

# Dimensions of communication

Petukhova, V.V.; Bunt, H.C.

*Publication date:*
2009

Tilburg centre for Creative Computing
Tilburg University
http://www.uvt.nl/ticc

P.O. Box 90153
5000 LE Tilburg, The Netherlands
Email: ticc@uvt.nl



**April 17, 2009**

**TiCC TR 2009–003**

# Dimensions of communication

**Volha V. Petukhova** **Harry C. Bunt**

TiCC, Tilburg University

## Abstract

This study is concerned with the identification and analysis of dimensions of communication in dialogue, with the aim to provide theoretical and empirical arguments for chosing the dimensions in the ISO standard for dialogue act annotation 24617-2 "Semantic annotation framework, Part 2: Dialogue acts". A 'dimension' in this context is a cluster of semantically related communicative functions which has a conceptual, theoretical and empirical significance. Five criteria are put forward for including a particular dimension in a multidimensional annotation schema: it should be (1) theoretically justified; (2) empirically observed; (3) recognizable by human annotators and by machine; ( 4) addressable independently of other dimensions; and (5) reflected in a significant number of existing dialogue act schemes.

Eight dimensions are identified which fulfil all five criteria, and can be considered as 'core' aspects of dialogue communication, namely Task, Feedback, Turn Management, Social Obligation Management, Discourse Structuring, Own Communication Management, Partner Communication Management, and Time Management. Contact Management is proposed to be given the status of an optional additional dimension.

# Dimensions of communication

**Volha V. Petukhova** **Harry C. Bunt**

TiCC, Tilburg University

## 1 Introduction

The research presented in this report has been carried out within project 24617-2 “Semantic annotation framework, Part 2: Dialogue acts” of the International Organisation for Standards ISO. This project aims to develop an international standard for the annotation of dialogues with dialogue act information, in order to support the creation of interoperable and reusable language resources ([54]). In line with the design of the most widely used existing dialogue act annotation schemas, the project takes a multidimensional approach to dialogue act annotation. This study is concerned with the identification and analysis of dimensions of communication as reflected in existing annotation schemas and theoretical models, with the aim to provide considerations and criteria for making well-founded choices of the dimensions in the standard that the ISO project aims to establish.

## 2 The notion of ‘dimension’

Multidimensional approaches to dialogue act annotation have their origin in the view that utterances in dialogue are often multifunctional, serving multiple purposes at the same time (see e.g. [5]; [19]). When annotating the utterances in a dialogue with information about the communicative acts that are performed, they should therefore be marked up with multiple tags.

The most frequently used multidimensional annotation scheme is DAMSL (Dialogue Act Markup using Several Layers ([4]). DAMSL distinguishes four so-called layers: Communicative Status, Information Level, Forward-Looking Function (FLF) and Backward-Looking Function (BLF); the last two are concerned with communicative functions. The FLF layer is subdivided into five classes, including (roughly) the classes of commissive and directive functions, well known from speech act theory. The BLF layer has four classes: Agreement, Understanding, Answer, and Information Relation. In [38] Core and Allen also refer to these eleven classes as dimensions.

Clustering related communicative functions, rather than using a flat lists of tags, has the advantage of making the annotation schema more transparant. Even more important is that a well-designed multidimensional annotation schema makes the possible multifunctionality of dialogue utterances explicit, by defining its dimensions in such a way that an utterances can maximally have one function in each dimension. Existing multidimensional schemas for dialogue act annotation have mostly not explicitly motivated their choice of dimensions, Usually, a dimension is formed by a set of tags corresponding to communicative functions that are (intuitively) semantically related and mutually exclusive.

In [23] it was shown that this approach to multidimensionality is unsatisfactory in several respects. For example, if the cluster of information-seeking functions for a range of question types and the cluster of information-providing functions for various kinds of informs and answers are considered as dimensions (as in the DAMSL schema), then an utterance may be tagged as having both information-seeking and information-providing, which is conceptually impossible since one cannot (for example) question the truth of a given proposition and state that it is true. ohence they address different communicative aspects, e.g. question about task domain and the answer about the processing of the previous utterance(-s). Also, consisting of mutually exclusive tags is not a good criterion for defining a dimension either, since some functions within one dimension may form specializations of more general functions. For example,

a warning is a special case of an inform; a check is a special kind of question; and a confirmation is a special kind of answer.

Popescu-Belis in [71] argues that dialogue act tag sets should seek a multidimensional theoretical grounding and defines the following aspects of utterance function that could be relevant for choosing dimensions in a multidimensional schema: (1) the traditional clustering of illocutionary forces in speech act theory into five classes: Representatives, Commissives, Directives, Expressives and Declarations; (2) turn management; (3) adjacency pairs; (4) topical organization in conversation; (5) politeness functions; and (6) rhetorical roles.

To arrive at a well-designed multidimensional annotation schema, It is essential to have a clear picture of what constitutes a theoretically and empirically satisfactory set of dimensions. In [22], Bunt proposed the following definition of the notion of a dimension in dialogue act analysis (see also [25]

(1) A dimension is an aspect of participating in dialogue which can be addressed:

- through linguistic and/or nonverbal behaviour that has a communicative function for this specific purpose;
- independently of addressing other aspects for which sets of communicative functions are distinguished (i.e., other dimensions).

The two criteria mentioned in this definition are necessary conditions for distinguishing a dimension; for choosing useful dimension, considerations of theoretical and empirical relevance should be added. We propose that each dimension in dialogue act scheme should be:

(2)
1. theoretically justified;
2. empirically observed in communicative functions of dialogue utterances;
3. recognizable by human annotators and by machine;
4. addressable independent of other dimensions.

Moreover, for the particular purpose of designing a dialogue act annotation standard that is useful for researchers in dialogue and designers of dialogue systems, an additional requirement is:

(3) 5. the dimension should be reflected in a significant number of existing dialogue act schemes.

This report aims to provide theoretical and empirical evidence motivating the choice of dimensions in a multi dimensional schema as a proposed ISO standard for dialogue act annotation.

# 3 Method

To address the requirements listed in (2) and (3), we studied the most influential and widely cited works of researchers in the area of dialogue modelling, and analysed 18 existing well-known dialogue act annotation schemes (see Section 5). For the latter we benefited from the work done in the MATE[1] [58] and [59], and LIRICS[2] [26] projects, which aimed to provide standards for various areas of language technology, including dialogue act annotation.

For the empirical evidence relating to communicative dimensions we analysed the following dialogue corpora:

- the DIAMOND corpus)[3] which consists of two-party human-human task-oriented instructional spoken dialogues in Dutch;

[1] **M**ulti level **A**nnotation **T**ools **E**ngineering

[2] Linguistic InfRastructure for Interoperable ResourCes and Systems (http://lirics.loria.fr)

[3] For more information about the project see Jeroen Geertzen, Yann Girard, and Roser Morante. 2004. The diamond project. Poster at the 8th Workshop on the Semantics and Pragmatics of Dialogue(CATALOG 2004).

- the AMI meeting recordings corpus[4] which consists of multimodal task-oriented human-human multi-party dialogues in English;
- the OVIS corpus[5] which consists of task-oriented human-computer dialogues over the telephone in Dutch.

The DIAMOND dialogues were orthographically transcribed; 952 utterances representing 1,408 functional segments from the human-human subset of the corpus were selected. The AMI data contain 17,335 words, which form 3,897 functional segments with an average length of 4.4 words (average turn length is 7.7 segments). The OVIS corpus contains 3942 functional segments. All corpora were manually tagged using the DIT$^{++}$ annotation scheme[6] in multiple dimensions. We analysed the distribution of the tags that were used in various communicative dimensions, and discuss the results of dialogue act recognition experiments which have been reported in [44] and [46].

In order to investigate the last defined criteria some dependency tests are performed and results are reported in 8. Section 9 outlines some discussion issues and draws conclusions.

# 4 Theoretical validation

Multidimensional approaches to dialogue act annotation, which incorporate a multifunctional view on dialogue behaviour, have been recognised by many researchers as empirically better motivated, and allowing the modeling of theoretical distinctions (e.g. [59], [61], [71], etc.). Studies of human dialogue behaviour indicate that natural dialogue involves several activities beyond those strictly related to performing the task or activity for which the dialogue is instrumental (such as obtaining certain information, instructing another participant, negotiating an agreement, etc.). In natural conversation, dialogue participants among other things constantly 'evaluate whether and how they can (and/or wish to) continue, perceive, understand and react to each other's intentions' [7]. They share information about the processing of each other's messages, elicit feedback, monitor contact and attention and manage the use of time, allocation of turns, contact and attention, etc. Communication is thus a complex, multi-faceted activity, and dialogue utterances are therefore most of the time multifunctional. A dialogue act tagset should contain the concepts needed to cover all these aspects of dialogue.

Popescu-Belis in [71] argues that dialogue act tagsets should seek a multidimensional theoretical grounding. The presence and definition of each dimension as a communicative aspect in dialogue should be theoretically justified. We studied the most influential and widely cited works and theories of researchers working in the area of dialogue understanding and modelling, to see what aspects of the interaction are considered and investigated, such as Bales [15] for a general account of interaction, Allen [1] among others for *plan-based approaches*, Clark [32], Traum [81], and Allwood [10] and [7] for *collaborative joint activity* models, Sacks et al. [74] for *conversational* analysis, Mann and Thompson [66] and Asher and Lascarides [14] for *rhetorical relations* in discourse.

## 4.1 Dialogue purpose and domain of discourse

Dialogues are usually motivated by goals, tasks, or activities which are non-communicative in their nature, e.g. to obtain certain information, to solve a problem, to improve relationships, to act in a game as team mates, and so on. Allen in [1] assumes that people are rational agents capable of forming and executing plans to achieve their goals and they are also capable of inferring the plans of other agents from observing their actions. Rationality is analysed by [7] in terms of adequate (efficient) and competent action. People communicate with the aim to achieve something and they do this in a rational fashion [19], organising the interaction so as to optimise the conditions for successful communication.

[4] **A**ugmented **M**ulti-party **I**nteraction (`http://www.amiproject.org/`)

[5] **O**penbaar **V**ervoer **I**nformatie **S**ystem (Public Transport Information System) http://www.let.rug.nl/v̈annoord/Ovis/

[6] For more information about the tagset visit: http://dit.uvt.nl/

### 4.2 Contact, presence, and attention

A basic requirement on communication is that the parties are in ***contact*** and are willing to be in continued contact [11]. 'If A attempts to communicate with B, he/she can expect B to respond, at least by indicating that no contact is possible, and any response from B is enough to manifest contact' [11]. For some types of dialogue this aspect of communication is of a particular importance, namely when there is no or limited visual contact between the participants. For example, telephone conversations are dependend on the quality of the communication channel. But also when dialogue participants have direct visual contact, they tend to permanently check the attention of their interlocutors and their readiness to continue the conversation. Participants utilise both their bodies and facial expressions (e.g. gaze is used to ensure contact between participants) and a variety of vocal phenomena to show the attention they are giving to the events of the moment and, reciprocally, the type of reaction they expect from others [49].

### 4.3 Grounding and feedback

To be successful, participants in a dialogue have to coordinate their activities on many levels other than that of the underlying task or activity. The coordination of knowledge and beliefs is a central issue in any communication, the basic coordination problem being that of building mutual or shared beliefs out of individual ones. Clark in [32] argues that speakers and addressees attempt to establish the mutual belief that the addressee has understood what is uttered. The process of establishing mutual understanding of each others intentions and actions is called ***grounding***. Traum in [81] proposes to distinguish a class of grounding acts; which are directly related to ***feedback***. Feedback is generally considered as an essential instrument for successful communication. Allwood in [7] agues that feedback morphemes and mechanisms, whether they occur as a single utterance or as a part of a large utterance, are probably the most important cohesion device in spoken language. Feedback mechanisms, their linguistics (verbal and non-verbal expressions, durational, temporal and prosodic properties) and related phenomena have been studied extensively, e.g. [42], [10], [34]. Bales in [15] noticed that dialogue participants address several levels of processing of the partner's previous utterances, taking each other into cognitive consideration and showing readiness to communicate, giving attention and receptiveness, recognition, interest and responsiveness to the partner's contribution(-s). Thus, feedback may be reported on various levels. Allwood in [10], Clark in [32] and Bunt in [21] distinguish several ***feedback levels***: attention (in [10] called contact), perception (in [32] called identification), understanding (in [21] called interpretation), evaluation (in [32] called consideration and in [10] attitudinal reaction), and execution defined in [21].

