISO standard ISO 24610-1 is followed for attaching information to digital documents. In the example, the dialogue participants are assumed to be identified in the metadata of the primary data as "p1" and "p2", and their utterances are segmented multidimensionally into the functional segments "fs1", "fs2.1", and "fs2.2".

(12) a.

P1: *What do you propose for scope of the smoking campaign?*

P2: *For scope, I think I prefer all outdoor smoking allowed, because it does not affect non-smokers.*

fs2.1 for scope [positiveAutoFeedback]

fs2.2 I think I prefer all outdoor smoking allowed [answer, uncertain]

fs2.3 because it does not affect non-smokers [inform, justify]

b.

```
<diaml xmlns:"http://www.iso.org/diaml/">
```

```
<dialogueAct xml:id="da1" sender="#p1" addressee="#p2"
target="#fs1" communicativeFunction="setQuestion"
dimension="task"
```

```
  <NegotiationSemantics>
```

```
    <NegotiationMove type="elicitOfferValue"/>
```

```
  </NegotiationSemantics>
```

```
</dialogueAct>
```

```
<dialogueAct xml:id="da2" sender="#p2" addressee="#p1"
target="#fs2.1" communicativeFunction="autoPositive"
dimension="autoFeedback" feedbackDependence="#fs1"
</dialogueAct>
```

```
<dialogueAct xml:id="da3" sender="#p2" addressee="#p1"
target="#fs2.2" communicativeFunction="answer"
qualifier="uncertain" dimension="task"
functionalDependence="#da1"
```

```
  <NegotiationSemantics>
```

```
    <NegotiationMove type="offerValue"/>
```

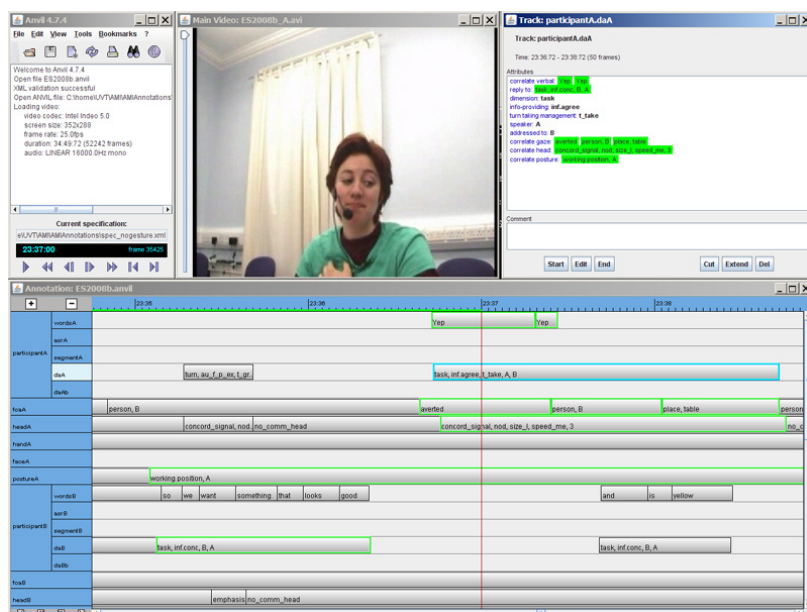


Figure 9: Annotation of multimodal dialogue acts using Anvil.

```
</NegotiationSemantics>
</dialogueAct>
```

```
<dialogueAct xml:id="da4" sender="#p2" addressee="#p1"
target="#fs2.3" communicativeFunction="inform"
qualifier="uncertain" dimension="task"
functionalDependence="#da1"
<rhetoLink dact="#da4" rhetoAntecedent="#da3" rhetoRel="justify"/>
</dialogueAct>
```

```
</diaml>
```

Annotation files of .diaml type consists of TEI-compliant primary data representation ISO (2006) and ISO 24617-2 dialogue act annotations attached to this primary data. Metalogue specific elements based on DiAML types are defined inside metalogue namespace. MlogCorpus.xsd adds TEI element (containing TEI-compliant primary data representation) into metalogue namespace. Using chameleon scheme and several xsd files allows validation of xml files that include any of the xml elements defined in diaml, metalogue and TEI schemas. All Metalogue corpus related XSD schemes are provided in the DiAML/Specifications folder. As for TEI schemes, we refer to the TEI standard specific documentation at <http://www.tei-c.org/index.xml>.

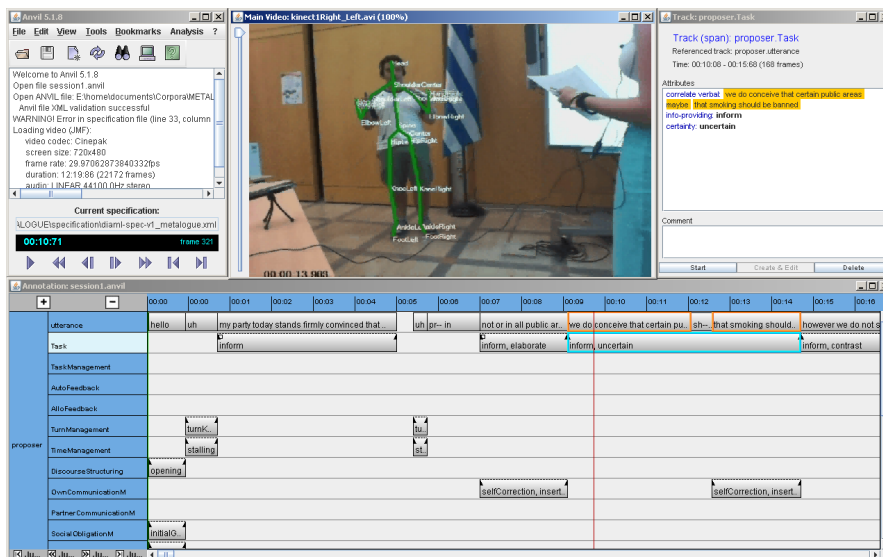


Figure 10: ISO 24617-2 dialogue act annotation in Metalogue using Anvil.

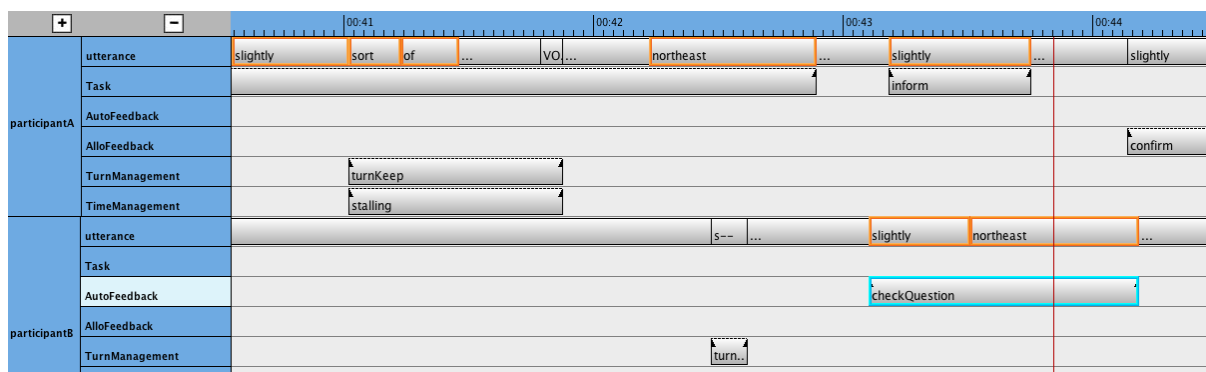


Figure 11: ANVIL cross-tier links to encode verbal correlates and dependencies. The selected annotation is marked blue, the orange frames indicate linked up elements. Note that some tracks are hidden for clarity.

7 ANVIL Annotation Tool

In the past few years, a dozen of tools has become available for annotation of digital audio-video data, The most widely used are ELAN⁹, ANVIL¹⁰, EXMARaLDA¹¹ and TASX¹². These tools have common features but differ with respect to usability and flexibility in adapting them to specific annotation needs. At workshop ‘Comparison of multimodal annotation tools’ in 2006 Rohlfing et al. (2006), the group of expert users, developers and researchers came together and analysed the above mentioned tools specifying strength and shortcomings of each. No ultimate decision could have been made in favor of one or the other. It mainly depends on researcher and his/her analysis which tool is the most appropriate one. The majority of tools, however, support data export in an XML format which enables the data exchange between them. Some of tools have build-in format conversion functionality. For instance, it is possible to import ELAN files into ANVIL. Given the analysis outcome summarized in Rohlfing et al. (2006), for Metalogue purposes ELAN and ANVIL are the top runners, since EXMARaLDA is a specialized tool for conversation analysis and therefore it is text-based, i.e. it lacks the temporal precision that many multimodality researchers need, and TASX is not longer available. While ELAN and ANVIL are compliant with the Metalogue requirements for usability and flexibility. Both allow user-defined coding schemes, offer various tier relationships and controlled vocabularies. ELAN is widely used in linguistic communities. However, it lacks rich elements so that every attribute need a separate tier and it does not allow any links between elements. Thus, in Metalogue the decision has been made in favor of the ANVIL tool.

The ANVIL annotation tool (Kipp (2001), Kipp (2008) and Kipp (2012)) was developed for the annotation of digital video, offering a graphical user interface for creating annotation elements on temporal, hierarchical, user-defined tiers. ANVIL has proved to be a very useful tool for the annotation of multimodal and spoken dialogue (see e.g. Petukhova and Bunt, 2009a; 2009b), where its tiered representation form is convenient for annotating the communicative behaviour of a dialogue participant in each modality in a separate tier (e.g. using one tier for speech, one for gaze direction, one for head movements, and one for body posture). See the illustrative example in Figure 9 as used in previous work on AMI corpus (Petukhova and Bunt, 2010).

ANVIL’s tiered format has also proved convenient for multidimensional annotation, when stretches of communicative behaviour are marked up with multiple tags, especially when the various tags are provide functional information relating to a particular dimension of interaction, such as feedback, turn taking, or time management (see Petukhova, 2011; Petukhova and Bunt, 2012). Figure 10 illustrates this for Metalogue corpus annotation.

Annotations are of stand-off (or stand-alone) type (see glossary definition 9.22) and performed using the ANVIL specification designed for ISO 24617-2¹³ The ANVIL tool allows annotations in multiple tiers so that for each participant we specified a speech tier and several tiers for each dimension. Dialogue act annotations are saved in `.anvil` format and converted into `.diaml` format.

⁹<https://tla.mpi.nl/tools/tla-tools/elan/>

¹⁰www.anvil-software.de

¹¹<http://www.exmaralda.org/>

¹²Last URL <http://medien.informatik.fh-fulda.de/tasxforce> not working any more.

¹³The specification is available in the Specification folder of this corpus release.

Type	Content	Format	Comment
Preference cards	9 negotiation cases	html for web-presentation	defined for City Councilor and Small Business Representative
Signals	speech flac files	mono, 16000Hz sample rate 16-bit sample format mono, 16-bit sample format	1 channel per speaker cut per speaker/per turn
Automatic Speech Recognition Transcriptions	turn (id, start, end, string) turn (id, start, end, string)	plain text plain text	automatic manual
Primary data (segmented)	utterance (id, start, end, string) functional segments (id, start, end, pointers)	xml, TEI compliant xml, TEI compliant	automatic automatic
DA annotations	dialogue act (sender, addressee, dimension, communicative function, qualifier functionalDependenceRelation feedbackDependenceRelation rhetoricalLinks negotiation moves disfluency types)	Anvil and DiAML	manual

Table 3: Metalogue Multi-Issue Bargaining Corpus overview.

7.1 Workflow

In ANVIL, the layout and functionality of tiers (also called tracks) is defined in a separate XML file, the so-called *specification file*. As already mentioned, for Metalogue, we provide a specific specification file in the Specification folder. A screenshot (just the annotation board) of an annotation session in progress can be seen in Figure 11. Each dialogue is annotated manually, resulting in ANVIL data files (.anvil). Detailed workflow and guidelines for using ANVIL tool can be found in Appendix 5.

8 Conclusions and future extensions

The manual presented the basic dialogue act annotation concepts, annotation tools, guidelines and structure of the Metalogue annotated corpus. The following entities are identified and annotated:

- 2.010 ISO 24617-2 dialogue act annotations in 9 ISO 24617-2 dimensions
- 1.100 DIT⁺⁺ Discourse Structuring acts
- 15 DIT⁺⁺ Contact Management acts
- 275 Task Management dialogue acts
- 1.375 Negotiation moves
- 650 ISO 24617-8 discourse relations
- 750 speech disfluencies

In total, the Metalogue multi-issue bargaining corpus contains 8.775 annotated entities.

The fact that the annotations were performed in compliance with the internationally accepted annotation and representation standards extending those for the Metalogue application specific purposes makes the Metalogue MIB corpus a valuable interoperable semantic resource,

supports data exchange and contributes to better computational modelling of the negotiation dialogues of such type. Table 3 provides corpora overview specifying type of data offered for release.

The Metalogue annotated corpus required substantial investment. We expect it to have a great impact on the rest of the project and outside. The Metalogue partner, Saarland University, will continue to maintain the corpus and to take an interest in its growth, e.g. next release - Metalogue Debate Corpus - is planned for 2017 and MIB Corpus Part 2 in 2018.

Bibliography

- Afantenos, S., Asher, N., Benamara, F., Cadilhac, A., Degremont, C., Denis, P., Guhe, M., Keizer, S., Lascarides, A., Lemon, O., Muller, P., Paul, S., Popescu, V., Rieser, V., and Vieu, L. (2012). Developing a corpus of strategic conversation in the settlers of catan. In *Proceedings of the GameNLP Workshop*.
- Alexandersson, J., Buschbeck-Wolf, B., Fujinami, T., Kipp, M., Koch, S., Meier, E., Reithinger, N., Schmitz, B., and Siegel, M. (1998). *Dialogue acts in Verbmobil-2*. DFKI Saarbrücken, University of Stuttgart; TU Berlin; University of Saarland.
- Allen, J. and Core, M. (1997). Draft of DAMSL: Dialog Act Markup in Several Layers. Available at <http://www.cs.rochester.edu/research/cisd/resources/damsl/>.
- Allen, J., Schubert, L., Fergusorr, G., Heeman, P., Hee Hwang, C., Kato, T., Light, M., Martin, N., Miller, B., Poesio, M., and Traum, D. (1994). The TRAINS project: a case study in building a conversational planning agent. TRAINS Technical Note 94-3, University of Rochester.
- Asher, N. and Lascarides, A. (2003). *Logics of Conversation*. Cambridge University Press.
- Austin, J. (1962). *How to do things with words*. University Press, Oxford.
- Besser, J. (2006). A corpus-based approach to the classification and correction of disfluencies in spontaneous speech. Master Thesis, Saarland University, Saarland, Germany.
- Bunt, H. (1994). Context and dialogue control. *THINK Quarterly* 3(1), pages 19–31.
- Bunt, H. (1999). Dynamic interpretation and dialogue theory. In Taylor, M., Neel, F., and D., B., editors, *The structure of multimodal dialogue II*, pages 139–166. John Benjamins, Amsterdam.
- Bunt, H. (2006). Dimensions in dialogue act annotation. In *Proceedings of the International Conference on Language Resources and Evaluation*, pages 919–924, Genoa, Italy.
- Bunt, H. (2009). The DIT++ taxonomy for functional dialogue markup. In Heylen, H., Pelachaud, C., Catizone, R., and Traum, D., editors, *Proceedings of the AAMAS 2009 Workshop 'Towards a Standard Markup Language for Embodied Dialogue Acts' (EDAML 2009)*, pages 13–25, Budapest.
- Bunt, H., Alexandersson, J., Choe, J.-W., Fang, A., Hasida, K., Lee, K., Petukhova, V., Popescu-Belis, A., Romary, L., Soria, C., and Traum, D. (2010). towards and ISO standard for dialogue act annotation. In *Proceedings of LREC 2010, Malta*, Paris. ELDA.

- Bunt, H. and Prasad, R. (2016). Iso dr-core (iso 24617-8): Core concepts for the annotation of discourse relations. In *Proceedings 12th Joint ACL-ISO Workshop on Interoperable Semantic Annotation*, pages 45–54, Portoroz, Slovenia.
- Bunt, H. and Romary, L. (2004). Standardization in multimodal content representation: Some methodological issues. In *Proceedings of the 4th International Conference on Language Resources and Evaluation (LREC 2004)*, Lisbon.
- Carletta, J. C., Isard, A., Isard, S., Kowtko, J., Doherty-Sneddon, G., and Anderson, A. (1996). *HCRC Dialogue Structure Coding Manual*. Human Communication Research Centre HCRC TR-82, University of Edinburgh.
- Dhillon, R., Bhagat, S., Carvey, H., and Shriberg, E. (2004). Meeting recorder project: dialogue labelling guide. ICSI Technical Report TR-04-002.
- Di Eugenio, B., Jordan, P., and Pytkkaenen, L. (1998). The COCONUT project: dialogue annotation manual. ISP Technical Report 98-1.
- Geertzen, J., Petukhova, V., and Bunt, H. (2007). A multidimensional approach to utterance segmentation and dialogue act classification. In *Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue*, pages 140–149, Antwerp, Belgium. Association for Computational Linguistics.
- Hindriks, K., Jonker, C., and Tykhonov, D. (2007). Analysis of negotiation dynamics. In et al., M. K., editor, *CIA 2007, LNAI 4676*, page 27–35. Springer-Verlag Berlin Heidelberg.
- Hobbs, J. (1985). On the coherence and structure of discourse. Research Report 85-37, CSLI, Stanford.
- Hovy, E. (1990). Approaches to the planning of coherent text. In Swartout, C. and Mann, W., editors, *Natural Language in Artificial Intelligence and Computational Linguistics*, pages 83–102. Kluwer, Boston.
- Hovy, E. and Maier, E. (1995). Parsimonious of profligate: how many and which discourse structure relations? unpublished manuscript.
- ISO (2006). *TEI-ISO 24610-1:2006 Language resource management: Feature structures, Part 1: Feature structure representation*. Iso, Geneve.
- ISO (2009). *ISO 24612:2009 Language resource management: Linguistic annotation framework (LAF)*. ISO, Geneva.
- ISO (2012). *Language resource management – Semantic annotation framework – Part 2: Dialogue acts. ISO 24617-2*. ISO Central Secretariat, Geneva.
- Jurafsky, D., Shriberg, E., and Biasca, D. (1997). *Switchboard SWBD-DAMSL Shallow-Discourse-Function Annotation: Coders Manual, Draft 13*. University of Colorado.
- Kipp, M. (2001). Anvil - a generic annotation tool for multimodal dialogue. In *Proceedings of Eurospeech 2001*, pages 1367–1370.
- Kipp, M. (2008). Spatiotemporal coding in anvil. In *Proceedings LREC 2008*.