Another important aspect of feedback functions according to Allwood is their direction [7]. The speaker in dialogue may provide feedback on his own processing of the partner's previous utterance(-s) (***feedback giving*** functions or ***auto-feedback***, in terms of [19]), or elicit feedback when he wants to know the processing status of the addressee (***feedback eliciting*** functions, or a part of ***allo-feedback***, in the terminology of [19], which is concerned with the addressee's processing of the speaker's utterance(-s)). In [34] it was noticed that addressees in dialogue cooperate by displaying and signalling their understanding, but speakers also monitor their addressees for understanding and, when necessary, alter their utterances or elicit feedback.

### 4.4 Taking turns

Another essential aspect of any interactive conversation is ***turn management***. Allwood ([7]) defines turn management as the distribution of the right to occupy the sender role in dialogue. He argues that this is rather normative than a behavioural unit. Accordingly, the decision to take the next turn or to offer the next turn to the partner(-s) depends on the speaker's needs or motivations and beliefs, and on the rights and obligations in a conversational situation. People do not start up talking just anywhere and do not just stop talking without any reason. 'Doing conversation' is behaving according to certain orderly procedures ([74]). In the widely quoted study of Saks, Schegloff and Jefferson ([74]) the following manifestations of turn-taking in human-human communication are observed:

1. Speaker change recurs, or at least occurs.
2. Overwhelmingly, one party talks at a time.

3. Occurrences of more than one speaker at time are common, but brief.
4. Transitions with no gap or overlap are common; together with transitions with a slight gap or overlap they form the majority of transitions.
5. Turn order is not fixed, but varies.
6. Turn size is not fixed, but varies.
7. Length of conversation is not fixed in advance.
8. What parties say is not fixed in advance.
9. Relative distribution of turns is not specified in advance.
10. Number of parties can vary.
11. Talk can be continuous, or discontinuous.
12. Turn allocation techniques are obviously used. Either the speaker selects the next speaker by addressing him or her, or speakers may self-select.
13. Various turn-constructional units are employed (word, phrase, sentence).
14. Repair mechanisms exist for dealing with turn-taking errors and violations. In particular, if two parties find themselves talking at the same time, one of them will stop.

In [74] Transition Relevance Places (TRPs) are defined as points where the turn is yielded to another participant, the following rules are formulated:

1. If the current speaker (S) selects the next speaker (N) in the current turn, S is expected to stop speaking, and N is expected to speak next.
2. If S's utterance or behaviour does not select the next speaker, then any other participant may self-select. Whoever speaks first gets the floor.
3. If no speaker self-selects, S may continue.

Recent years have seen a number of solid qualitative and quantitative findings on turn-taking mechanisms and related phenomena, analysing the ways dialogue participants indicate that they intend to start speaking, finish speaking, resume speaking, or give the right to speak to someone else; e.g. [30], [31], [76], [17].

### 4.5 Social obligations and politeness

Participating in a dialogue is a social activity, where one is supposed to do certain things and not to do others, and to act in accordance with the norms and conventions regulating social behaviour. Each participant in dialogue not only has functional but also ethical tasks and obligations, and performs ***social obligation*** acts to fulfill these. The golden rule of ethics *'Do unto others what you would have them do unto you'* means in communication *'make it possible for others to be rational, motivated agents'* [11]. Bales in [15] pays a lot of attention to social obligation acts such as *acts of active solidarity and affection*, *status-raising acts* and *acts for giving help and reward.*

Social obligation acts are closely related with politeness phenomena. Lakoff in [60] formulates three politeness rules:

1. don't impose (a speaker who acts according this rule will avoid, mitigate, or ask permission or apologize for making the addressee do anything which the addressee does not want to do);

2. offer options (speaker should express himself in such way that his opinion or request can be ignored without being contradicted or rejected, e.g. the use of indirect speech acts rather than direct ones);

3. encourage feelings of camaraderie (in general to show active interest in the other and his opinion).

Brown and Levinson's Theory of Politeness [18] influenced most work on politeness and linguistic style. The key idea is that speakers are polite in order to save the hearer's *face*: a public self-image that every person wants to pursue. The concept of face is divided into positive face, the need for a person to be approved of by others, and negative face, the need for autonomy from others. All in all, people communicate with each other according to the norms and conventions for pleasant and comfortable

interaction [19]. People commonly employ in dialogues so-called 'politeness acts': greetings, apologies, expression of gratitude, valediction, etc. Bunt [20] noticed that social obligation acts are not just 'social', they are also useful for improving the conversational transparency of the dialogue. For example, people greet each other to establish their presence, and say good-bye to close the conversation; they often apologise when interrupting another speaker, and so on.

### 4.6 Dialogue structure

Dialogue participants may at several dialogue stages indicate their view of the state of the dialogue and make the hearer acquainted with his plans for the continuation of the conversation. The speaker can give indications that he is going to close the discussion of certain topic(s); or that he wants to concentrate the hearer attention on a new topic. ***Dialogue structuring*** acts are based on the speaker's view of the present linguistic context, on his plan for continuing the dialogue, and on the assumed need to structure the discourse for his partner.

Organization of discourse structure is extensively studied by [66], formulating Rhetorical Structure Theory for monologues; by [14] doing something similar for dialogues, for argumentative dialogues ([35]), for interviews ([75], [52]) and for dialogues that are highly interactive in nature and are characterized by rapid turn switching among participants, such as task-oriented dialogues ([51]). Some researchers distinguish macro-, meso- and micro-levels in discourse structuring (e.g. [67] and [62]). The micro-level is concerned with relations within a turn or within a single utterance, such as rhetorical relations; the meso-level is about the relations within a subdialogue, e.g. common ground units; and the macro-level is concerned with topic structure and plan-based analysis, topic shifts, opening and closing of dialogue, etc.

Studies have also been made of nonverbal behaviour as clues for structuring the discourse. Cassell et al. ([31]), for example, studied posture-shift, gaze, and hand and head movements in correlation with the start of a new discourse segment, turn management behaviour, and information structure (e.g. emphasizing certain information).

### 4.7 Speech production and editing

An aspect of communication which has been addressed in the literature as well as extensively studied from a practical point of view in the context of designing spoken dialogue systems, concerns the speaker's speech production and monitoring.. Speakers continuously monitor the utterance that is currently being produced or prepared to produce [34], and when problems or mistakes are discovered, they stop the flow of the speech and signal to the addressee that there is trouble and that a repair follows (*error signalling*). A speaker may make mistakes in verbal fluency, e.g. stuttering, or mispronouncing words and may wish to reformulate a part of his utterance or to start from the beginning of the phrase within the same turn (*retractions*). Retractions frequently occur at the beginning of an utterance and within other hesitations and phrasal breaks. Sometimes a speaker just repeats a phrase or part of it without reformulations within the same turn (*restart* or *refresh*), and this may have several reasons. When the speaker has produced a (partial) result, recognises that he made an error, and corrects it within the same turn one speaks of *self-correction*.

In [64] seven reasons for repairs are mentioned:

- lexical error or flaw in formulation, e.g. 'For example if you needed to add a voice recognition then your user interface would be *split* **broken down** into more components which you have a microphone the VR and stuff like that';
- syntactical or morphological errors, for example, word ordening, agreement, etc., e.g. 'What *I'm* **I'd** be a bit worried about is if someone *was* **had** previously developed habits of expecting to control surround sound';
- sound form errors, tongue slips, e.g. 'And then the desired *devi* **design** will consist in specifically implementing and detailing the choice we've made in the second';
- articulation errors, such as speaking too loud or too fast;
- dialogue act errors, e.g. '*are there any like* ***what are our options*** **Is this the only way that we go about it or are there other thin**';

- speaking style errors, and also errors in choice of social register, according to social standards;
- conceptual errors, e.g. more information should be provided, an ambiguity should be avoided, etc., e.g. '*They* find them ugly **Most people** find them ugly'.

Garret in [43] argues that speech errors can be corrected by deletion (a unit is missed out from the intended target), preservation (a unit occurs both in the right place and later in the utterance), exchange (two units are swapped), blend (two units are combined), substitution (a word is substituted for a different word) or cognitive intrusions (units from outside the message level are inserted into the utterance).

According to Allwood et al. [12] Own Communication Management (OCM) is concerned with how a speaker continuously manages the planning and execution of his/her own communication, and is a basic function in dialogue. Partner Communication Management (PCM) is concerned with monitoring the partner's speech by the speaker, either providing assistance by completing an utterance that the partner is struggling to complete (*completion*), or correcting (part of) a partner's utterance, believing that partner made a speaking error (*correct-misspeaking*).

### 4.8 Timing

In dialogue conversation fluent speech is rare [33]. Another aspect of communication which is concerned with disfluent speech production is ***time management***, where the speaker suspends the dialogue for one of several reasons and resumes it after minor (*stalling*) or prolongned (*pause*) delay. Delays take place at all major levels of planning - from retrieving a word to deciding what to talk about next [33], in other terms 'micro-' (e.g. word searching problems) and 'macro-structure' delays (uncertainty [77], new topic introduction [79] or turn-keeping [78]) [48].

According to Clark's *theory of performance* [32] speakers in dialogue proceed along two tracks of communication simultaneously: (1) primary track referring to the task or topic of the dialogue; and (2) collateral track referring to the performance itself - to rephrasing, mistakes, repairs (own communication management), intentions to speak (turn management), timing, delays (time management), and the like. Clark notices that time delays can be signalled by modifying a syllable, word or phrase within a primary utterance, e.g. prolongned syllables, non-reduced words; by using filled and silent pauses, e.g. 'um' and 'uh'; and by using other modalities, e.g. certain head nods, eye gaze, over-speech laughter, pointing, etc. (studied by [16] and [49] among others). Criticising Maclay and Osgood [65], Clark shows that stalling acts are not simply ways of holding the floor but signal imminent delays. He analysed monologues and observes that in monologue there is no issue of holding the floor, yet stalling acts are used just as in dialogues.

### 4.9 Concluding observations

To sum up, in the literature several aspects of communication are addressed, which involve several activities beyond those strictly related to performing the motivating **task** or **activity**, notably the actions concerned with the processing of each other's messages, giving and eliciting feedback (**auto-feedback** and **allo-feedback**), managing the use of **time**, the allocation of **turns**, **contact**, difficulties in the speaker's utterance production (**own communication management**), or those of other interlocutors (**partner communication management**), structuring the dialogue (**dialogue structuring**), and giving attention to social aspects of the interaction (**social obligations management**). In the next section we investigate to what extent these aspects of communication are reflected in existing dialogue act annotation schemes.

## 5 Dimension related concepts in existing DA annotation schemes

### 5.1 Task and Task Management

Multidimensional dialogue taxonomies, such as **DAMSL**, **MRDA**, **DIT++** and **LIRICS**, define a *Task* dimension for those dialogue acts that relate directly to the performance of the task (or 'activity') that motivates the dialogue. **DAMSL** has two separate dimensions for this aspect, *Task* and *Task Management* ('about task' in **MRDA** and **SWBD-DAMSL**). The latter explicitly addresses the way in which

the task is performed and interpreted. The **MRDA** category 'about-task' covers similar information applied to meetings, and is defined as 'reference to meeting agendas or direction of meeting conversation'. It was, however, noticed in [40] that it is often difficult to distinguish between Task Management and Communication Management, or Task Management and Task, especially for dialogues which involve solving a problem or developing a plan. Indeed, the observed agreement and annotation accuracy on the **DAMSL** Task Management dimensions are low. We performed small-scale annotation experiments with 5 naive annotators (non-linguistic undergraduate students) who had been introduced to the **DAMSL** annotation scheme and the underlying theory as part of a course, and who were asked to annotate a dialogue from the TRAINS corpus (about 20 utterances). The observed agreement between the annotators on this task was 72%, but the annotation accuracy was only 42%. The Task Management dimension is clearly difficult to apply, and even though annotators reached quite good agreement between each other, they agreed on wrong choices, as displayed in annotation accuracy scores. Task Management was very often confused with Communication Management or Task.

One-dimensional schemes invariably address the Task dimension in their tagsets. In fact, the majority of the communicative functions in most annotation schemes are meant to be used for the Task dimension.

The Task dimension is usually addressed using *information related* (information-seeking and information -providing) and *action related* functions (commissive and directive). Some schemes define categories which are specific to a particular task or domain. For example, the **Coconut** scheme, which applies the multidimensional approach defined in **DAMSL**, has some domain-specific tags related to furniture items (*needItem, getItem, haveItem*, etc.).