-
- Kipp, M. (2012). Multimedia annotation, querying and analysis in anvil. In *In M. Maybury (ed.) Multimedia Information Extraction*. IEEE Computer Society Press.
- Levinson, S. (1983). *Pragmatics*. Cambridge University Press, Cambridge, UK.
- Mann, W. and Thompson, S. (1988). *Rhetorical structure theory: toward a functional theory of text organisation*. The MIT Press, Cambridge, MA.
- Petukhova, V. (2011). *Multidimensional Dialogue Modelling*. PhD Thesis. Tilburg University, The Netherlands.
- Petukhova, V. and Bunt, H. (2010). Towards an integrated scheme for semantic annotation of multimodal dialogue data. In *Proceedings of the seventh international conference on language resources and evaluation*, pages 2556–2563. Paris: ELRA.
- Petukhova, V., Stevens, C., de Weerd, H., Malchanau, A., and Alexandersson, J. (2016a). Corpus annotation report. METALOGUE Consortium Deliverable 4.1.
- Petukhova, V., Stevens, C., de Weerd, H., Taatgen, N., Cnossen, F., and Malchanau, A. (2016b). Modelling multi-issue bargaining dialogues: Data collection, annotation design and corpus. In *Proceedings 9th International Conference on Language Resources and Evaluation (LREC 2016)*. ELRA, Paris.
- Povey, D. (2011). The kaldi speech recognition toolkit. In *Proceedings of the 2011 IEEE Workshop on Automatic Speech Recognition and Understanding*.
- Prasad, R., Dinesh, N., Lee, A., Miltsakaki, E., Robaldo, L., Joshi, A., and Webber, B. (2008). The Penn Discourse Treebank 2.0. In *Proceedings of the 6th International Conference on Language Resources and Evaluation (LREC 2008)*, Marrakech, Maroc.
- Rohlfing, K., Loehr, D., Duncan, S., Brown, A., Franklin, A., Kimbara, I., Milde, J.-T., Parrill, F., Rose, T., Schmidt, T., Sloetjes, H., and Thies, A. and Wellinghof, S. (2006). Comparison of multimodal annotation tools – workshop report. In *Gespraechsforschung - Online Zeitschrift fuer verbale Interaktion*, volume 2006:7, pages 99 – 123.
- Sanders, T., Spooren, W., and Noordman, L. (1992). Toward a taxonomy of coherence relations. *Discourse Processes*, 15:1–35.
- Searle, J. (1969). *Speech acts*. Cambridge University Press, Cambridge.
- Watkins, M. (2003). Analysing complex negotiations. *Harvard Business Review*.

9 Glossary

9.1 addressee

Dialogue participant oriented to by the speaker in a manner to suggest that his/her utterances are particularly intended for him/her, and that some response is therefore anticipated from him/her, more so than from the other participants

NOTE Source: Goffman (1981)

9.2 annotation

process of adding linguistic information to primary data or the linguistic information itself, independent of its representation

NOTE: the word “annotation” used as a verb refers to the addition of linguistic information to primary data, whereas its use as a noun refers to the linguistic information itself. NOTE Source: ISO Linguistic Annotation Framework ISO 24612:2009

9.3 annotation scheme

Set of descriptors together with their syntax, semantics and condition of use, intended to provide descriptive or interpretive information relevant to a language resource

NOTE Source: Terminology and other language and content resources – Specification of data categories and management of a Data Category Registry for language resources ISO 12620:2009

9.4 communicative function

Property of a **dialogue act**, specifying how the act’s **semantic content** changes the addressee’s information state when (s)he understands the dialogue act

NOTE Source: ISO Semantic Annotation Framework – Dialogue Acts ISO 24617-2

9.5 data Representation of information in a formalized manner suitable for communication, interpretation and processing

NOTE Source: ISO 1087-2:2000

9.6 dialogue

exchange of utterances between two or more participants

NOTE Source: ISO Semantic Annotation Framework – Dialogue Acts ISO 24617-2

9.7 dialogue act

Communicative activity of a participant in dialogue interpreted as having a certain communicative function and semantic content, and possibly also having certain functional dependence relations, rhetorical relations and feedback dependence relations

NOTE Source: ISO Semantic Annotation Framework – Dialogue Acts ISO 24617-2

9.8 dialogue act annotation

Activity of marking up stretches of dialogue with information about the dialogue acts performed, and is usually limited to marking up their communicative functions using a given set of such functions (a ‘tag set’)

9.9 dialogue participant

Person or artificial agent involved in dialogue

NOTE Source: ISO Semantic Annotation Framework – Dialogue Acts ISO 24617-2

9.10 dimension

Aspect of participating in dialogue which can be addressed by **dialogue acts**

NOTE Source: ISO Semantic Annotation Framework – Dialogue Acts ISO 24617-2

A class of dialogue acts concerned with one particular aspect of communication, corresponding to a particular type of semantic content, which a dialogue act can address independently from other dimensions

NOTE Source: Petukhova 2011

9.11 feedback dependence relation

Relation between a feedback act and the stretch of communicative behaviour whose processing the act provides or elicits information about

NOTE Source: ISO Semantic Annotation Framework – Dialogue Acts ISO 24617-2

9.12 functional dependence relation Relation between a dialogue act DA_1 and one or more previous dialogue acts $\{DA_2, \dots, DA_N\}$ iff the meaning of DA_1 depends on the meaning of $\{DA_2, \dots, DA_N\}$ due to the responsive character of DA_1

NOTE Source: ISO Semantic Annotation Framework – Dialogue Acts ISO 24617-2

9.13 functional segment

Minimal stretch of communicative behaviour that expresses one or more **dialogue acts**

NOTE Source: Geertzen et al. 2007

NOTE Source: ISO Semantic Annotation Framework – Dialogue Acts ISO 24617-2

9.14 linguistic annotation

Linguistic information that is added to segments of language data and/or nonverbal communicative behaviour

NOTE Source: ISO Linguistic Annotation Framework ISO 24612:2009

9.15 primary data

Data observed or collected directly from first-hand experience such as representation of written (e.g. text), spoken (e.g. orthographic transcriptions of audio) and multimodal (e.g. images or videos) behaviour

NOTE: do not confuse with primary linguistic data (PLD) in language acquisition

Electronic representation of language

EXAMPLES: text, image, speech signal, etc.

NOTE: Typically, primary data objects are represented by “locations” in an electronic file, for example, the span of characters comprising a sentence or word, or a point at which a given temporal event begins or ends (as in speech annotation). More complex data objects may consist of a list or set of contiguous or non-contiguous locations in primary data. NOTE Source: ISO Linguistic Annotation Framework ISO 24612:2012

9.16 qualifier

Predicate that can be associated with a communicative function

NOTE Source: ISO Semantic Annotation Framework – Dialogue Acts ISO 24617-2

9.17 rhetorical relation

Relation between two dialogue acts, indicating a pragmatic connection between the two

NOTE Source: ISO Semantic Annotation Framework – Dialogue Acts ISO 24617-2

9.18 representation

Format in which the annotation is rendered, independent of its content

NOTE Source: ISO Linguistic Annotation Framework ISO 24612:2012

9.19 semantic content (of a dialogue act)

Objects, propositions, events, actions, and other entities that a dialogue act refers to or uses as arguments of predicates

NOTE Source: ISO Semantic Annotation Framework – Dialogue Acts ISO 24617-2

9.20 semantic content type Kind of objects, events, actions, and other entities that a dialogue act refers to or uses as arguments of predicates, according to some typology of kinds of entities

NOTE See also **dimension**

9.21 signal Variation of a physical quantity used to represent data NOTE Source: ISO/IEC 2382-1:1993

9.22 stand-off annotation

Annotation layered over primary data and serialized in a document separate from that containing the primary data

NOTE Stand-off annotations refer to specific locations in the primary data, by addressing character offsets, elements, etc. to which the annotation applies. Multiple stand-off annotation documents for a given type of annotation can refer to the same primary document (e.g. two different part of speech annotations for a given text)

NOTE Source: ISO Linguistic Annotation Framework ISO 24612:2012

9.23 transcription

Form resulting from a coherent method of writing down speech sounds

NOTE Source: Language resource management – Morpho-syntactic annotation framework (MAF) ISO 24611:2012

Representation of the pronunciation of a given language by the characters of a writing system or by a specially devised system of notations

NOTE: These special systems can be alphabetic [e.g. International Phonetic Association's (IPA) symbols] or analphabetic (e.g. Kenneth Fike's notation in his classic book 'Phonetics', 1943).

9.24 data encoding

Process of putting a sequence of characters (letters, numbers, punctuation, and certain symbols) into a specialized digital format for efficient transmission or transfer

A (concrete) representation of an abstract functional specification in the form of a sequence of data items (e.g. characters, 8-bit bytes, etc.). The identification of each function and its parameter values are encoded to produce a function representation. Similarly, responses from an interpreter of the data stream are also encoded

NOTE Source: ISO/IEC 9636-1:1991

9.25 encoding process Process of converting source sample values into a code stream

NOTE Source: ISO/IEC 29199

9.26 turn

stretch of communicative activity produced by one participant who occupies the speaker role, bounded by periods where another participant occupies the speaker role

NOTE Source: ISO Semantic Annotation Framework – Dialogue Acts ISO 24617-2

9.27 speaker role

role occupied by a dialogue participant who has temporary control of the dialogue and speaks for some period of time

NOTE Source: ISO Semantic Annotation Framework – Dialogue Acts ISO 24617-2

NOTE Source: DAMSL Revised Manual

9.28 utterance

Anything said, written, keyed, gesticulated, or otherwise expressed by a dialogue participant

NOTE Source: ISO Semantic Annotation Framework – Dialogue Acts ISO 24617-2

Linguistically defined contiguous stretches of (linguistic) behaviour. An utterance is the issuance of a sentence, a sentence-analogue, or sentence-fragment, in an actual context

NOTE Source Levinson 1983

10 Annexes

10.1 Annex 1. ISO 24617-2 communicative functions definitions

This Annex provides ISO 24617-2 definitions of communicative functions (Annex D informative). Detailed information about the ISO 24617-2 dialogue act annotation standard and relevant publications are available at http://dit.uvt.nl/#iso_24617-2.

10.1.1 General-purpose functions: information-seeking functions

Conceptual domain	/propositionalQuestion/
Definition	/checkQuestion/ Dialogue act where the speaker, S, wants to know whether a certain proposition is true or false. S assumes that the addressee, A, possesses that information, and puts pressure on A to inform S whether the proposition is true or false.
– Source	Commonplace
– Note	Related terminology in other schemes: QUERY-YN (HCRC MapTask), Yes-No-Question (SWBD-DAMSL) and YNQ (TRAINS).
Explanation	A propositional question corresponds to what is commonly termed a YN-question or polarity question in the literature. The term 'propositional' is preferred because: (a) it clearly separates form from function by removing any oblique reference in the label to syntactic criteria for the identification of such acts; and (b) it is not a language specific term. SWBD-DAMSL for example conflates form and function by distinguishing between propositional questions that are marked explicitly by subject inversion (yes-no questions) and those that are marked by intonation alone (declarative questions). However, though they may have different realisations, these are in fact performing the same function.
Example	"Have you got a haystack on your map?"
–Source	HCRC MapTask

Broader concept	/checkQuestion/
Definition	/propositionalQuestion/ Dialogue act where the speaker, S, wants to know whether a given proposition is true, about which S holds an uncertain belief that it is true S. S assumes that A knows whether the proposition is true or not, and puts pressure on A to provide this information
– Source	Commonplace
– Note	Related terminology in other schemes: Check (DIT, HCRC MapTask), Tag Question (SWBD-DAMSL), Tag (TRAINS) and Request_Comment (Verbmobil)
Example	"The meeting starts at ten, right?"

Conceptual domain	/setQuestion/
Definition	/choiceQuestion/ Dialogue act where the speaker, S, wants to know which elements of a certain set have a named property. S puts pressure on the addressee, A, to provide S with this information. S believes that at least one element of the set has the named property, and S assumes that A knows which are the elements of the set that have the property.
– Source	Commonplace
– Note	Related terminology in other schemes: QUERY-W (HCRC MapTask), WH-Question (SWBD-DAMSL) and WHQ (TRAINS).
Explanation	A set question corresponds to what is commonly termed a WH-question in the literature. The term set is preferred because: (a) it clearly separates form from function by removing any oblique reference in the label to syntactic criteria for the identification of such acts; and (b) it is not a language specific term (it may be further noted that even in English, not all questioning words begin with 'wh', e.g. "How?").
Example	"What time does the meeting start?"

Broader concept	/choiceQuestion/
Definition	/setQuestion/ Dialogue act where the speaker, S, wants to know which one from a given list of alternative propositions is true; S believes that exactly one element of that list is true; S assumes that the addressee, A, knows which of the alternative propositions is true, and S puts pressure on A to provide this information.
– Source	DAMSL; DIT
– Note	Related terminology in other schemes: Alternatives Question (DIT, LIRICS), QUERY-W (HCRC Map-Task), Or-Question/Or-Clause (SWBD-DAMSL, MRDA).
Explanation	It is not very common in annotation schemes to specifically distinguish the concept of choice questions from that of set questions (although it is common in the literature on interrogatives, see for instance: Tsui 1994). However, whereas it is common for the concept set question to carry the expectation that all members of the set with a given property should be returned by the addressee, for a choice-question the expectation is that there will be exactly one. The different preconditions and effects indicate that these are semantically different concepts, and they have been treated here as such.
Example	"Does she live in Amsterdam or in Rotterdam?"

10.1.2 General-purpose functions: information-providing functions

Conceptual domain	/inform/
Definition	/agreement/ /disagreement/ Dialogue act where the speaker, S, wants to make certain information known to the addressee, A; S believes that the information is correct.
– Source	Commonplace
– Note	Related terminology in other schemes: Assert (DAMSL), Explain (HCRC MapTask), Update (LINLIN), Statement (SWBD-DAMSL) and Inform (DIT, TRAINS, Verbmobil).
Explanation	The inform function may also have more specific rhetorical functions such as: explain, elaborate, exemplify and justify, but these all fall under the more generic function here defined.
Example	"The 6.34 to Breda leaves from platform 2."

Definition	/agreement/
– Source	Dialogue act where the speaker, S, wants to inform the addressee, A, that the information which S has reason to believe that A believes is correct, is in fact correct.
– Note	DIT
Explanation	Related terminology in other schemes: Accept (DAMSL, SWBD-DAMSL, TRAINS, Verbmobil). DAMSL and SWBD-DAMSL use "Agreement" to refer to various degrees in which a speaker accepts some previous proposal, plan, opinion or statement; "accept" is one of these degrees; "reject" is another. Note: in this definition /agreement/ inherits the elements in the definition of /inform/.
Example	"Exactly"; Dutch "Precies!"; Danish: "Netop!"
–Source	DIT

	/disagreement/
Conceptual domain	/correction/
Definition	Dialogue act where the speaker, S, wants to inform the addressee, A, that the information which S has reason to believe that A believes is correct, is in fact incorrect.
– Source	DIT
– Note	Related terminology in other schemes: Reject (DAMSL, SWBD-DAMSL, Verbmobil) and Denial (TRAINS).
Explanation	DAMSL and SWBD-DAMSL use "Agreement" to refer to various degrees in which a speaker accepts some previous proposal, plan, opinion or statement; "accept" is one of these degrees; "reject" is another. Note: in this definition /disagreement/ inherits the elements in the definition of /inform/.
Example	"I'm afraid you're wrong."

	/correction/
Definition	Dialogue act where the speaker, S, wants to inform the addressee, A, that the information which S has reason to believe that A believes is correct, is in fact incorrect and should be replaced by the information that S offers.
– Source	Commonplace
– Note	In this definition /correction/ inherits the elements in the definition of /disagreement/.
Example	"To Montreal, not to Ottawa."

	/answer /
Broader concept	/inform/
Conceptual domain	/confirm/ /disconfirm/
Definition	Dialogue act where the sender, S, wants to make certain information available to the addressee, A, which S believes A wants to know.
– Source	Commonplace

	/confirm/
Broader concept	/answer/
Definition	Dialogue act where the sender, S, believes that the addressee, A, wants to know whether his (A's) uncertain belief that the information queried by a check is correct.
– Source	DIT; Verbmobil
– Note	Related terminology in other schemes: Reply-Y (HCRC MapTask), Yes-Answer (SWBD-DAMSL).
Example	"Indeed"

	/disconfirm/
Broader concept	/answer/
Definition	Dialogue act where the sender, S, believes that the addressee, A, wants to know whether his (A's) uncertain belief that the information queried by a check is incorrect.
–Source	DIT
–Note	Related terminology in other schemes: Reply-N (HCRC MapTask) and No-Answer (SWBD-DAMSL).
Example	French "Si"; Danish "Jo"; Dutch: "Toch niet" and "Toch wel"

10.1.3 General-purpose functions: commissives

Conceptual domain	/promise/
Definition	/acceptRequest/ /declineRequest/ Dialogue act where the sender, S, commits himself unconditionally to perform a certain action in the manner or with the frequency described. S believes that the addressee, A, prefers that the action be performed (rather than not be performed).
– Source	DIT, Searle (1969)
– Note	Related terminology in other schemes: Promise (TRAINS)
Example	"I will send you an email"

Definition	/offer/ Dialogue act where the sender, S, commits himself to perform a certain action, conditional on A's consent that S do so.
– Source	DAMSL; DIT
– Note	Related terminology in other schemes: Offer (TRAINS).
Example	"Shall I start?"; "Would you like to have some coffee?"

Definition	/addressRequest/ Communicative function of a dialogue act by which the sender, S, commits himself to perform an action that he was requested to perform, possibly with certain conditions concerning manner or frequency of performance.
– Source	DIT
– Note	The addressRequest function covers a range of possible responses to a request. If no condition is attached, then the speaker commits himself unconditionally to perform the requested action; this is the special case of /acceptRequest/. If the condition is attached that the action be performed zero times, then the speaker in fact declines to perform the requested action. See also the data categories for the qualifiers /conditional/ and /partial/.
Example	"Maybe"

Definition	/acceptRequest/ Dialogue act where the sender, S, commits himself to perform an action that was requested.
– Source	DIT
– Note	Related terminology in other schemes: Accept (DAMSL, SWBD-DAMSL, TRAINS, Verbmobil).
Example	"Sure"

Definition	/declineRequest/ Dialogue act where the sender, S, indicates unwillingness to perform an action that was requested.
– Source	DIT
– Note	Related terminology in other schemes: Reject (DAMSL, SWBD-DAMSL, TRAINS, Verbmobil).
Example	"Not now"

Definition	/addressSuggest/ Communicative function of a dialogue act by which the sender, S, commits himself to perform an action that was suggested to him, possibly with certain restrictions or conditions concerning manner or frequency of performance.
– Source	DIT
– Note	Related terminology in other schemes: Accept (DAMSL, SWBD-DAMSL, TRAINS, Verbmobil).
Example	"Let's have this in mind"

	/acceptSuggest/
Definition	Communicative function of a dialogue act by which the sender, S, commits himself to perform an action that was suggested to him.
– Source	DIT
– Note	Related terminology in other schemes: Accept (DAMSL, SWBD-DAMSL, TRAINS, Verbmobil).
Example	"Let's do that"

	/declineSuggest/
Definition	Communicative function of a dialogue act by which the sender, S, indicates unwillingness to perform an action that was suggested to him.
– Source	DIT
– Note	Related terminology in other schemes: Reject (DAMSL, SWBD-DAMSL, TRAINS, Verbmobil).
Example	"I don't think so"

10.1.4 General-purpose functions: directives

	/instruct/
Broader concept	/request/
Conceptual domain	/acceptOffer/ /declineOffer/
Definition	Dialogue act where the sender, S, wants the addressee, A, to carry out a named action in the manner or with the frequency described; S assumes that A is able and willing to carry out the action.
– Source	DIT; HCRC Map Task
– Note	Related terminology in other schemes: Action-directive (DAMSL).
Example	"Go right round until you get to just above that."
–Source	HCRC MapTask

	/suggest/
Definition	Dialogue act where the sender, S, wants the addressee, A, to be aware that a named action is potentially promising for achieving a certain goal, which is either named explicitly or contextually salient.
– Source	DIT; TRAINS; Verbmobil
– Note	Related terminology in other schemes: Open-option (DAMSL).
Example	"Let's wait for the meeting to finish."