## 5.2 Feedback

*Feedback* is an important aspect of communication. This is reflected in almost all existing dialogue act taxonomies except **Linlin** [39] and **Primula** [70]. In **DAMSL** [4] and schemes based on DAMSL such as **Switchboard-DAMSL** [55], **Coconut** [40] and **MRDA** [41] various levels of feedback are defined, ranging from merely hearing what was said to identifying the speaker's intention. T he functions *signal-understanding* and *signal-non-understanding* are available for coding successes and failures in perception and interpretation of the partner's utterance(-s) (see Table 1) . The *acknowledgment* function signals that the previous utterance was understood without necessarily signalling acceptance, and *repeat-rephrase* (except for [41]) is used to signal that the previous speaker has been understood, but like acknowledgments, no further commitment is made as to whether the responder agrees with or believes the antecedent. **SWBD-DAMSL** and **MRDA** have one more feedback function, called *assessment/appreciation* which express the speaker's evaluation, emotional involvement, or support of what the partner has said, e.g. *'That would be nice'*. **SWBD-DAMSL** has also *summarize-reformulate* as a feedback function, which is used when a speaker is proposing a summarization or paraphrase of what was said by another speaker. To code expressions of negative auto-feedback **MRDA** defines an *understanding check*, for when the speaker checks whether he correctly understands what the previous speaker said, and *repetition request* when a speaker was unable to perceive or interpret another speaker's previous utterance and wishes to hear that portion again, e.g. *'Please repeat'*. **Coconut** defines *clarification request*, which can be used for signalling understanding failures by the speaker.

The **AMI** scheme [13] defines the *assess* function to express evaluative feedback, and is comparable to the *assessment/appreciation* of **SWBD-DAMSL** and **MRDA**. **AMI** also has *backchannels* as special cases which are not really dialogue acts but which are labelled in order to avoid gaps in the annotation, and signal that someone who has just been listening to a speaker says something in the background, without stopping that speaker. *Backchannels* signal that what the speaker has just said presents no difficulty to the person who utters the backchannel, so that the speaker can continue. *Backchannels* defined in the **Verbmobil** scheme [3] are more comparable with **DAMSL** *acknowledgments* and are used to signal understanding, acknowledging successful communication without expressing acceptance, rejection, or (dis)agreement. *Acknowledgments* are also defined in the **HCRC Maptask** scheme [28] for a verbal response which minimally shows that the speaker has heard the utterance to which it responds. **Verbmobil** defines other feedback functions (which in Verbmobil are not considered as dialogue control

| DIT | Pos. attention | Pos. percep-tion | Pos. interpreta-tion | Pos. evaluation | Pos. execu-tion | Neg. atten-tion | Neg. percep-tion | Neg. interpre-tation | Neg. evalua-tion | Neg. execu-tion |
|---|---|---|---|---|---|---|---|---|---|---|
| **LIRICS** | Positive auto-feedback | | | | | Negative auto-feedback | | | | |
| **DAMSL** | Signal understanding | | Acknow-ledgment | | | Signal-non-understanding | | | | |
| **SWBD-DAMSL** | Signal understanding | | Acknow-ledgment Repeat-rephrase | Assessment | Summarize-reformulate | Signal-non-understanding | | | | |
| **MRDA** | Signal understanding | | Acknow-ledgment | Assessment Apprecia-tion | | Signal-non-understanding: repetition request | | Understan ding check | | |
| **Coconut** | Signal understanding | | Acknow-ledge Repeat-rephrase | | | Signal-non-understanding | | Clarificatio n check | | |
| **AMI** | Comment-about-understanding POS | | | Assess | Inform POS | Comment-about-understanding NEG | | | | Inform NEG |
| **HCRC MapTask** | | | Acknow-ledgment | | | Check | | | | |
| **Verbmobil** | backchannel | | | Acknow-ledgement | Pos. feedback | Request clarify | | | | Neg. feed-back |
| **SLSA** | Pos. contact | Pos. percepti on | Pos. understan-ding | Pos. acceptance/ attitude | | Neg. contact | Neg. percep-tion | Neg. understand ing | Neg. attitude | |
| **TRAINS** | Acknowledgment | | | Pos. evaluation | | | | Neg. evaluation | | |
| **SPAAC** | echo | | Acknow-ledge | appreciate | | pardon | | | | |
| **MALTUS** | Pos. attention | Repeat-rephrase | | appreciation | | Neg. attention | | | | |
| **Chiba** | Follow up: pos. understand | | | Pos. response | | Follow up: neg. understand | | | Neg. response | |
| **Alparon** | | | Acknow-ledgment | | | | | | | |
| **C-Star** | | | Acknow-ledge | | | | | | | |

Table 1: ***Auto-feedback*** *communicative functions in different dialogue act taxonomies.*

acts but belong to the 'Task-Promote' layer) such as *reject*, *explain_reject*, *accept* and *confirm*. Feedback at the level of execution can be labelled in **AMI** using the Inform function plus a relevant relation tag (e.g. NEGative or POSitive). To code expressions of auto-feedback at levels of perception and interpretation, **AMI** has *comment-about-understanding* where the speaker can indicate either that he did understand (or simply hear) what a previous speaker said, or that he didn't.

In the **TRAINS** scheme [2] and [80], grounding acts are defined which address feedback phenomena, such as *acknowledgment*, which signals understanding of a previous utterance and includes (1) repetition or paraphrase of all or part of the utterance; (2) backchannel responses; and (3) implicit signalling of understanding by initiating a new unit, e.g. an answer to a question. Acknowledgments are confirmations or acceptances (agreements). In **SLSA** [68] feedback aspects are part of the interaction communication management dimension. A distinction is made between *giving* and *eliciting* feedback at the levels of *contact*, *perception* and *understanding*, which are comparable to the levels defined in **DIT** [19] and [23] as *attention*, *perception* and *interpretation*. Additionally, **SLSA** defines *acceptance* attitudes, which imply the successful execution of the previous utterance, e.g. acceptance to carry out a request, or acceptance of a turn. Emotional acceptance attitudes are also tagged, such as surprise, anger, happiness, etc.

The **SPAAC** [63] annotation scheme defines three communicative functions for positive feedback, namely *echo* (in which the speaker simply echoes or 'parrots' something the other person said in a preceding turn, generally to make sure that what that speaker said has been correctly heard and decoded), *acknowledgement* (a backchannel, signalling that the speaker is following or taking on board what the other speaker is saying) and *appreciate* (where a speaker responds appreciatively to a previous turn in which the addressee has indicated something from which speaker is presumed to benefit, e.g. *'That's great'*). There is one communicative function to address negative feedback (negative perception or interpretation), for utterances such as *pardon* which is a general request for repetition, expressing that the speaker was unable to hear or understand what was said.

**DIT**++, LIRICS and some other schemes make a distinction between auto-feedback, which is about the speaker's processing of the previous discourse, and *allo-feedback*, which is about the addressee's

| | | | | | | |
|---|---|---|---|---|---|---|
| **DIT** | Turn-take | Turn-grab | Turn-accept | Turn-keep | Turn-assign | Turn-release |
| **LIRICS** | Turn-take | Turn-grab | Turn-accept | Turn-keep | Turn-assign | Turn-release |
| **DAMSL** | | | | Turn maintain | | |
| **SWBD-DAMSL** | | | Hold before answers | Turn maintain | Turn exit | |
| **MRDA** | Regain turn | grabber | Hold before answers | holder | | |
| **Coconut** | | | | Turn maintain | | |
| **SLSA** | Turn taking | interruption | Turn opening | Turn-holding | Turn closing | |
| **TRAINS** | Turn take | | | Turn-keep | Turn assign | Turn release |
| **SPAAC** | | | | hold | | |
| **MALTUS** | | Turn grabber | | Turn holder | | backchannel |
| **Primula** | | Turn grabber | | Turn holder | | backchannel |
| **Chiba** | | | | hold | | |

Table 2: ***Turn Management*** *communicative functions in different dialogue act taxonomies.*

processing (see the above distinction between giving and eliciting feedback made by **SLSA**). In [34] it was noticed that addressees in dialogue cooperate by displaying and signalling their understanding, and that the speaker also monitors addressees for their understanding, and when necessary alter their utterances or elicit feedback. **SWBD-DAMSL** and **MRDA** define *backchannels in question form* for utterances like *'right?'*. Additionally, **MRDA** has *'follow-me'* questions where the speaker wants to verify that what he is saying is being understood, e.g. *'Do you know what I mean?'* **Coconut** introduces a *correct assumption* function which is used to correct both speaker's and addressee's wrong assumptions at the semantic level, while in **DAMSL** *correct misspeaking* was used for correction at the level of speakings. The **AMI** scheme has several functions defined to signal feedback elicitation: *elicit inform*, which is used by a speaker to request that someone else give some information which maybe about the task but also about feedback (unspecified here); *elicit assessment*, where the speaker attempts to elicit an assessment about what has been said or done so far; and *elicit comment-about-understanding*, where the speaker attempts to elicit a comment about whether or not what has been said or done so far has been understood. The **TRAINS** scheme has *request acknowledgment* and *request repair* to code feedback elicitation, and the **Verbmobil** scheme has *request comment*. Thus, feedback elicitation is an important communicative aspect; this is reflected both in theoretical studies and in the majority of dialogue act annotation taxonomies (just 6 of the 18 analysed schemes do not have feedback eliciting functions).

## 5.3 Taking Turns

The majority of DA schemes define communicative functions dealing with *turn management* (see Table 2 for an overview). **DIT++** and **LIRICS** define 6 communicative functions in this dimension: *turn accepting, grabbing* and *taking* as turn-initial functions, and *turn keeping, assigning* and *releasing* as turn-final (or closing) functions.

All multidimensional annotation schemes, like those based on **DAMSL**, define turn management functions. **SWBD-DAMSL** and **MRDA** have *hold* before answers, which corresponds with **DIT** *turn accept* and indicates that the speaker has some reasons or evidence to believe that she was selected for the next turn by the previous utterance and performs some actions to signal acceptance of the turn. Speakers to whom the next turn is assigned may simply start speaking without performing any extra actions. Sometimes, however, speakers do indicate explicitly that they agree to take the turn. We detected 33 functional segments in our AMI data having the communicative function of *turn accepting*, about 0.8% of the data. This means that every fifth *turn assignment* was followed by explicitly expressed *turn accepting*. The **SLSA** scheme defines the *turn opening* function to code explicit turn acceptance.

Like DIT, **MRDA** distinguishes a *turn grabbing* function for utterances which are used by the speaker to interrupt the partner who has the turn. Interruptions are important elements in conversation; they play a key role in signalling and resolving imbalances in information adequacy and desired topic direction, and they may be competitive, cooperative, clarification requests and unintentional interruptions ([82]). The interruptive behaviour of dialogue participants has been studied both from interpersonal and intercultural perspectives. For example, the turn-taking process was seen as a way of exercising

influence in groups. Subjects scoring high on dominance hold the floor longer, and attempt more interruptions ([73]). In the AMI data 171 segments were detected having the communicative function of *turn grabbing*, which accounts for 4.4% of all functional segments in corpus. About 89% of the interruptions were completed successfully, leading to speaker switch.

The **SLSA** and **TRAINS** annotation schemes have a *turn taking* function. According to [80] any instances of starting to talk (also interrupting the current speaker) can be seen as a take-turn attempt. According to **DIT**, turn taking events occur when the speaker wants to have the turn which is available. These events take place after the previous speaker released the turn so that anybody may continue the conversation (Sacks's rule Nr 2). In the AMI data 477 functional segments were identified that have an explicitly signalled *turn taking* function; this accounts for 12% of all functional segments in corpus.

Segments where the speaker indicates that she wants to have the next turn are in general quite well detectable and successfully automatically classified with an accuracy of 97% (using the RIPPER rule inducer). These scores outperform the baseline of 41%, which in this case was the percentage of the first tokens in a segment that do not have a turn-initial function. It was noticed in [41] that while turn-initial utterances share a very similar vocabulary (e.g. 'well' can be used to grab, take or accept the turn), they are very different in sound. Presences of pauses before and after a segment, durational, and acoustic properties help facilitate the detection of turn-initial segments.

As for turn-final functions, almost every analysed taxonomy defines a function for *turn-keeping* (in **TRAINS**: *turn keep*; **DAMSL, SWBD-DAMSL, Coconut**: *turn maintain*; **MRDA, SLSA, SPAAC** and **Chiba**: *turn (floor) hold*). Sometimes the speaker may want to continue with the next or part of the old contribution and signals that he wishes to stay in the sender role. In this case, no reallocation of the speaker role occurs. The efforts that the speaker makes in order to achieve this constitute a *turn keeping* act. Functional segments with the communicative function of *turn keeping* frequently occur in our data (28.2%).