	/request/
Conceptual domain	/instruct/
Definition	Dialogue act where the sender, S, wants the addressee, A, to perform a named action in the manner or with the frequency described, conditional on A's consent.
– Source	DIT; TRAINS; Verbmobil
Example	"Please turn to page five"

	/addressOffer/
Definition	Dialogue act where the sender, S informs the addressee, A, that S is committed to perform the action described in the semantic content dependent on S's consent to do so.
– Source	DIT
– Note	Related terminology in other schemes: Accept (DAMSL, SWBD-DAMSL, TRAINS, Verbmobil).
Example	"I will think about it"

	/acceptOffer/
Definition	Dialogue act where the sender, S, informs the addressee, A, that S agrees to A performing the action that A has offered to perform.
– Source	DIT
– Note	Related terminology in other schemes: Accept (DAMSL, SWBD-DAMSL, TRAINS, Verbmobil).
Example	"Yes please"; French: "Je vous en prie"

	/declineOffer/
Definition	Dialogue act where the sender, S, informs the addressee, A, that S does not agree to A performing the action that A has offered to perform.
– Source	DIT
– Note	Related terminology in other schemes: Reject (DAMSL, SWBD-DAMSL, TRAINS, Verbmobil).
Example	"No thanks"

10.1.5 Feedback functions

	/autoPositive/
Definition	Dialogue act where the sender, S, wants the addressee A to know that S believes that S's attention to, perception, interpretation, evaluation or execution of the previous utterance was successful.
– Source	DIT
– Note	Related terminology in other schemes: Signal-Understanding (DAMSL), Acknowledge (HCRC Map-Task, SWBD-DAMSL) Ack (TRAINS) and Feedback_Positive (Verbmobil). This type of feedback could be further broken down into more specific areas (dealing with the sender's attention, perception, interpretation, evaluation and execution), as exemplified in the DIT schema. Such fine distinctions have hitherto not been made in other annotation schemes however, so a simplified top level data category is defined here.
Example	"Uh-huh"; Nonverbally: nodding; "Yes"

	/alloPositive/
Definition	Dialogue act where the sender, S, wants the addressee, A, to know that S believes that A's attention to, perception, interpretation, evaluation or execution of the previous utterance was successful.
– Source	DIT
– Note	The distinction between whether feedback is about S's (auto) understanding or A's (allo) is only made within the DIT scheme. This type of feedback could be further broken down into more specific areas (dealing with the addressee's attention, perception, interpretation, evaluation and execution).
Example	"You got that right"

Definition	/autoNegative/ Dialogue act where the sender, S, wants the addressee, A, to know that S believes that S's attention to, perception, interpretation, evaluation or execution of the previous utterance encountered a problem.
– Source	DIT
– Note	Related terminology in other schemes: Signal-Non-Understanding (DAMSL) and Feedback _Negative (Verbmobil). This type of feedback could be further broken down into more specific areas (dealing with the sender's attention, perception, interpretation, evaluation and execution), as is exemplified in the DIT schema. Such fine distinctions have hitherto not been made in other annotation schemes however, so a simplified top level data category is defined here.
Example	"Sorry?"; "What?"; Spanish: "Que?"; Italian: "Como?"

Definition	/alloNegative/ Dialogue act where the sender, S, wants the addressee, A, to know that S believes that A's attention to, perception, interpretation, evaluation or execution of the previous utterance encountered a problem.
– Source	DIT
– Note	The distinction between whether feedback is about S's (auto) understanding or A's (allo) is only made within the DIT scheme. This type of feedback could be further broken down into more specific areas (dealing with the addressee's attention, perception, interpretation, evaluation and execution).
Example	"No no no no no"

Definition	/feedbackElicitation/ Dialogue act where the sender, S, wants to know whether A's attention to, perception, interpretation, evaluation or execution of the previous utterance was successful.
– Source	DIT
– Note	Feedback elicitation could be further broken down into more specific areas dealing with the addressee's attention, perception, interpretation, evaluation and execution.
Example	"Okay?"; Italian: "Capisce?"

10.1.6 Turn management functions

Definition	/turnAccept/ Dialogue act where the sender, S, agrees to take the turn when he is requested to do so.
– Source	DIT
– Note	Related terminology in other schemes: Take-Turn (TRAINS).
Example	A: "Would you like to say something at this point?" B: "Certainly."; Nonverbally: nodding

Definition	/turnAssign/ Dialogue act where the sender, S, wants the addressee, A, to take the turn.
– Source	Common in literature on turn taking in conversation
– Note	Occurs especially in multiparty dialogue. Related terminology in other schemes: Assign-Turn (TRAINS).
Example	"Adam?", characteristically accompanied by the speaker directing his gaze to Adam, possibly also nodding or pointing in his direction and raising the eyebrows.

	/turnGrab/
Definition	Dialogue act where the sender, S, wants to take the turn from another participant.
– Source	DIT
– Note	Related terminology in other schemes: Take-Turn (TRAINS).
Example	"Hold on"

	/turnKeep/
Definition	Dialogue act where the sender, S, wants to keep the turn.
– Source	DIT
– Note	Related terminology in other schemes: Keep-Turn (TRAINS).
Explanation	Utterances used for turn keeping often also have a stalling function.
Example	"Erm"

	/turnRelease/
Definition	Dialogue act where the sender, S, wants to give other dialogue participants the opportunity to take the turn
– Source	Common in literature on turn taking in conversation
– Note	Related terminology in other schemes: Release-Turn (TRAINS).
Example	Sender uses declining intonation towards the end of a contribution and subsequently pauses.

	/turnTake/
Definition	Dialogue act where the sender, S, wants to take the turn when it is available.
– Source	DIT
– Note	Related terminology in other schemes: Take-Turn (TRAINS)
Example	"Ehm..." as a turn-initial segment

10.1.7 Time management functions

	/stalling/
Definition	Dialogue act where the sender, S, wants to have a little more time to construct his contribution.
– Source	DIT
– Note	Related terminology in other schemes: Hold (DAMSL).
Example	"Let me see...", "Erm..."; Nonverbally: slowing down
–Note	Utterances used for stalling often also have a turn keeping function.

	/pausing/
Definition	Dialogue act where the sender, S, wants to suspend the dialogue for a while because he needs some time to do something.
– Source	DIT
– Note	Related terminology in other schemes: Hold (DAMSL).
Explanation	Pausing occurs either in preparation of continuing the dialogue, or because something else came up which is more urgent for the sender to attend to.
Example	"Just a moment"; Dutch: "een ogenblikje"

10.1.8 Own and partner communication management functions

	/completion/
Definition	Dialogue act where the sender, S, wants to assist the addressee, A, by finishing or adding to the clause that A is in the middle of constructing.
– Source	DAMSL; DIT; TRAINS
Example	S: "which should leave us plenty of time to uhhh", A: "get to city H"
– Source	TRAINS

	/correctMisspeaking/
Definition	Dialogue act where the sender, S, wants to correct (part of) an utterance by the addressee, A, assuming that A made a speaking error.
– Source	DAMSL; DIT
– Note	Related terminology in other schemes: Correction suggestion (TRAINS).
Example	S: "second engine E3 is going to uhh city H to pick up the bananas, back to A, drop...", A: "to pick up the oranges", S: "sorry, pick up the oranges"
–Source	TRAINS

	/selfError/
Definition	Dialogue act where the sender, S, wants the addressee, A, to know that S has made a mistake in speaking.
– Source	DIT
Example	"We're going out on Tues- no, er, not on Tuesday"

	/selfCorrection/
Definition	Dialogue act where the sender, S, wants to correct an error that he made, or to improve on an infelicitous formulation that he used, within the same turn.
– Source	Common in literature on conversation studies
– Note	Related terminology in other schemes:
Example	"We're going out on Tues- no, er, Thursday"

	/retraction/
Definition	Communicative function of a dialogue act performed by the sender, S, in order to withdraw something that he just said within the same turn.
Example	"then we're going to gÅŞ- to go"

10.1.9 Discourse structuring functions

	/interactionStructuring/
Definition	Dialogue act where the sender, S, wants to explicitly indicate to the addressee, A, the function or topic of his next contribution(s).
– Source	DIT
– Note	Interaction structuring covers such phenomena as topic introduction, dialogue act announcement and topic closing.
Examples	"Question"

Definition	/opening/ Dialogue act where the sender, S, wants the addressee, A, to know that S is ready and willing to engage in a dialogue with A.
– Source	DIT
Example	"Okay ..."

Definition	/closing/ Dialogue act where the sender, S, wants the addressee, A, to know that S is ready and willing to close a dialogue with A.
– Source	DIT
Example	"That's it"

Definition	/topicIntroduction/ Dialogue act where the sender, S, wants to introduce the topic mentioned in the semantic content.
– Source	DIT
Example	"Now we will discuss taxes"

Definition	/topicShiftAnnouncement/ Dialogue act where the sender, S, wants to change the topic.
– Source	DIT
Example	"Moving further to the scope"

Definition	/topicShift/ Dialogue act where the sender, S, wants to shift the topic to the one mentioned in the semantic content.
– Source	DIT
Example	"Antismoking campaign"

10.1.10 Social obligations management functions

Definition	/initGreeting/ Dialogue act where the sender, S, wants the addressee, A to know that S is present and aware of A's presence; S puts pressure on A to acknowledge this.
– Source	DIT
– Note	Related terminology in other schemes: Greet (Verbmobil).
Explanation	Greetings usually come in initiative-response pairs within a dialogue; this data category corresponds to the first element of such a pair.
Example	"Hello!"; "Good morning"

	/returnGreeting/
Definition	Dialogue act where the sender, S, wants to acknowledge that S is aware of the presence of the addressee, A, and of A having signalled his presence to S; S has been pressured to respond to an initialGreeting by A.
– Source	DIT
– Note	Related terminology in other schemes: Greet (Verbmobil).
Explanation	Greetings usually come in initiative-response pairs within a dialogue; this data category corresponds to the second element of such a pair.
Example	"Hello!"; "Good morning"

	/initSelfIntroduction/
Definition	Dialogue act where the sender, S, wants to make himself known to the addressee, A; S puts pressure on A to acknowledge this.
– Source	DIT
Explanation	Introductions usually come in initiative-response pairs within a dialogue; this data category corresponds to the first element of such a pair.
Example	"I'm Jack"

	/returnSelfIntroduction/
Definition	Dialogue act where the sender, S, wants to make himself known to the addressee, A; S has been pressured to respond to an initialSelfIntroduction by A.
– Source	DIT
Explanation	Introductions usually come in initiative-response pairs within a dialogue; this data category corresponds to the second element of such a pair.
Example	"And I'm Jill"

	/apology/
Definition	Dialogue act where the sender, S, wants the addressee, A, to know that S regrets something; S puts pressure on A to acknowledge this.
– Source	DIT; SWBD-DAMSL
Example	"Sorry about that."

	/acceptApology/
Definition	Dialogue act where the sender, S, wants to mitigate the addressee, A's feelings of regret; S has been pressured to respond to an apology by A.
– Source	DIT
Example	"No problem."

	/thanking/
Definition	Dialogue act where the sender, S, wants the addressee, A, to know that S is grateful for some action performed by A; S puts pressure on A to acknowledge this.
– Source	DIT
– Note	Related terminology in other schemes: Thank (Verbmobil).
Explanation	Utterances used for thanking often also indicate that the sender wants to end the dialogue.
Example	"Thanks a lot."

Definition	/acceptThanking/ Dialogue act where the sender, S, wants to mitigate or respond to the addressee, A's feelings of gratitude; S has been pressured to respond to an act of thanking by A.
– Source	DIT
Example	"Don't mention it"

Definition	/initGoodbye/ Dialogue act where the sender, S, wants the addressee, A, to know that S intends the current utterance to be his final contribution to the dialogue; S puts pressure on A to acknowledge this.
– Source	DIT
– Note	Related terminology in other schemes: Bye (Verbmobil).
Explanation	Goodbyes usually come in initiative-response pairs within a dialogue; this data category corresponds to the first element of such a pair.
Example	"Bye, see you later"

Definition	/returnGoodbye/ Dialogue act where the sender, S, wants to acknowledge his awareness that the addressee, A, has signalled his final contribution to the dialogue and S signals in return his agreement to end the dialogue; S has been pressured to respond to an initialGoodbye by A.
– Source	DIT
– Note	Related terminology in other schemes: Bye (Verbmobil).
Explanation	Goodbyes usually come in initiative-response pairs within a dialogue; this data category corresponds to the second element of such a pair.
Example	"Bye, see you later"

10.1.11 Contact management functions

Definition	/contactIndication/ Dialogue act where the sender, S, wants to make it known to the addressee, A, that S is ready to send messages to, and receive messages from, A.
– Source	DIT
Example	"Yes?"

Definition	/contactCheck/ Dialogue act where the sender, S, wants to establish whether the addressee, A, is ready to receive messages from, and send messages to, S.
– Source	DIT
Example	"Hello?!"

10.1.12 Other functions

Conceptual Domain	/directQuestion/ <i>/set Question/ /propositional Question/ /alternative Question/ /check Question/</i>
Definition	Dialogue act where the sender, S, wants to know something which S assumes the addressee, A, to know. S puts pressure on A to provide this information
– Source	Commonly used as contrasting with <i>/indirectQuestion/</i>

10.2 Annex 2. Rhetorical (discourse) relations

Definition – Source	/elaboration/ Arg1 and Arg2 are the same situation, but Arg2 contains more detail. ISO 24617-8
Definition – Source	/exemplification/ Arg1 describes a set of situations; Arg2 describes an element of that set. ISO 24617-8
Definition – Source	/contrast/ One or more differences between Arg1 and Arg2 are highlighted with respect to what each predicates as a whole or to some entities they mention. ISO 24617-8
Definition – Source	/cause/ Arg1 is an explanation for Arg2. ISO 24617-8
Definition – Source	/synchrony/ Some degree of temporal overlap exists between Arg1 and Arg2. All forms of overlap are included. ISO 24617-8
Definition – Source	/asynchrony/ Arg1 temporally precedes Arg2. ISO 24617-8
Definition – Source	/condition/ Arg1 is an unrealized situation which, when realized, would lead to Arg2. ISO 24617-8
Definition – Source	/negativeCondition/ Arg1 is an unrealized situation which, when <i>not</i> realized, would lead to Arg2. ISO 24617-8
Definition – Source	/purpose/ Arg1 enables Arg2. ISO 24617-8

Definition – Source	/manner/ Arg1 describes how Arg2 comes about or occurs. ISO 24617-8
Definition – Source	/concession/ An expected causal relation between Arg1 and Arg2 is canceled or denied by Arg2. ISO 24617-8
Definition – Source	/exception/ Arg2 indicates one or more circumstances in which the situation(s) described by Arg1 does not hold. ISO 24617-8
Definition – Source	/similarity/ One or more similarities between Arg1 and Arg2 are highlighted with respect to what each predicates as a whole or to some entities they mention. ISO 24617-8
Definition – Source	/substitution/ Arg1 and Arg2 are alternatives, with Arg2 being the favored or chosen alternative. ISO 24617-8
Definition – Source	/disjunction/ Arg1 and Arg2 bear the same relation to some other situation evoked in the discourse, explicitly or implicitly. Their disjunction indicates that they are alternatives with respect to that situation, with the disjunction being non-exclusive so that both Arg1 and Arg2 may hold. ISO 24617-8
Definition – Source	/conjunction/ Arg1 and Arg2 bear the same relation to some other situation evoked in discourse. Their conjunction indicates that they both hold with respect to that situation. ISO 24617-8
Definition – Source	/restatement/ Arg1 and Arg2 describe the same situation, but from different perspectives. ISO 24617-8
Definition – Source	/expansion/ Arg2 is a situation involving some entity/entities in Arg1, expanding the narrative forward of which Arg1 is a part, or expanding on the setting relevant for interpreting Arg1. The Arg1 and Arg2 situations are distinct. ISO 24617-8

Definition – Source	/motivation/ Comprehending Arg2 increases the reader's desire to perform the action presented in Arg1. Hovy and Maier (1995)
Definition – Source	/instantiation/ Arg2 is a situation involving some entity/entities in Arg1, narrowing the narrative forward of which Arg2 is a part, or specifying the setting relevant for interpreting Arg1.
Definition – Source	/justification/ Arg2 justifies the reader's intention to perform the action presented in Arg1. Arg2 describes why Arg1 comes about or occurs. Hovy and Maier (1995)
Definition – Source	/explanation/ Arg2 explains situation described in Arg1. Arg2 describes specific circumstances of Arg1 occurrence. DIT
Definition – Source	/evaluation/ Arg2 conveys assessment of the situation described in Arg1. Hovy and Maier (1995)
Definition – Source	/conclude/ ArgN + 1 contains a judgment or decision reached by reasoning from Arg1, Arg2,..., ArgN. 'summary' in Hovy and Maier (1995)

10.3 Annex 3. Negotiation Moves

Definition Example	/elicitOfferValue/ Speaker wants the partner to express his commitment to a certain value What do you prefer? What do you want? What about ...? How do you feel about ...
Definition Example	/offerValue/ Speaker expresses his commitment to a certain value (utility value) I want no change in tobacco taxes. I can go for just public transportation
Definition Example	/addressOfferValue/ Speaker is considering to commit to a certain value dependent on condition specified in the semantic content It might be possible. I will have this option in mind.

Definition	/acceptOfferValue/ Speaker believes that he will commit to the value
Example	That's fine for me. Yes, that's it
Definition	/declineOfferValue/ Speaker believes that he will not commit to the value
Example	No it's not possible for me. I don't like that
Definition	/counterOfferValue/ Speaker proposes an alternative value towards which he has stronger commitment than one expressed by the partner in his/her previous offer
Note	This does not necessary imply that the Speaker rejects partner's previous offer
Example	P1: I was going to say no change at all P2: I go twenty five then if you're so bad
Definition	/exchangeValues/ Speaker proposes a certain value in exchange for another value or he commits to a certain value conditionally on the addressee's consent on another value typically of a different type/in a different issue
Example	If you can give me ten percent taxes, I would agree on all outdoor smoking allowed.
Definition	/concession/ Speaker acknowledges a potential or apparent incompatibility between his and partner's offers but either (1) regards partner's offer as possible/compatible in the light of his goals (= below his max. goal/sub-optimal option) or (2) proposes his incompatible offer seeking for concession
Example	Well I think it's not good for you but no smoking should be allowed in public transportation, parks, and open air events
Definition	/bargainIn/ Speaker is committed to alter his previous offer by going up
Note	This might be the case that Speaker noticed that his initial offer was too low and he may gain more
Example	I want more, I want all outdoor smoking banned
Definition	/bargainDown/ Speaker is committed to alter his previous offer by going down (= below his max. goal/sub-optimal option)
Example	Okay, I can bargain down here
Definition	/deal/ Speaker believes that both speaker and partner are committed to certain values and are able to wrap up negotiations
Example	We have a deal

Definition	/noDeal/ Speaker believes that both speaker or partner are not committed (refrain to commit) to certain values and wants to close negotiation
Example	No agreement here.. We don't have deal on...

Definition	/withdraw/ Speaker withdraws all his previous offers concerning a certain value (or values) and as a consequence may terminate negotiation
Example	I change my mind, I want ...