Like DIT, **TRAINS** [80] distinguishes between *turn-release* and *turn-assign* utterances. **SLSA** has a *turn closing* function covering these two types of utterances, which signal explicit turn allocation. According to Sacks's first rule, after finishing his dialogue contribution the speaker may select the next speaker for the next turn. The act of indicating to the addressee that he may take the turn, constitutes a *turn assigning* act. About 4.6% of all functional segments in the AMI data have the communicative function *turn assign*. When the speakers offer the speaker role without selecting the next speaker and without putting any pressure on the addressee to take the turn, this behaviour constitutes a *turn releasing* act. To release the turn the speaker may just stop speaking. Ceasing to speak could by default be annotated as an indication of the *turn release* function. We studied, however, explicit *turn release* acts. About 1.3% of all functional segments in the AMI data have the explicitly signalled communicative function of *turn release*. Turn releasing utterances can be signalled by the following expressions:

- *anybody*, *anything* or *any* for example: '**Anybody** anything to add?'; '**Anything** else to say at all?'; '**Any** thoughts on that at all'
- *everybody*, for example: 'Is that what **everybody** got?'
- *we* or *all* for example:'Shall **we** make the decision?'; '**All** ready to go?
- *you* in general meaning, for example:

  (4) *B1: First of all just to kind of make sure that we all know each other*
  *B2: I'm Laura and I'm the project manager*
  *B3: Do **you** want to introduce yourself again?*

### 5.4 Social obligations and politeness

Except for the **Chiba** [53] and **HRCR Maptask** [28] dialogue annotation schemes, all other taxonomies address the dimension of *social obligations* and *politeness*, albeit to a different extent (see Table 3 for an overview). Some schemes have two functions defined for *greeting* and *good-bye*, such as **DAMSL, Coconut, LinLin** and **Alparon** [83], or only *greeting* as **SLSA**. Some others have additional communicative functions to address this aspect of communication, such as *self-introduction* (**DIT**, **LIRICS**, **Verbmobil**, **SPAAC** and **C-Star**), *thanking* (**DIT**, **LIRICS**, **SWBD-DAMSL**, **MRDA**, **Verbmobil**, **C-Star** and **SPAAC**), *apology* (**DIT**, **LIRICS**, **SWBD-DAMSL**, **TRAINS**, **SPAAC** (where it is called

| | Social Obligation Management | | | | | Discourse Structuring | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| **DIT** | Greeting/ return greeting | Self-introduction/ return self-introduction | Goodbye/ return goodbye | Apology/ accept apology | Thanking/ accept thanking | opening | Pre-closing | Topic introduction | Topic shift | Topic shift announcement |
| **LIRICS** | Greeting/ return greeting | Self-introduction/ return self-introduction | Goodbye/ return goodbye | Apology/ accept apology | Thanking/ accept thanking | Interaction structuring | | | | |
| **DAMSL** | greeting | | Good-bye | | | opening | closing | | | |
| **SWBD-DAMSL** | greeting | | | Apology/down-player | Thanking/ downplayer | opening | closing | | | |
| **MRDA** | | | | Downplayer/ sympathy | thanking | | | Topic change | | |
| **Coconut** | greeting | | Good-bye | | | opening | closing | topic | | |
| **AMI** | Be-positive/ be-negative | | | | | Argument structure; topic segmentation schemes | | | | |
| **HCRC MapTask** | | | | | | | | Ready (for topic shifts) | | |
| **Verbmobil** | greet | introduce | bye | Polite (apologies and compliments) | thank | | Task close | Task initiate | digress | |
| **Linlin** | | | | | | opening | ending | Topic layer | | |
| **SLSA** | greet | | | | | opening | closing | opening | continuation | |
| **TRAINS** | greet | | | apologize | | | | | | |
| **SPAAC** | greet | Identify-self | bye | Express regret | thank | | | Initialise; Raise issue | topic | |
| **MALTUS** | politeness | | | | | | | | Topic change | |
| **Primula** | Politeness; face-threatening/ face-saving | | | | | opening | closing | Topic opening | Topic closing/ topic change | |
| **Chiba** | | | | | | opening | closing | Topic break | | |
| **Alparon** | greet | | bye | | | | | | | |
| **C-Star** | greeting | Sef-introduction | | apologize | thanking | | closing | Introduce topic | | |

Table 3: ***Social Obligation Management*** *and* ***Discourse Structuring*** *communicative functions in different dialogue act taxonomies.*

*express regret*) and **C-Star**), and reaction to the latter two like *downplayer* (**DIT**, **LIRICS**, **SWBD-DAMSL** and **MRDA** (which also has the *sympathy* function)). **AMI** and **Verbmobil** have some unspecified social obligation functions. For example, *be-positive* in **AMI** includes any social acts that are intended to make an individual or group happier, including acts of politeness like greeting one another or saying ”please”, ”sorry”, and ”thank you” for smooth social functioning in the group, as do things like good-natured jokes, positive comments about someone’s appearance or intelligence, and expressions that say they are doing a good job. *Be-negative* in AMI includes any social acts that express negative feelings towards an individual or group, e.g. hostile comments, jokes if the point is to run down someone, and expressions of frustration or withdrawal. *Politeness_formula* in **Verbmobil** is for asking about the partner’s good health or formulating compliments.

## 5.5 Discourse and topic structure

Except for **AMI**, **TRAINS** and **Alparon** all other taxonomies define communicative functions for *Discourse Structuring*. It should be noted, however, that within **AMI** separate taxonomies hava been designed to analyse topical and argumentative structures in meetings (see [84] and [72]).

As for individual communicative functions, *opening* and *closing* are the most frequently defined ones (**DIT**, **DAMSL**, **Coconut**, **Linlin**, **SLSA**, **Chiba** and **C-Star**). There are some variations in terminology and in the level of granularity. Some schemes leave *topic* functions unspecified (e.g. **Linlin** and **SPAAC**). Some other taxonomies have more specific functions such as *topic change/shift* (**DIT** and **MRDA**), or *ready* (**HCRC Maptask**), *topic introduction/opening* (**SLSA**, **C-Star**) and *task introduction* and *digress* (**Verbmobil**). Still others are very domain specific, for example, **Coconut** has labels like topic proper (furniture items) with *needItem*, *haveItem*, *getItem*, etc. The **SPAAC** scheme has *init(ialize)* as a dialogue control act for initiating a new phase of the dialogue.

## 5.6 Monitoring one’s own and the addressee’s speech

For monitoring and editing one’s own speech (*own communication management*), the majority of annotation schemes address this aspect of communication (10 from the 18; see Table 4). Bales in [15] notices

| | Own Communication Management | | | Partner Communication Management | | Time Management | | Contact Management | |
|---|---|---|---|---|---|---|---|---|---|
| **DIT** | Error-signalling | retraction | Self-correction | Correct-misspeaking | completion | stalling | pausing | Contact check | Contact indication |
| **LIRICS** | Error-signalling | Self-correction | | Correct-misspeaking | completion | stalling | pausing | Contact check | Contact indication |
| **DAMSL** | | Speech repair | | Correct-misspeaking | completion | Communication management: delay | | Communication channel | |
| **SWBD-DAMSL** | | Speech repair | | Correct-misspeaking | completion | Stalling; delay; Hold before answers | | Communication channel | |
| **MRDA** | | Self-correction | | Correct-misspeaking | Collaborative completion | Hold before answers | | | |
| **Coconut** | | Correct-assumption; speech repair | | Correct-misspeaking | completion | delay | | Communication channel | |
| **AMI** | | | | | | stall | | | |
| **Verbmobil** | | | | | | deliberate | | Refer-to-settings | |
| **SLSA** | change | | | | | choice | | | |
| **TRAINS** | | repair | | | | keep | | | |
| **SPAAC** | | Correct-self | | correct | complete | hold | | | |
| **MALTUS** | | Restated info with repetition; restated info with correction | | restated info with correction | | | | | |
| **Alparon** | | | | | | | pause | | |
| **C-star** | | | | | | | Please-wait | | |

Table 4: ***Own, Partner Communication Management***, ***Time*** *and* ***Contact Management*** *communicative functions in different dialogue act taxonomies.*

that it is important for cooperative communicative partners to signal and admit an error or oversight in dialogue. **DAMSL** and **Coconut** mention this phenomenon in their Communication Management dimensions without defining individual communicative functions. DAMSL-based schemes have a dialogue act tag for *speech repair* (**SWBD-DAMSL**) or *self-correct misspeaking* (**MRDA**: marks when a speaker corrects his own errors with regard to either pronunciation or word choice). **Coconut** has additionally the *correct assumption* function for both partner- and self-corrections at the semantic level. The **TRAINS** scheme has the *repair* function defined for utterances which replace any of the content of the current dialogue unit [80]. It is also noticed that these changes could be made in order to make the content of an utterances or a presupposition explicit. They are often prefaced by editing phrases like *'I mean'* or apologies. The **SPAAC** scheme has the communicative function *correct-self* for speaker's own utterances.

*Partner communication management* is concerned with monitoring the partner's speech by the speaker, providing assistance by completing an utterance that the partner is struggling to complete (*completion*) or correcting (part of) partner's utterance, believing that the partner made a speaking error (*correct-misspeaking*). **DAMSL** and DAMSL-based schemes define these functions within the dimension of Understanding (Feedback). **SPAAC** also defines the function *correct* (correction of what the partner just said including misspeaking and utterance content) and *complete* (completing the partner's move). **MALTUS** [70] defines the *restated info with correction* function, leaving unspecified whether speaker or partner was corrected.

## 5.7 Time

The majority of the analysed schemes (12 of 18) define dialogue acts that address the *management of time* in dialogue. *Stalling* is the function of utterances where the speaker indicates that he needs a little bit of time to formulate an utterance. This function is defined in **DAMSL** and **Coconut** in the Communicative Management dimension (called *turn delays*). In **SWBD-DAMSL** *stallings for time*, *delays* and *holds before answering* address this aspect of communication. **Verbmobil** calls the utterances, used to gain time by thinking aloud or using certain formulas, *deliberate*. **AMI** defines *stallings* as special cases; it is argued that these utterances are not really a dialogue act, since the speaker doesn't convey an intention in these segments. **SLSA** has *choice* as a mechanism enabling the speaker to gain time for processes having to do with the continuation of the interaction (involving hesitation, memory search, planning, and keeping the floor), but these are thought to address the OCM dimension. The **Alparon** scheme has the dialogue act *pause* defined, in **C-Star** called *please-wait*. In **TRAINS** this function

is covered by the *turn-maintaining* tag, e.g. for 'filling' pauses like 'uhh' where the speaker wants more time to work out his intended utterance. Finally, **SPAAC** defines *hold* as a dialogue act where the speaker indicates that he needs time and asks the partner to hold the line. Thus, two tendencies are observed here: (1) defined but considered as special cases, not as intentional acts; and (2) defined to address other dimensions: Turn Management or OCM.

### 5.8 Contact and attention

6 of the 18 studied dialogue act schemes define tags addressing the monitoring of contact and attention. **DAMSL**, **SWBD-DAMSL** and **Coconut** have *communication channel* estaishment in the Communication Management dimension, for utterances like *'Are you there?'* (*contact check* in **DIT** and **LIRICS**) and the answer 'I'm here' (*contact indication* in **DIT** and **LIRICS**). **Verbmobil** defines *refer-to-settings* tag which addresses the settings of interaction, e.g. noise in the room, or the output quality of the computer used in the interaction. **HRCR Maptask** has *align* for checks of the attention or agreement of the partner, or his/her readiness for the next move (the second part of the definition is particularly relevant here).

### 5.9 Summary

To summarize, the following aspects of communication are reflected in the majority of dialogue act taxonomies:

- Task (17 of 18; not defined in SLSA);
- Auto-Feedback (16 of 18; not defined in Linlin and Primula);
- Allo-Feedback (elicitation) (12 of 18; not defined in DAMSL, LinLin, SPAAC, Primula, Chiba and Alparon);
- Turn management (12 of 18; not defined in HCRC MapTask, AMI, Verbmobil, Linlin, Alparon and C-Star);
- Discourse Structuring (16 of 18; not defined in TRAINS and Alparon);
- Social Obligation Management (16 of 18; not defined in Chiba and HCRC MapTask);
- Own Communication Management (10 of 18; not defined in AMI, HCRC MapTask, Verbmobil, Linlin, Primula, Chiba, Alparon and C-Star);
- Time Management (12 of 18; not defined in MRDA, HCRC MapTask, Linlin, Maltus, Primula and Chiba);

In addition, Contact Management is addressed by all multidimensional dialogue taxonomies, by Verbmobil and by HCRC MapTask. Partner Communication Management is reflected in the multidimensional dialogue taxonomies only.