Definition	/breakDown/ Speaker states that no act or process is possible; negotiation is failing to function or continue
Example	I believe we will not reach any agreement here, we better stop

Definition	/exitDeal/ Speaker cancels the previously achieved deals/agreements with respect to a certain value
Example	We need to re-discuss

10.4 Annex 4. Speech production disfluency types

Definition	/substitution/ Substitution disfluency repeats some parts of the Reparandum in the Reparans, new information substitutes some parts that would be necessary for a grammatically correct sentence.
-------------------	---

Definition	/insertion/ This is the opposite of the Deletion as it adds material to the Reparans while repeating the Reparandum.
-------------------	--

Definition	/deletion/ The Deletion repeats some parts of the Reparandum in its Reparans while omitting the rest. The deleted parts must be from the center of the Reparandum.
-------------------	--

Definition	/omission/ Omissions are cases where the speaker omitted a word that would be necessary for a grammatically correct sentence.
-------------------	---

Definition	/repetition/ These denote expressions where the Reparandum and the Reparans have equal content. This holds for words but not for word fragments.
-------------------	--

Definition	/replacement/ The Replacement disfluency repeats some parts of the Reparandum in the Reparans, new information substitutes the rest.
-------------------	--

Definition	<p>/restart/</p> <p>The Restart contains completely new material in the Reparans. The disfluency must not be at the beginning of the utterance.</p>
Definition	<p>/order/</p> <p>Denotes disfluent speech in which the word order has to be changed in order to have a correct sentence.</p>
Definition	<p>/mistake/</p> <p>A Mistake is an error that leads to a grammatically incorrect sentence which was not corrected by the speaker.</p>

10.5 Annex 5. Guidelines on using ANVIL annotation tool

1. Install ANVIL tool following instructions on <http://www.anvil-software.org/download/index.html>
2. Download Metalogue ANVIL specification file from the Specification folder (example given also below)
3. Open your working .anvil file: if you work with existing annotations (e.g. correct or view them) from Data/Annotations/Anvil, Anvil may ask for new updated paths to specification and video files, browse your folders and select correct files as shown in the Figure 12 NOTE:Anvil requires videos, since Metalogue MIB dialogue sessions were not video recorded we created black (fake) videos with original flac files using VirtualDub free software <https://sourceforge.net/projects/virtualdub/> and compress them using codecs supported by Anvil, e.g. Cinepak. Created black screen videos are not provided with the corpus.

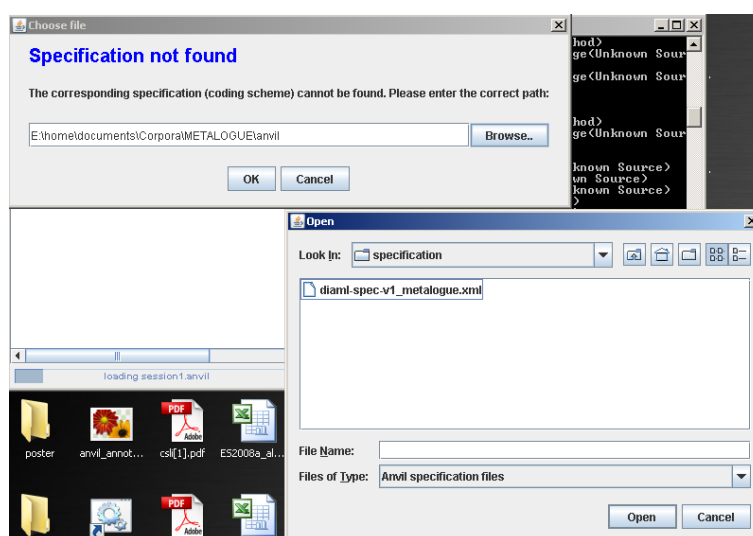


Figure 12: Selection of the correct specification file in ANVIL

Alternatively you could open the annotation file in any text editor and specify the paths manually in the head of the file, for example:

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<annotation>
  <head>
    <specification src="D:/Corpora/Metalogue/Metalogue_MIB-corpus/Specifications/
Metalogue_MIB_Anvil_specification.xml" />
    <video src="D:/Corpora/Metalogue/Negotiations/Pre-Pilot/Annotations/session1/
pilot0-casel.avi" master="true" />
  </head>
```

4. ANVIL will load the .anvil file with windows for control, video, track summary and annotation like shown in Figure 9 in Section 7 of this Manual
5. Perform dialogue act annotation according to the guidelines highlighted in Section 4.8 by first deciding on the right dimension (Anvil track) then marking the start and end of an identified segment and choosing 'Create&Edit' option with right-mouse-click, 'add element' window will appear as shown in Figure 13

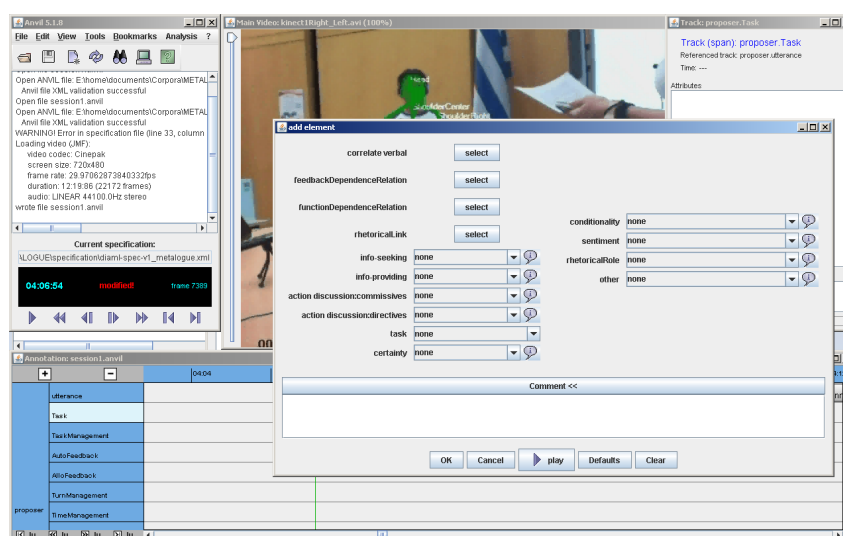


Figure 13: Dialogue act annotation with ANVIL.

6. Select the communicative behaviour to which you would like to attach a communicative function label by clicking 'correlate verbal' and choosing relevant elements from the 'utterance' track as illustrated in Figure 14
7. Choose then the right communicative function label from drop-down lists as shown in Figure 15
8. Consult help for all defined conceptual tags when necessary by clicking on 'i' symbol as shown in Figure 16

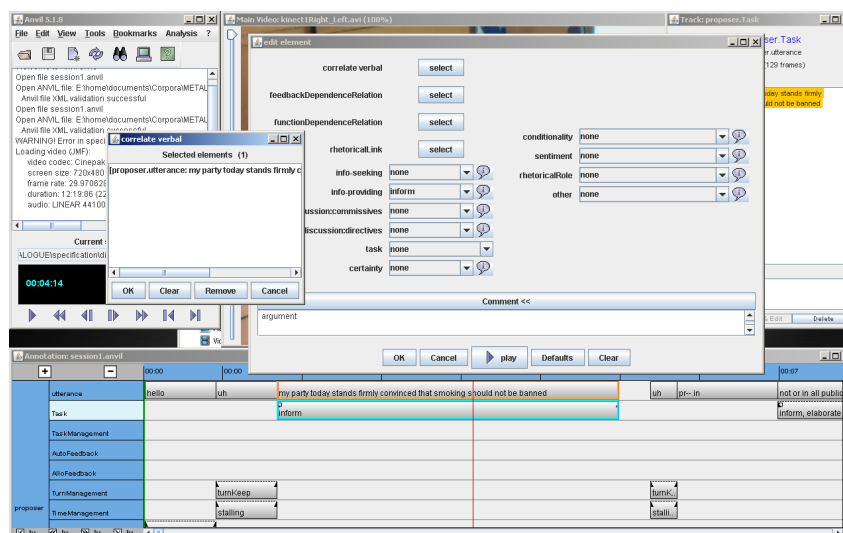


Figure 14: Dialogue segmentation in ANVIL.

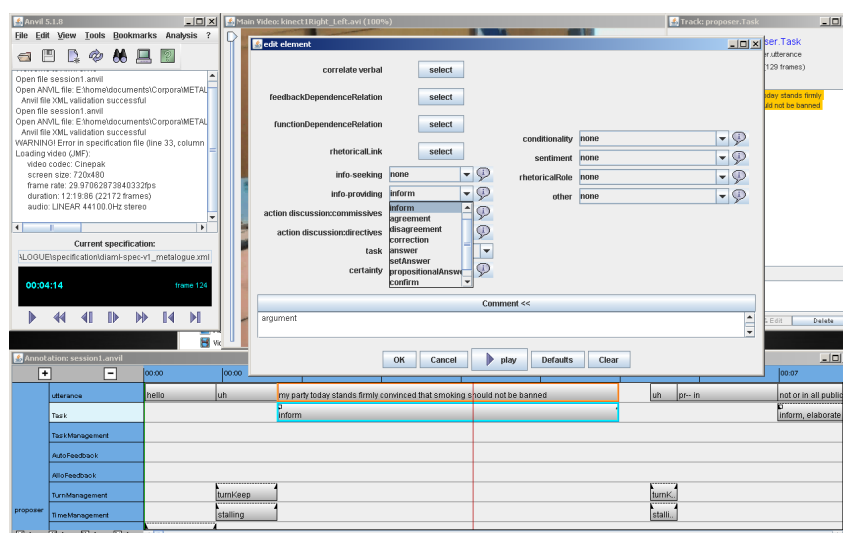


Figure 15: Communicative function labeling with ANVIL.

9. Do not forget to annotate links between segments such as functional and feedback dependence, and rhetorical links by clicking corresponding 'Select' buttons and choosing relevant elements/segments out of annotated or transcribed behaviour, i.e. 'utterance' or annotated dialogue act elements in one of the 'dimension' tracks for feedback and functional dependence, and rhetorical relations. Note that for rhetorical relation you additionally need to specify the type of relation by choosing values out of 'RhetoricalRole' list.
10. Repeat steps 5 to 9 for new identified segments
11. Save modified annotations

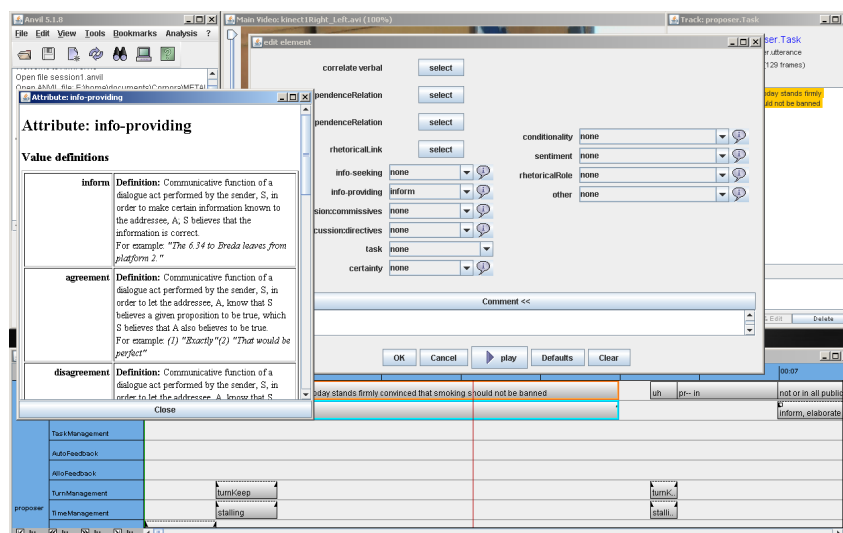


Figure 16: Using help for communicative function definitions in ANVIL.

10.5.1 Example of Metalogue specification for dialogue act annotations using Anvil annotation tool

Specification is provided with this corpus distribution