## 6 Empirical observations from dialogue corpora

The majority of utterances in most dialogues involve performing the *task* or *activity* that motivates the dialogue, as Table 5 shows. The second largest category of utterances in AMI and DIAMOND data addresses *auto-feedback*, showing its importance for communication. In fact we observed that in AMI meetings one minute of conversation contains on average 9.4 positive auto-feedback utterances; even more auto-feedback utterances (13.4) were observed in the middle and near the end of a dialogue. In OVIS dialogues a significantly larger portion of *allo-feedback* was observed. This is not surprising since these are human-machine dialogues and the system's processing of user's utterances often fails due to faulty input from the ASR module. The OVIS system constantly checks its correct understanding of user utterances, and the user reports back on the correctness of the system's understanding, addressing the dimension of allo-feedback. A considerable amount of turn and time management utterances was observed in AMI and DIAMOND dialogue corpus data. Being multiparty interactions, AMI-meetings clearly involve more complex turn management mechanisms where participants perform certain actions to take the turn rather than just start speaking (more than the half of all segments was preceded by certain turn-obtaining events (59%)); they interrupt each other (4.4%) and speak simultaneously (20%

| | AMI | DIAMOND | OVIS |
|---|---|---|---|
| Task | 33 | 47.7 | 48.8 |
| Auto-Feedback | 20 | 14 | 18 |
| Allo-Feedback | 0.7 | 3.8 | 39 |
| Turn Management | 15 | 14.2 | 1 |
| Social Obligation Management | 0.3 | 5 | 3.8 |
| Discourse Structuring | 2.2 | 2.3 | 2.4 |
| Own Communication Management | 8.7 | 0.7 | 0.3 |
| Time Management | 16.8 | 10.7 | 0.6 |
| Partner Communication Management | 0.3 | 0.3 | 0.1 |
| Contact Management | 0.1 | 1.3 | 12.3 |

Table 5: *Distribution of utterances across dimensions for analysed dialogue corpora in (%).*

of all segments partly overlap). The OVIS dialogue system exhibits behaviour that is not natural for humans. Features that are characteristic for human dialogue behaviour such as hesitations, time delays, self-corrections, misspeaking, etc. were observed for the human user but not for the computer system.

Another noticeable difference between different types of dialogues is *contact* management. Since AMI participants have face-to-face contact there are not so many utterances dealing with this aspect of communication, and contact is managed by using non-verbal means most of the time, e.g. by securing eye-contact, by posture shifts forward or to the speaker, or by short head nods indicating active listening. Since these are phone conversations, the participants in OVIS dialogues are less certain about the partner's presence and readiness to start or continue the interaction; this explains a significantly larger amount of utterances used for this purpose.

*Social obligation* acts are used more frequently in DIAMOND and OVIS dialogues. In OVIS dialogues the main producer of socialy polite utterances is the system. It always greets the user in the beginning of the dialogue and introduces itself; the user, by contrast, usually does not return the greeting. The system is designed to apologize if its processing of the user's utterances fails. DIAMOND participants also act in accordance with social norms and obligations by greeting, apologising and thanking each other. Social obligation acts were observed in the AMI corpus especially during the introduction phase of the first meeting, when participants need to get to know each other. When closing a meeting, the participants always express gratitude to each other for successful cooperation.

Thus, all dimensions mentioned in Section 4 are observed in dialogue corpus data, though not in equal proportions. The distribution of the data across dimensions is one of the main distinguishing features of different types of dialogue, such as multi- vs. two-party interactions, face-to-face vs remote conversations, human-human vs human-machine, and formal vs informal, instructive vs information seeking vs meeting dialogues.

# 7 Dimension recognition

How important is (human and machine) recognition of dimensions, and inter-annotator agreement on the assignment of dimensions to a markable? Dimension recognition is not important in relation to the use of dimension-specific communicative functions, e.g. Turn Take or Grab, or Greeting, Topic introduction, because these functions may occur only in one particular dimension (are specific to it) and specifying the dimension is redundant, for example:

(5) Auto-feedback: Overall Positive *Okay*
Allo-feedback: Evaluation Elicitation *Okay?*
Turn management: Turn Assign *Craig?*
Time management: Stalling *Well, you know,..*
Contact management: Contact Checking *Hello?*
Own communication management: Self-correction *I mean...*
Partner communication management: Completion ... *completion*
Dialogue structuring: Topic Shift Announcement *Something else*
Social obligation management: Valediction *Bye*
Task/domain: Open Meeting *I open this meeting*

Dimension recognition is, by contrast, essential in connection with the use of general-purpose functions. For instance, an Auto-Feedback Inform as expressed by *'I didn't hear what you said'* is semantically equivalent to the use of the feedback-specific function Perception-Negative (in the Auto-Feedback dimension) as may be expressed by *'I beg you pardon?'* or *'What?'* accompanied with a hand gesture behind an ear. This semantic equivalence would not be brought out at all if the utterance *'I didn't hear what you said'* was annotated just as Inform (rather than Auto-Feedback Inform). More generally, the intended update effect associated with the use of a general-purpose function crucially depends on the dimension, or kind of semantic content, that the function is combined with to form a full-blown dialogue act. There are other examples of Informs in various dimensions:

(6) *The KL204 leaves at 12.30* (Task/domain)
*I see what you mean* (Auto-feedback)
*You misunderstood me* (Allo-feedback)
*I would like to hear Peters opinion* (Turn mananagement)
*Im listening* (Contact management)
*... I mean Toronto* (Own communication management)
*We should also discuss the agenda* (Discourse structuring)
*Im very grateful for you help* (Social obligation management)

Table 6 shows the agreement observed between two expert annotators tagging the DIAMOND and OVIS data.

| DIMENSION | OBSERVED AGREEMENT |
|---|---|
| Task | 84.99 |
| Auto-feedback | 91.32 |
| Allo-feedback | 93.31 |
| Time-Management | 98.55 |
| Turn-Management | 92.59 |
| Contact-Management | 99.28 |
| Own-Communication-Management | 99.10 |
| Partner-Communication-Management | 99.46 |
| Dialogue-structuring | 98.73 |
| Social-Obligations-Management | 99.10 |
| Weighted AVG | 96.04 |

Table 6: *Observed agreement between two expert annotators on the DIAMOND and OVIS data.*

Inter-annotator agreement is commonly calculated for the qualitative evaluation of a tagset using Cohne's kappa statistic [29], [36]. When the inter-annotator agreement scores for data annotated with a particular tagset indicate high reliability of the annotations[7], this does not not guarantee high agreement on the assignment of the right concept. Even though it is not likely to happen often, annotators occasionally show perfect agreement in assigning a specific concept, but disagree with an expert on what would be the correct concept to be assigned. In other words, to obtain reliable annotations inter-annotator agreement scores should be complemented with annotation accuracy. This is done by comparing the data produced by annotators with a gold standard [44]. Table 7 presents both inter-annotator agreement for expert annotators expressed in terms of kappa and tagging accuracy. The table shows that there are no systematic differences between annotators in assigning values for dimensional tag.

While human annotators are quite successful in dimension recognition, the question arises whether comparable scores can be obtained in macine recognition. A wide variety of machine-learning techniques has been used for NLP tasks with various instantiations of feature-sets and target class encodings; for dialogue processing, it is still an open issue which techniques are the most suitable for which task. We used the rule induction algorithm *Ripper* [37]. The advantage of such an algorithm is that the regularities discovered in the data are represented as human-readable rules. It is also shown in [46] that

[7] In case of Cohen's kappa, this is often taken to be between 0.8 and 1.0.

| Dimensions | Inter-annotator agreement | | | Tagging accuracy | | |
|---|---|---|---|---|---|---|
| | $P_o$ | $P_e$ | k | $P_o$ | $P_e$ | k |
| Task | 0.85 | 0.16 | 0.82 | 0.91 | 0.16 | 0.9 |
| Auto-Feedback | 0.92 | 0.57 | 0.82 | 0.94 | 0.48 | 0.88 |
| Allo-Feedback | 0.95 | 0.24 | 0.89 | 0.95 | 0.22 | 0.94 |
| Turn Management | 0.94 | 0.58 | 0.76 | 0.92 | 0.67 | 0.94 |
| Time Management | 0.99 | 0.87 | 0.92 | 0.99 | 0.88 | 0.94 |
| Discourse Structuring | 0.92 | 0.38 | 0.88 | 0.87 | 0.34 | 0.81 |
| Contact Management | 0.95 | 0.48 | 0.81 | 0.91 | 0.48 | 0.83 |
| Own Communication Management | 0.99 | 0.38 | 0.98 | 1 | 0.38 | 1 |
| Partner Communication Management | 0.99 | 0.38 | 0.98 | 1 | 0.38 | 1 |
| Social Obligations Management | 0.99 | 0.24 | 0.99 | 0.95 | 0.23 | 0.94 |

Table 7: *Inter-annotator agreement and tagging accuracy per dimension.*

| Dimension | DIAMOND | | AMI | | OVIS | |
|---|---|---|---|---|---|---|
| | BL | Accuracy (%) | BL | Accuracy (%) | BL | Accuracy (%) |
| Task | 64.9 | 70.5 | 66.8 | 72.3 | 60.8 | 73.5 |
| Auto-Feedback | 71.1 | 85.1 | 77.9 | 89.7 | 66.1 | 75.9 |
| Allo-Feedback | 86.9 | 96.6 | 96.7 | 99.3 | 52.5 | 80.1 |
| Turn Management | 69.5 | 90.0 | 59 | 93 | 89.8 | 99.2 |
| Time Management | 65.6 | 82.2 | 69.7 | 99.4 | 95.9 | 99.4 |
| Discourse Structuring | 59.0 | 67.9 | 98 | 92.5 | 76.3 | 89.4 |
| Contact Management | 88.0 | 95.2 | 99.8 | 99.8 | 87.7 | 98.5 |
| Own Communication Management | 77.4 | 83.1 | 89.6 | 94.1 | 99.7 | 99.7 |
| Partner Communication Management | 45.4 | 62.6 | 99.7 | 99.7 | 99.8 | 99.8 |
| Social Obligation Management | 80.3 | 92.2 | 99.6 | 99.6 | 96.2 | 98.4 |

Table 8: *Success scores in terms of accuracy (in %) comparing to baseline scores (BL) for each dimension and data set.*

Ripper performed best on our data comparing to statistical learners (e.g. Naive-Bayes classifiers) and memory-based learners (e.g. IB1).

Every communicative function is required to have some reflection in observable features of communicative behaviour, i.e. for every communicative function there are devices which a speaker can use in order to allow its successful recognition by the addressee, such as linguistic cues, intonation properties, properties of dialogue history, etc. State-of-the-art automatic dialogue understanding uses all available sources to interpret a spoken utterance. Features and their selection play a very important role in supporting accurate recognition and classification of utterances and their computational modelling may be expected to contribute to improved automatic dialogue processing. The features included in the data sets considered here are those relating to *dialogue history*, *prosody*, and *word occurrence*.

For dialogue history we used of the tags of the 10 (AMI and OVIS) or 4 (DIAMOND) previous turns. Additionally, the tags of utterances to which the utterance in focus was a response, as well as timing, are included as features. For the data that is segmented per dimension, some segments are located inside other segments. This occurs for instance with backchannels and interruptions, that do not cause turn shifting; the occurrence of such events is encoded as a feature.

Prosodic features that are included are minimum, maximum, mean, and standard deviation of *pitch* (F0 in Hz), *energy* (RMS), *voicing* (fraction of locally unvoiced frames and number of voice breaks), and *duration*. Word occurrence is represented by a bag-of-words vector[8] indicating the presence or absence of words in the segment. In total, 1,668 features are used for AMI data, 947 for DIAMOND data and 240 for OVIS data. For the AMI data we additionally indicated the speaker (A, B, C, D) and the addressee (other participants individually or the group as a whole).

[8]With a size of 1,640 entries for AMI data, 923 for DIAMOND data and 219 for OVIS data.

Table 8 presents the resulting scores using the Ripper classifier obtained in 10-fold cross-validation experiments[9].

As our results show, the 10 dimensions defined in DIT++ and LIRICS are recognizable as well by human annotators and by machine. As for the *Task Management* dimension defined in DAMSL, we noticed earlier in this report the observed agreement was 72%, the tagging accuracy, however, was only 42%. This dimension was often confused with Communication Management or Task.

# 8 The independence of dimensions

The distinction of a dimension only makes sense if it can be separated from the other dimension that are considered. Therefore, in [23] it was proposed as part of the definition of 'dimension' that it corresponds to an aspect of communication that an utterance may address independently of other aspects that it might also address. This means that an utterance may in principle be assigned any tag in a given dimension, regardless of whatever tags have been assigned to it it in other dimensions. This is only *in principle*, though; empirically, there are restrictions of assigning tags multiple dimensions. For example, accepting an offer cannot have a negative feedback function, because an answer presupposes that the speaker believes to have understood the preceding question; similarly, a farewell greeting closing a dialogue can not have a feedback elicitation function or a turn-assigning function. So the assignment of a communicative functions in a certain dimension may entail restrictions on the possible tagging in another dimension. Such occasional restrictions on the co-assignment of tags in different dimensions correspond to empirical facts about communication, and do not affect the independence of the dimensions. Two dimensions are not independent if there are systematic relations between the tags in one dimension and those in the other, in particular if the tag in one dimension can be computed from that in the other.

We define the independence (or 'orthogonality') of a set of dimensions as follows. First, we define the pairwise independence of two dimensions:

(7) **Definition.** Two dimensions $D_1$ and $D_2$ are called *pairwise independent* iff:

1. a markable may be assigned a $D_2$ tag, regardless of whether a $D_1$ tag is assigned (and vice versa);
2. if a markable is assigned both a $D_1$ tag and a $D_2$ tag, then the $D_2$ tag is in general not determined by the $D_1$ tag (and vice versa).

(8) **Definition.** A set $D$ of dimensions is independent iff every pair $< D_i, D_j > \in D$ is pairwise independent.

The independence of a set of dimensions can be determined empirically and theoretically. Theoretically, dependency relations can be uncovered by analyzing the definitions of dimensions and their function tags, in particular for the existence of logical relations between the preconditions of communicative functions. For example, a *dialogue opening* is logically related to a *contact indication* act, because the precondition for a *contact indication* act, which says that the speaker wants the addressee to know that the speaker is ready to communicate with the addressee, is among the preconditions of a *dialogue opening*.

Empirically, dependency relations can be found by analyzing annotated dialogue data. Tags which always co-occur are either logically related or else show an empirical fact about communication; similarly for zero co-occurrence scores. Besides co-occurrence scores, we also provide a statistical analysis using the phi coefficient as a measure of relatedness. The phi measure is related to the chi-square statistic, used to test the independence of categorical variables, and is similar to the correlation coefficient in

[9] In order to reduce the effect of imbalances in the data, it is partitioned ten times. Each time a different 10% of the data is used as test set and the remaining 90% as training set. The procedure is repeated ten times so that in the end, every instance has been used exactly once for testing and the scores are averaged. The cross-validation was stratified, i.e. the 10 folds contained approximately the same proportions of instances with relevant tags as in the entire dataset.

| within | Task | Auto-F. | Allo-F. | Turn M. | Time M. | DS | Contact M. | OCM | PCM | SOM |
|---|---|---|---|---|---|---|---|---|---|---|
| Task | - | 0.05(67.9) | 0(24.9) | 10.2(97.5) | 1.4(2.4) | 1.4(1.5) | 0 (0.4) | 5.1(69.6) | 0(0.1) | 0(0.7) |
| Auto-F. | 0.7(78.9) | - | 0(0) | 9.1(98.7) | 0.6(1.4) | 0.3(1.2) | 0(20.2) | 0(0.7) | 0(65.0) | 0(0.7) |
| Allo-F. | 0(24.9) | 0 | - | 59.2(94.8) | 1.2(35.7) | 0(2.1) | 0(1.2) | 0(7.9) | 0.6(0.7) | 0(0.3) |
| Turn M. | 50.2(76.0) | 3.5(66.2) | 5.6(19.4) | - | 8.0(42.9) | 1.2(3.9) | 0.1(13.8) | 25(99.6) | 0.2(1.0) | 0.2(0.5) |
| Time M. | 28.2(13.4) | 0.5(11.3) | 2.8(7.8) | 96.9(98.6) | - | 0.7(1.7) | 0(0) | 2.5(83.2) | 0(0.5) | 0(0) |
| DS | 28.3(92.2) | 0.4(58.3) | 0(29.1) | 22.6(87.5) | 4.2(4.9) | - | 0(25.0) | 0(3.7) | 0(0) | 3.2(12.5) |
| Contact M. | 0(2.4) | 0(97.1) | 0(1.6) | 18.2(98.8) | 0(0) | 0(2.4) | - | 0 (0.3) | 0(0) | 0(0) |
| OCM | 75.5(82.2) | 0(0.8) | 0(2.5) | 82.9(96.9) | 3.4(7.8) | 1.3(3.9) | 0(13.5) | - | 0 (0.9) | 0.2(0.6) |
| PCM | 0(11.8) | 0(65.0) | 4.9(11.8) | 12.2(79.1) | 0(12.2) | 0(0) | 0(0) | 0(0) | - | 0(0) |
| SOM | 0(0.7) | 0(80.0) | 0(10.0) | 6(90.0) | 0(0) | 10.0(30.0) | 0(0) | 0(2.0) | 0(0) | - |

Table 9: *Co-occurrences of communicative functions across dimensions in the AMI corpus, expressed in relative frequency in %, implicated and entailed functions excluded and included (in brackets).*

its interpretation. In addition, to investigate whether dimensions are concerned with very different information, we defined the similarities between dimensions in terms of distances between dimension vectors in a multidimensional space, where orthogonal vectors convey unique, non-overlapping information.

If a dimension is not independent from other dimensions, then there would be no utterances in the data which address only that dimension. Looking for utterances which address *only* one dimension is therefore another test. Finally, we also investigate whether a dimension is addressed always in reaction to a certain other dimension. If that is the case, then the presence of a dimension in a multidimensional scheme depends on the presence of another dimension. For example, the *answer* dimension as defined in DAMSL cannot be seen as an independent dimension because *answers* need *questions* in order to exist. The test here is to examine for each dimension the relative frequencies of pairs <dimension tag, previous dimension tag>; if a tag always co-occurs with a certain previous tag, then there is apparently a dependence between the two.

To sum up, we perform 5 tests, examining:

1. the relative frequency of *communicative function co-occurrences* across dimensions;
2. *the extent of relatedness between dimensions* measured with the phi coefficient;
3. *dimension vector distances* in multidimensional space;
4. for all dimensions whether there are utterances *addressing only that dimension*;
5. the relative frequency of pairs of *dimension* and *previous dimension*.

All three corpora were manually segmented and tagged using the DIT++ annotation scheme. The test results presented in this section are similar for all three corpora.

The co-occurrence results in Table 9 show no dependences between dimensions, although some combinations of dimensions are relatively frequent, e.g. time and turn management acts often co-occur. A speaker who wants to win some time to gather his thoughts and uses Stalling acts mostly wants to continue in the sender role, and his stalling behaviour may be intended to signal that as well (i.e., to be interpreted as a Tun Keeping act). But stalling behaviour does not always have that function; especially an extensive amount of stallings accompanied by relatively long pauses may be intended to elicit support for completing an utterance.

It is also interesting to have a look at co-occurences of communicative functions taking implicated and entailed functions into account (the corpora were reannotated for this purpose). An *implicated* function is for instance the positive feedback (on understanding and evaluating the preceding utterance(s) of the addressee) that is implied by an expression of thanks; examples of *entailed* functions are the positive feedback on the preceding utterance that is implied by answering a question, by accepting an invitation, or by rejecting an offer.

Co-occurrence scores are higher when entailed and implicated functions are taken into account (the scores given in brackets in Table 9). For example, questions, which mostly belong to the Task dimension, much of the time have an accompanying Turn Management function, either releasing the

turn or assigning it to another dialogue participant, allowing the question to be answered. Similarly, when accepting a request the speaker needs to have the turn, so communicative functions like Accept Request will often be accompanied by functions like Turn Take or Turn Accept. Such cases contribute to the co-occurrence score between the Task and Turn Management dimensions. Nevertheless, again, no clear dependences between dimensions can be observed.

Table 9 shows that some dimensions do not occur in combination. We do not find combinations of Contact and Time Management, Contact and Partner Communication Management, or Partner Communication Management and Discourse Structuring, for example. Close inspection of the definitions of the tags in these pairs of dimensions does not reveal any clear restrictions on the possible co-assignment of tags in these dimensions, and hence no dependences between the dimensions.

Table 10 presents the extent to which dimensions are related when the corpus data are annotated without taking implicated and entailed functions are not taken (white cells) and when they are (grey cells), according to the calculated phi coefficient.

| Dimensions | Task | Auto-F. | Allo-F. | Turn M. | Time M. | Contact M. | DS | OCM | PCM | SOM |
|---|---|---|---|---|---|---|---|---|---|---|
| Task | | .1 | .3 | .06 | -.4 | -.6 | .03 | -.03 | -.1 | .04 |
| Auto-F. | -.5 | | -.6 | .1 | -.3 | .2 | -.02 | -.02 | -.1 | .04 |
| Allo-F. | -.2 | -.03 | | .09 | -.1 | -.2 | .03 | -.01 | -.02 | -.01 |
| Turn M. | -.03 | -.04 | .14 | | .6 | .04 | -.06 | .02 | .02 | -.03 |
| Time M. | -.4 | -.06 | .14 | .6 | | -.1 | -.02 | .04 | -.03 | -.02 |
| Contact M. | -.05 | -.006 | -.003 | .001 | -.007 | | .04 | -.01 | -.04 | -.03 |
| DS | -.2 | -.02 | -.01 | -.01 | -.02 | -.002 | | -.01 | -.01 | .2 |
| OCM | .01 | -.05 | .02 | .4 | -.03 | -.006 | -.003 | | -.03 | -.007 |
| PCM | -.1 | -.01 | .01 | -.006 | .01 | -.001 | -.005 | -.01 | | -.003 |
| SOM | -.1 | -.01 | -.007 | -.02 | -.02 | -.001 | .05 | -.007 | -.003 | |

Table 10: *Extent of relation between dimensions for AMI corpus expressed in the Phi coefficient (implicated and entailed functions excluded (white cells) and included (grey cells)).*

No strong positive (phi values from .7 to 1.0) or negative (-.7 to -1.0) relations are observed. There is a weak positive association (.6) between Turn and Time Management (see co-occurence analysis above) and between OCM and Turn Management (.4). Weak negative associations are observed between Task and Auto-feedback (-.5) when entailed and implicated functions are not considered; between Task and Contact Management (-.6); and between Auto- and Allo-feedback (-.6) when entailed and implicated functions are included in the analysis. The weak negative association means that an utterance does not often have communicative functions in these two dimensions simultaneously. Some negative associations become positive if we take entailed and implicated functions into account, because, as already noted, dialogue acts like answers, accepts and rejects, imply positive feedback.

For the third test we represented all annotated utterances by vectors with 8 prosodic values (duration, min, max, mean, sd in pitch, fraction voiced/unvoiced frames, voice breaks and intensity), 220 values for dialogue history and 1623 values for word tokens occurred in the utterance. To simplify the distance measures between dimensions we constructed for each dimension a dummy dimension at the centre of the dimension cloud, which is basically the centroid $C = (c_1,c_2,...,c_t)$, in which every $c_j$ is the mean of all the values of j:

$$c_i = \frac{1}{N}\sum_{j=1}^{N} w_{i,j}$$

where w is the weight value for each feature. We then measured the distances between dimension vectors pair-wise using Euclidean distance:

$$euclid(\vec{d}_j,\vec{d}_k) = \sqrt{\sum_{i=1}^{n}(w_{ij} - w_{ik})^2}$$

Table 11 presents the results of distance measures between centroid dimension vectors. There are no vectors which cross or overlap each other, although some dimension vectors are closer to each other in space, e.g. the Task dimension is closer to the Discourse Structuring dimension because they share more

| | task | auf | alf | turn | time | contact | ds | topic | ocm | pcm | som |
|---|---|---|---|---|---|---|---|---|---|---|---|
| task | .000 | 82.911 | 70.952 | 110.264 | 120.260 | 118.979 | 87.027 | 27.132 | 110.561 | 92.694 | 33.101 |
| auf | 82.911 | .000 | 33.906 | 39.855 | 53.530 | 132.211 | 136.307 | 76.489 | 42.106 | 31.326 | 72.339 |
| alf | 70.952 | 33.906 | .000 | 44.267 | 53.668 | 148.877 | 141.244 | 74.180 | 44.225 | 32.236 | 65.130 |
| turn | 110.264 | 39.855 | 44.267 | .000 | 13.833 | 171.665 | 174.537 | 110.668 | 3.597 | 19.186 | 104.994 |
| time | 120.260 | 53.530 | 53.668 | 13.833 | .000 | 185.428 | 187.414 | 122.505 | 11.951 | 28.982 | 116.671 |
| contact | 118.979 | 132.211 | 148.877 | 171.665 | 185.428 | .000 | 57.944 | 94.947 | 173.970 | 161.210 | 101.148 |
| ds | 87.027 | 136.307 | 141.244 | 174.537 | 187.414 | 57.944 | .000 | 69.747 | 176.066 | 159.736 | 80.440 |
| topic | 27.132 | 76.489 | 74.180 | 110.668 | 122.505 | 94.947 | 69.747 | .000 | 111.661 | 94.574 | 24.187 |
| ocm | 110.561 | 42.106 | 44.225 | 3.597 | 11.951 | 173.970 | 176.066 | 111.661 | .000 | 19.543 | 105.827 |
| pcm | 92.694 | 31.326 | 32.236 | 19.186 | 28.982 | 161.210 | 159.736 | 94.574 | 19.543 | .000 | 90.655 |
| som | 33.101 | 72.339 | 65.130 | 104.994 | 116.671 | 101.148 | 80.440 | 24.187 | 105.827 | 90.655 | .000 |

Table 11: *Distances between dimensions.*

| Dimension | Frequency (in %) | | |
|---|---|---|---|
| | AMI | OVIS | DIAMOND |
| task | 28.8 | 37.9 | 29.9 |
| Auto-feedback | 14.2 | 16.3 | 20.9 |
| Allo-feedback | 0.7 | 4.1 | 6.8 |
| turn | 7.4 | 0.9 | 8.5 |
| time | 0.3 | 0.4 | 0.7 |
| contact | 0.1 | 0.3 | 0.7 |
| Discourse structuring | 1.9 | 1.8 | 0.9 |
| Own Communication Management | 0.5 | 0.8 | 2.7 |
| Partner Communication Management | 0.2 | 3.1 | 0.4 |
| Social Obligation Management | 0.3 | 6.4 | 0.7 |

Table 12: *Overview of dimensions being addressed without any other dimension also being addressed in AMI, OVIS and DIAMOND data, expressed in relative frequency in%.*

or less the same vocabulary; Turn Management is close to Own Communication Management because they have similar prosodic properties, like duration and pitch (sd, mean, min and max); Turn and Time Management very often share the same vocabulary and some prosodic properties, like intensity and standard deviation in pitch; Contact Management and Discourse Structuring are close due to the shared vocabulary.