```
<?xml version="1.0" encoding="UTF-8"?>
<!--
```

Specification for annotation of extended ISO 24617-2 dialogue acts and negotiation moves. Written by Volha Petukhova – Spoken Language System Group, Saarland University, Germany
email: v.petukhova@lsv.uni-saarland.de.

Feel free to use in your projects and modify (see remark at the very end of the file).

Please see and cite the following paper at LREC 2016:

Volha Petukhova, Christopher Stevens, Harmen de Weerd, Niels Taatgen, Fokie Cossen and Andrei Malchanau. (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. ELRA, Paris

-->

```
<annotation-spec>
```

```
<!-- HEAD -->
```

Below is the example of specification for Metalogue *task* track grouped per Metalogue dialogue participant, i.e. cityCouncilor. 'Utterance' track in a group is a primary track. The dimensional tracks are span tracks and are meant to assign a communicative function(-) in the specified dimension, one track per dimension to make multidimensional segmentation possible.

```
<group name="cityCouncilor">
  <track-spec name="utterance" type="primary">
    <attribute display="true" name="token" valuetype="String" />
  </track-spec>

  <track-spec name="task" ref="cityCouncilor.utterance" type="span">
    <attribute link-color="orange" name="correlate verbal"
      valuetype="MultiLink" />
    <attribute link-color="blue" name="feedbackDependence"
      valuetype="MultiLink" />
    <attribute link-color="blue" name="functionalDependence"
      valuetype="MultiLink" />
    <attribute link-color="blue" name="rhetoricalLink"
      valuetype="ReciprocalLink" />
    <attribute display="true" name="info-seeking"
      valuetype="informationSeekingFunctions" />
    <attribute display="true" name="info-providing"
      valuetype="informationProvidingFunctions" />
    <attribute display="true" name="action discussion:commissives"
      valuetype="commisiveFunctions" />
    <attribute display="true" name="action discussion:directives"
      valuetype="directiveFunctions" />
    <attribute display="true" name="certainty"
      valuetype="CertaintyQualifiers" />
    <attribute display="true" name="conditionality"
      valuetype="ConditionalityQualifiers" />
    <attribute display="true" name="sentiment"
      valuetype="SentimentQualifiers">
      <doc>The ISO standard does not specify the values of the sentiment
        attribute. These values are only example values</doc>
    </attribute>
    <attribute display="true" name="rhetoricalRole" valuetype="RhetRole" />
  <attribute display="true" name="other" valuetype="other" />
</track-spec>
```

10.6 Annex 6. Transcription format specifications

10.6.1 Plain text formats

Manually produced transcriptions with timestamps information are stored. The targeted tab separated .txt files contain three fields as follows (1) 'start time'; (2) 'end time' and (3) string of

tokens spanned into a functional segment:

```
1.67 5.97 my party today stands firmly convinced that smoking should
      not be banned
7.97 10.08 not or in all public areas
10.08 15.68 we do conceive that certain public areas maybe that smoking
      should be banned
15.68 18.45 however we do not stand convinced that it should be banned
      in all public areas
18.85 21.99 my party will be trying to prove today two main things
22.72 24.72 that it should not be banned in all public areas
25.66 30 that even if we do conceive that it can be banned then that
      will not be effective
31.03 34.3 i will be talking to you about three fundamental ideas
35 36.5 about the impact to the individual
37.74 39.07 about the actual effect on this
```

10.6.2 XML-based format: ANVIL and Text Encoding Initiative format (TEI)

Transcriptions were automatically converted to .anvil transcription files of the following format:

```
<track name="system.utterance" type="primary">
  <el index="0" start="0" end="0.41">
    <attribute name="token">SIL</attribute>
  </el>
  <el index="1" start="0.41" end="0.8">
    <attribute name="token">alright</attribute>
  </el>
  <el index="2" start="0.8" end="1.09">
    <attribute name="token">SIL</attribute>
  </el>
  <el index="3" start="1.09" end="1.63">
    <attribute name="token">hello</attribute>
  </el>
  <el index="4" start="1.63" end="1.98">
    <attribute name="token">again</attribute>
  </el>
  ...
```

Transcriptions and meta-data are provided in TEI compliant format. The latter enables data exchange since it is supported by many annotation tools, e.g. ELAN and EXMARaLDA.

The TEI header contains corpus meta-data.

```
<?xml version="1.0" encoding="ISO-8859-1"?>
<dialogueSession xmlns="http://www.metalogue.eu/content_spec">
  <TEI xmlns="http://www.tei-c.org/ns/1.0" xml:lang="en" version="5.0">
    <teiHeader>
      <fileDesc>
```

```

<titleStmt>
  <title type="full">
    <title type="main">Metalogue Multi-Issue Bargaining Corpus</title>
    <title type="sub">Pilot 1, Case 4</title>
  </title>
  <author>
    ...
  </author>
<funder>European Commission, Grant Agreement 611073</funder>
</titleStmt>
<publicationStmt>
  <authority>
    <orgName>Spoken Language Group, Saarland University, Germany</orgName>
  </authority>
  <publisher>distributed by Linguistic Data Consortium</publisher>
  <availability>
    <licence>Available under LDC licence only<date from="2017"/>
  </licence>
</availability>
</publicationStmt>
<sourceDesc>
  <recordingStmt>
    ...
  </recordingStmt>
</sourceDesc>
</fileDesc>
<profileDesc>
  <particDesc xml:id="p1">
    <person role="cityCouncilor"/>
  </particDesc>
  <particDesc xml:id="p2">
    <person role="businessRepresentative"/>
  </particDesc>
</profileDesc>
</teiHeader>

```

Further, meta-data is followed by transcribed utterances including timeline that links elements.

```

<text>
  <timeline xml:id="TL01" unit="s">
    <when xml:id="TW0" absolute="00:00:00"/>
    <when xml:id="TWSp10" interval="2.06869" since="#TW0"/>
    <when xml:id="TWEp10" interval="2.56918" since="#TW0"/>
    <when xml:id="TWSp11" interval="2.56918" since="#TW0"/>
    <when xml:id="TWEp11" interval="3.30323" since="#TW0"/>
    <when xml:id="TWSp12" interval="3.30323" since="#TW0"/>

```

```

...
</timeline>

<body>
  <said who="#p1">
    <u xml:id="wp10" start="#TWSp10" end="#TWEp10">YEAH</u>
    <u xml:id="wp11" start="#TWSp11" end="#TWEp11">I AGREE</u>
    <u xml:id="wp12" start="#TWSp12" end="#TWEp12">\$LAUGHTER</u>
    <u xml:id="wp13" start="#TWSp13" end="#TWEp13">YEAH</u>
    <u xml:id="wp14" start="#TWSp14" end="#TWEp14">FIFTEEN</u>
    <u xml:id="wp15" start="#TWSp15" end="#TWEp15">I THINK</u>
    <u xml:id="wp16" start="#TWSp16" end="#TWEp16">THAT'S</u>
    <u xml:id="wp17" start="#TWSp17" end="#TWEp17">THAT'S</u>
    <u xml:id="wp18" start="#TWSp18" end="#TWEp18">A LOT</u>
    <u xml:id="wp19" start="#TWSp19" end="#TWEp19">I THINK IT CAN
      CAUSE LOT OF TROUBLES TO SOME BUSINESSES</u>
    <u xml:id="wp110" start="#TWSp110" end="#TWEp110">FIFTEEN</u>
  ...
  </said>
  <said who="#p2">
    <u xml:id="wp20" start="#TWSp20" end="#TWEp20">SMOKING SHOULD
      BE ALLOWED</u>
    <u xml:id="wp21" start="#TWSp21" end="#TWEp21">OKAY</u>
    <u xml:id="wp22" start="#TWSp22" end="#TWEp22">AND</u>
    <u xml:id="wp23" start="#TWSp23" end="#TWEp23">UH</u>
    <u xml:id="wp24" start="#TWSp24" end="#TWEp24">I THINK FOR
      THE TAXATION TO HAVE SOME EFFECT</u>
    <u xml:id="wp25" start="#TWSp25" end="#TWEp25">IT SHOULD
      BE AT LEAST</u>
    <u xml:id="wp26" start="#TWSp26" end="#TWEp26">FIVE</u>
    <u xml:id="wp27" start="#TWSp27" end="#TWEp27">FIFETEN
      PERCENT</u>
  ...
  </said>

```

As a result of annotation process utterance are segmented into functional segments as follows

```

<spanGrp xml:id="vesp1TSKCV0" type="functionalVerbalSegment">
  <span xml:id="tsp1TSKCV011" from="#wp10"/>
  <span xml:id="tsp1TSKCV012" from="#wp11"/>
</spanGrp>
<fs xml:id="fsp1TSKCV0" type="functionalSegment">
  <f name="verbalComponent" fVal="#vesp1TSKCV0"/>
</fs>
<spanGrp xml:id="vesp1TSKCV1" type="functionalVerbalSegment">
  <span xml:id="tsp1TSKCV111" from="#wp13"/>
</spanGrp>

```

```
<fs xml:id="fsp1TSKCV1" type="functionalSegment">  
  <f name="verbalComponent" fVal="#vesp1TSKCV1"/>  
</fs>
```