Concerning the very simple fourth test, Table 12 shows that each dimension may be addressed by an utterance without any other dimension being addressed. This proves that each of the defined dimensions exists independently, and is an autonomous aspect of communication.

Finally, we investigated the occurrences of dimension tags given the tag of the previous utterances in order to find out whether there are dependencies in using utterances addressing a certain communicative aspect and if a particular dimension is addressed previously. We took the range of 5 previous utterances saved in dialogue history, because there is often more distance between related utterances in multi-party interaction (e.g. AMI) than in two-party dialogues. Table 13 shows that there is no evident dependence in dimensions relations across the dialogue history; there is no need for the speaker to address a particular aspect of communication as a response to partner's previous contributions. There are certainly some observed logical patterns. For example, retractions and self-corrections often follow hesitations because the speaker, while monitoring his own speech and noticing that the utterance of part of what he just produced needs revision, needs some time before he continues with the improved part.

# 9 Conclusions and Discussion

In this report we discussed the notion of dimension as an aspect of communication which an utterance can address in a dialogue context. Five criteria were defined for including a dimension in an annotation scheme: (1) theoretically and (2) empirically motivated; (3) recognized by human annotators and automatically; (4) reflected in existing annotation schemes; and (5) independently addressable. Table 14

| | Task | AuF | AlF | Turn | Time | Contact | DS | OCM | PCM | SOM |
|---|---|---|---|---|---|---|---|---|---|---|
| Task | 21.2 | 27.4 | 27.7 | 20 | 32.5 | 0 | 7.1 | 16.4 | 15.2 | 32.1 |
| AuF | 15 | 24.4 | 25 | 21.4 | 15.4 | 27.8 | 12.3 | 7.5 | 22.7 | 12.8 |
| AlF | 0.4 | 1.3 | 5.6 | 0.5 | 0.5 | 0 | 0.6 | 0.4 | 0 | 0 |
| Turn | 14.3 | 4.7 | 0 | 6.5 | 5.2 | 0 | 6.5 | 2.2 | 7.6 | 6.4 |
| Time | 22.2 | 16.3 | 16.7 | 23.5 | 15 | 0 | 35.5 | 47.1 | 37.9 | 19.2 |
| Contact | 0 | 0.1 | 0 | 0.2 | 0 | 27.8 | 0 | 0 | 0 | 0 |
| DS | 2 | 2 | 0 | 0.5 | 0.5 | 27.8 | 5.2 | 0.4 | 0 | 0 |
| OCM | 7.7 | 6.3 | 5.6 | 7.7 | 11.2 | 0 | 0 | 7 | 0 | 0 |
| PCM | 0.4 | 0.4 | 0 | 0 | 0.08 | 0 | 0 | 0.2 | 0 | 0 |
| SOM | 0.1 | 0.3 | 0 | 1.2 | 0.08 | 0 | 0.6 | 0 | 0 | 6.4 |

Table 13: *Overview of relative frequency (in%) of dimensions given the dimensions addressed by previous utterances observed in AMI data, per dimension, using the last 5 utterances in the dialogue history.*

| | Theoretically validated | Observed in dialogue corpus data | Recognizable by human annotator and automatically | Reflected in the majority of existing DA taxonomies | Dependency tests passed |
|---|---|---|---|---|---|
| Task | + | + | + | + | + |
| Task Management | + | + | - | - | **na** |
| Auto-Feedback | + | + | + | + | + |
| Allo-Feedback | + | + | + | +/- | + |
| Turn Management | + | + | + | + | + |
| Social Obligation Management | + | + | + | + | + |
| Own Communication Management | + | + | + | + | + |
| Discourse Structuring | + | + | + | + | + |
| Partner Communication Management | + | + | + | - | + |
| Time Management | + | + | + | + | + |
| Contact Management | + | + | + | - | + |

Table 14: *Summary of survey and testing results in identifying the proper dimension set.*

gives an overview of the results of our investigations with respect to these criteria.

The analysis shows that eight dimensions, namely **Task**, **Feedback**, **Turn Management**, **Social Obligations Management**, **Own Communication Management**, **Discourse Structuring**, **Partner Communication Management** and **Time Management** fulfil all five criteria, and can be considered as 'core' aspects of dialogue communication. They have been studied extensively, from both theoretical and practical points of view; they are observed in actual dialogues; they are reliably annotated and successfully classified automatically; they are defined in most existing annotation schemes; and they address a certain aspect of communication independently of others.

Our conclusion with respect to Feedback is moreover that a distinction should be made at least between **Feedback giving** and **Feedback eliciting** aspects, since dialogue participants not only report about successes and failures of their own processing of previous utterances, but also constantly evaluate the partner's cognitive state, message processing, and degree of involvement in the communication, and may elicit information about these aspects. Making only the distinction between feedback-giving and feedback-eliciting acts, however, does not to justice to the fact that feedback-giving acts can report not only on the speaker's own processing of previous dialogue but also on the speakers beliefs about the addressee's processing - a distinction which is semantically important and which is captured by the distinction between Auto- and Allo-Feedback. Note also that the phi-coefficient (-0.3) indicates that Auto- and Allo-Feedback are not very closely related. These arguments support the suggestion to distinguish the two as separate dimensions.

Time Management was shown to be a 'core' dimension as it meets all five criteria. There are different opinions, however, between researchers as to whether it should be considered as a separate dimension on its own. Communicative functions defined for Time Management seem to be closely related to Own Communication Management when the speaker, monitoring his own dialogue contribution, es-

timates that he/she needs some more time to produce an utterance, which leads to hesitations expressed by filled or unfilled pauses. On the other hand, several types of pauses may have other reasons than own communication monitoring and management. For example, a speaker might need some time for opening a file, or consulting or making notes, or he might be distracted by the partner's (lack of) )activity and wants to get his attention by producing an extensive amount of stallings. Note also the the phi-measure (-0.3) indicates that Time Management and Own Communication Management are not closely related. So there are good arguments for keeping the two apart.

Time Management acts are also close to Turn Management acts, since speakers often need a bit of time to formulate their contribution when they take (or have and want to keep) the turn. This consideration applies only to *stallings* under certain context conditions, however; *pausing*, by contrast, does not imply that the speaker wants to keep the turn. It should be also noticed that *stallings* do not always imply that the speaker wants to keep the turn; extensive amounts of protraction accompanied by certain non-verbal behaviour may indicate that the speaker needs assistance. It was noticed by Butterworth [27] that an excessive amount of gaze aversion may also lead a listener to infer that the speaker is having difficulty formulating a message. Moreover, as Clark in [32] shows, time delays are not always are used for turn-keeping purposes, because even in monologues where speakers do not need to keep the turn, time delays are frequently used. Time and Turn Management are therefore better kept apart rather than considered as one dimension.

A third view on **Time Management** acts is that they are produced unintentionally, *stallings* in particular. They should therefore perhaps not be regarded as dialogue acts. An act that is not consciously intentional may still be relevant, however; for example, humans produce a lot of facial expressions unconsciously, but they display the emotional or cognitive state of the dialogue participant, which is obviously important for dialogue analysis. In other words, they affect the information states of dialogue participants if they have shared encoded meaning. Goffman [47] points out that the receiver is always responsible for the interpretation of an act as intentional or not. Kendon [56] also notices that whether an action is deemed to be intended or not is something that is dependent entirely upon how that action appears to others. So this does not provide a good argument against viewing Time Management as a dimension of dialogue communication.

**Partner Communication Management** also satisfies all criteria, although it is not recognized in many existing annotation schemas. This is perhaps related to its relatively low frequency in many types of dialogue (but notice its substantial frequency in the OVIS corpus; see Table 12). Some dialogue act taxonomies regard these functions as Allo-Feedback functions, claiming that *completion* and *correct-misspeaking* reflect the speaker's processing of the partner utterance(-s). We rather think that completions and correct-misspeakings *imply* positive Auto-Feedback, since one can only correct or complete what the current speaker is saying if one believes to have understood what has been said up to this point, just like Auto-Feedback is implied by an *answering* a question or *accepting*/*declining* an offer or request. This is confirmed by the co-occurrence data in Table 9, which show that 65% of all Partner Communication Management acts imply an Auto-Feedback act. Note also that the low phi-coefficient (0.1) indicates that Partner Communication Management and Allo- or Auto-Feedback are not closely related.

Our conclusion is that **Contact Management** could be considered as 'optional' dimension, since this aspect of communication is not reflected in most existing dialogue act annotation schemes (6 out of 18). It was noticed, however, that for some types of dialogues, e.g. phone conversations or teleconferences (as in the OVIS corpus), this aspect may be important.

# References


[1] Allen, J. (1983). *Recognising intensions from natural language utterances*. In: Brady, M., and Berwick, R.C. (eds). Computational Models of Discourse. MIT Press, Cambridge, MA, pp. 107–166.

[2] Allen, J. et al. (1994). *The TRAINS Project: a case study in building a conversational planning agent*. TRAINS Technical Note 94-3. University of Rochester.

[3] Alexandersson, J., et al. (1998). *Dialogue acts in Verbmobil-2*. Second edition. Report 226. DFKI Saarbruecken, University of Stuttgart; TU Berlin; University of Saarland.

[4] Allen, J., Core, M. (1997). *Draft of DAMSL: Dialog Act Markup in Several Layers*. Available at `http://www.cs.rochester.edu/research/cisd/resources/damsl/`.

[5] Allwood, J. (1992). *On dialogue cohesion*. Gothenburg Papers in Theoretical Linguistics 65, Gothenburg University, Department of Linguistics.
[6] Allwood, J. (1994). *Obligations and options in dialogue*. In: THINK Quarterly 3(1), pp. 9-18.
[7] Allwood, J. (2000). *An activity-based approach to pragmatics*. In: Abduction, Belief and Context in Dialogue, H. Bunt and W. Black (eds.), Amsterdam: John Benjamin, 47-81.
[8] Allwood, J. (2000). *Bodily Communication - Dimensions of Expression and Content.* In: Multimodality in Language and Speech Systems, B. Granstroem, D. House and I. Karlsson (eds.), Kluwer Academic Publishers, Dordrecht.
[9] Allwood, J., Ahlsèn, E., Nivre, J. and Larsson, S. (1997). *Own Communication Management: Kodningsmanual*. Goeteborg University: Department of Linguistics.
[10] Allwood, J., Nivre, J. and Ahlsèn E. (1993). *On the Semantics and Pragmatics of Linguistic Feedback*. Journal of Semantics, 9-1: 1-26.
[11] Allwood, J. and Traum, D. and Jokinen, K. (2000). *Cooperation, dialogue and ethics.* In: International Journal of Human Computer Studies, pp. 871–914.
[12] Allwood, J., Ahlsèn, E., Lund, J., and Sundqvist, J. (2005). *Multimodality in own communication management*. In: Proceedings from the Second Nordic Conference on Multimodal Communication, Göteborg.
[13] Augmented Multi-party Interaction Concortium. (2005). *Guidelines for Dialogue Act and Addressee Annotation Version 1.0*.
[14] Ascher, N., and Lascarides, A. (2003). *Logics of Conversation*. Campbridge University Press.
[15] Bales, R.F. (1951). *Interaction process analysis: a method for the study of small groups*. Addison-Wesley, Cambridge.
[16] Bavelas, J.B., and Chovil, N. (2000). *Visible acts of meaning. An integrated message model of language in face-to-face dialogue*. Journal of Language and Social Psychology, 19, pp. 163–194.
[17] ten Bosch, L., Oostdijk, N., and de Ruiter, J.P. (2004). *Durational aspects of turn-taking in spontaneous face-to-face and telephone dialogues.* In: P. Sojka, I. Kope ek, and K. Pala (eds.), Proceedings of the 7th International Conference, TSD 2004. Lecture Notes in Artificial Intelligence LNCS/LNAI 3206, Brno, Czech Republic, September 2004. Spinger Verlag. pp. 563–570.
[18] Brown, P., and Levinson, S. C. (1987). *Politeness - Some universals in language usage*. Cambridge University Press, 1987.
[19] Bunt, H. (1994). *Context and dialogue control*. THINK Quarterly 3(1), pp. 19–31.
[20] Bunt, H. (1996). *Interaction management functions and context representation requirements*. In: Luperfoy, S., Nijholt, A., and Veldhuizen van Zanten, G. (eds.), Dialogue Management in Natural Language Systems, pp.187–198.
[21] Bunt, H. (2000). *Dialogue pragmatics and context specification.* In: Abduction, Belief and Context in Dialogue; studies in computational pragmatics, H. Bunt and W. Black (eds.).John Benjamins, Amsterdam, pp.81–105.
[22] Bunt, H. (2004). *A Framework for Dialogue Act Specification*. First Workshop on Interoperable Semantic Annotation, Tilburg, January 2004. Available at `http://let.uvt.nl/research/ti/sigsem/wg`.
[23] Bunt, H. (2006). *Dimensions in Dialogue Act Annotation*. Proceedings LREC 2006, Genova.
[24] Bunt, H. (2007). *Multifunctionality and Multidimensional Dialogue Act Annotation*. In: E. Ahlsèn et al. (ed.) Communication - Action - Meaning, A Festschrift to Jens Allwood. Göteburg University Press, August 2007, pp. 237 – 259.
[25] Bunt, H., and Girard, Y. (2005). *Designing an open, multidimensional dialogue act taxonomy*. In: Gardent, C., and Gaiffe, B. (eds). Proceedings of the Ninth Workshop on the Semantics and Pragmatics of Dialogue, Nancy, June 2005, pp. 37–44.
[26] Bunt, H., and Schiffrin, A. (2007). *Documented compilation of semantic data categories*. Deliverable D4.3.
[27] Butterworth, B. (1980). *Evidence from pauses in speech.* In: Language Production: Speech and Talk, vol.1, B.Butterworth(eds.). Academic Press, London, pp. 155–177.
[28] 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, Scotland.
[29] Carletta, J. (1996). *Assessing agreement on classification tasks: The kappa statistic*. Computational Linguistics, 22(2):249-254.
[30] Cassell, J. and Torres, O. and Prevost, S. (1999). *Turn Taking vs. Discourse Structure: How Best to Model Multimodal Conversation*. Machine Conversations, Wilks, Y.(eds.). Kluwer, The Hague, pp. 143–154.
[31] Cassell, J., Nakano, Y.I., and Bickmore, T.W. (2001). *Non-Verbal Cues for Discourse Structure*. In: Proceedings of Association for Computational Linguistics Annual Conference (ACL), pp. 106–115.
[32] Clark, H. (1996). *Using Language*. University Press, Cambridge, UK.
[33] Clark, H., and Fox Tree, J. (2002). *Using 'uh' and 'um' in spontaneous speech.* Cognition, 84, pp. 73–111.
[34] Clark, H. and Krych, M.A. (2004). *Speaking while monitoring addressees for understanding*. Journal of Memory and Language, 50: 62-81.
[35] Cohen, R. (1984). *A computational theory of the function of clue words in argument understanding*. Coling-ACL 1984, Standford, pp. 251–258.
[36] Cohen, J. (1960). *A coefficient of agreement for nominal scales*. Education and Psychological Measurement, 20:37-46.
[37] Cohen, W. W.(1995).*Fast effective rule induction*. In: Proceedings of the 12th International Conference on Machine Learning (ICML'95), pp. 115–123.
[38] Core, M.G., and Allen, J.F. (1997). *Coding dialogues with the DAMSL annotation scheme*. In: Working Notes: AAAI Fall Symposium on Communicative Action in Humans and Machines, pages 28–35.
[39] Dahlbaeck, N., and Jonsson, A. (1998). *A coding manual for the Linköping dialogue model*. Unpublished manuscript.
[40] Di Eugenio, B. Jordan, P.W., and Pylkkaenen, L. (1998). *The COCONUT project: dialogue annotation manual*. ISP Technical Report 98-1.

[41] Dhillon, R., Bhagat, S., Carvey, H., and Schriberg, E. (2004). *Meeting recorder project: dialogue labelling guide*. ICSI Technical Report TR-04-002.
[42] Duncan, S. and Fiske, D. (1977). *Face-to-face interaction: research, methods, and theory*. Lawrence Erlbaum Associates, New Jersey.
[43] Garrett, M.F. (1980). *Levels of processing in sentence production*. In: Language Production: Speech and Talk, vol.1, B.Butterworth(eds.), Academic Press, London, pages 177–221.
[44] Geertzen, J., Petukhova, V. and Bunt, H. (2008). *Evaluating dialogue act tagging with naive and expert annotators*. In: Proceedings of the Sixth International Conference on Language Resources and Evaluation, Marrakech, Morocco.
[45] Geertzen, J., and Bunt, H.. (2006). *Measuring annotator agreement in a complex hierarchical dialogue act annotation scheme*. In Proceedings of the 7th SIGdial Workshop on Discourse and Dialogue. Sydney, Australia.
[46] Geertzen, J., Petukhova, V., and Bunt, H. (2007). *A Multidimensional Approach to Utterance Segmentation and Dialogue Act Classification*.Proceedings of the 8th SIGdial Workshop on Discourse and Dialogue, Antwerp, pages 140–149.
[47] Goffman, E. (1963). *Behavior in Public Places*. Basic Books, New York.
[48] Goldman-Eisler, F. (1968). *Psycholinguistics: experiments in spontaneous speech*. New York, Academic Press.
[49] Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers.* Academic Press. New York.
[50] Grosz, B. J. and Sidner, C. L.(1986). *Attention, Intentions, and the Structure of Discourse*. Computational Linguistics, Vol.12, pp. 175–204.
[51] Heeman, P.A. and Allen, J.F. (1999). *Speech repairs, intonational phrases and discourse markers: Modelling speakers utterances in spoken dialogue*. Computational Linguistics, 12(3): 1–45.
[52] Hirschberg, J. and Litman, D. (1993). *Empirical studies on the disambiguation of cue phrases*. Computational Linguistics, 25(4): 501–530.
[53] Ichikawa, A. et. al. (1998). *Standardising Annotation Schemes for Japanese Discourse*. In: Proceedings of the First International Conference on Language Resources and Evaluation, Spain.
[54] ISO. (2008). *Language resource management – Semantic annotation framework – Part 2: Dialogue acts*. ISO document ISO/TC 37/SC 4 N442 rev 03. ISO Central Secretariat, Geneva.
[55] Jurafsky, D., Schriberg, E., and Biasca, D. (1997). *Switchboard SWBD-DAMSL Shallow-Discourse-Function Annotation: Coders Manual*, Draft 13. University of Colorado.
[56] Kendon, A. (2004). *Gesture: visible action as utterance*. Cambridge University Press.
[57] Klein, M., et al. (1998). *Supported Coding Schemes*. MATE Deliverable D1.1.
[58] Klein, M., and Soria, C. (1998). Dialogue acts. In: Klein, M. et al. (eds.) *MATE Deliverable 1.1: Surported Coding Schemes*. MATE European Project LE4-8370.
[59] Klein, M. (1999). *Standardization efforts on the level of dialogue act in the MATE project*. Available at `http://acl.ldc.upenn.edu/W/W99/W99-0305.pdf`.
[60] Lakoff, R. (1973). *The logic of politeness: or minding your P's and Q's*. In: Corum, C., Smith-Stark, T.C., and Weiser A. (eds.). Papers from the Ninth Regional Meeting of the Chicago Linguistic Society. Chicago, Chicago Linguistic Society, pp. 292–305.
[61] Larsson, S. (1998). *Coding Schemas for Dialogue Moves*. Technical report from the S-DIME project.
[62] Louwerse, M., and Mitchell, H. (2003). *Toward a taxonomy of a set of discourse markers in dialogue: A theoretical and computational linguistic account*. Discourse Processes 35 (3), 243-281.
[63] Leech, J., McEnery, T., and Weisser, M. (2003). *SPAAC Speech-Act Annotation Scheme*. University of Lancaster.
[64] Levelt, W.J.M. (1989).*Speaking: From Intention to Articulation*. Cambridge, Massachusetts, etc.: MIT Press.
[65] Maclay, H., and Osgood, C.E. (1959). *Hesitation phenomena in spontaneous English speech*. Word, 15, pp. 19–44.
[66] Mann, W. and Thompson, S. (1988). *Rhetorical structure theory: toward a functional theory of text organisation*. The MIT Press, Cambridge, MA.
[67] Nakatani, C., and Traum, D. (1999). *Draft: Discourse structure coding manual*. (version 1.0). Technical Report UMIACS-TR-99-03, University of Maryland.
[68] Nivre, J., Allwood, J. and Ahlsèn, E. (1998). *Interactive Communication Management: Coding Manual*. Version 1.0. Göteborg University: Department of Linguistics.
[69] Petukhova, V. and Bunt, H. (2007). *A multidimensional approach to multimodal dialogue act annotation*. In: Proceedings of the Seventh International Workshop on Computational Semantics (IWCS), pp. 142–153.
[70] Popescu-Belis, A. (2003). *Dialogue act tagsets for meeting understanding: an abstraction based on the DAMSL, Switchboard and ICSI-MR tagsets*. Technical report, IM2.MDM-09, v1.1.
[71] Popescu-Belis, A. (2005). *Dialogue Acts: One or More Dimensions?* ISSCO Working Paper 62, ISSCO, Geneva. November 2005. Available at: `http://www.issco.unige.ch/publicaitons/working-papers/papers/apb-issco-wp62b.pdf`
[72] Rienks, R., and Verbree, D. (2005). *Twente Argument Schema Annotation Manual v 0.99b*. University of Twente.
[73] Roger, D. and Bull, P. (1989). *Conversation: an interdisciplinary perspective*. Multilingual Matters, Clevedon.
[74] Sacks, H. and Schegloff, E. and Jefferson, G. (1974). *A simplest systematics for the organization of turn-taking for conversation*. Language, 50(4), pp. 696–735.
[75] Schiffrin, D. (1987). *Discourse Markers*. Cambridge: University Press.
[76] Selting, M. (2000). *The construction of units in conversational talk*. Language in Society, 29, pp. 477–517.
[77] Smith, V.L., and Clark, H. (1993). *On the course of answering questions*. Journal of Memory and Language, 32, pp. 25–38.
[78] Stenstroem, A.-B. (1990). *Pauses in monologue and dialogue*. In: Svartvik, J. (ed.). London-Lund Corpus of Spoken English: Description and Research. Lund, Lund University Press.

[79] Swerts, M., and Ostendorf, M. (1997).*Prosodic and lexical indications of discourse structure in human-machine interactions*. Speech Communication, Vol. 22 , pp. 25–41.
[80] Traum, D. R., and Hinkelman E. A. (1992). *Conversation Acts in Task-oriented Spoken Dialogue*. Computational Intelligence, 8(3), pp. 575–599.
[81] Traum, D. (1999). *Computational models of grounding in collaborative systems.* In Brennen, S.E., Giboin, A., and Traum, D. (eds). Working Papers of the AAAI Fall Symposium on Psychological Models of Communication in Collaborative Systems, Menlo Park, CA, USA. American Association for Artificial Intelligence, pp. 124 – 131.
[82] Li-chiung Yang. (2001). *Visualizing Spoken Discourse: Prosodic Form and Discourse Functions of Interruptions*. In: Proceedings 2nd SigDial Workshop on Discourse and Dialogue, Aalborg, Denmark.
[83] Vark van, R.J., Vreught de, J.P.M., and Rothkrantz, L.J.M. (1996). *Analysing OVR dialogue coding scheme 1.0*. Report 96-137.
[84] Weiqun Xu, Carletta, J., Kilgour, J., and Karaiskos, V. (2005). *Coding Instructions for Topic Segmentation of the AMI Meeting Corpus (version 1.1)*. School of Informatics, University of Edinburgh.