Chapter 19

Computational Pragmatics

Harry Bunt

1 General and Computational Pragmatics

Pragmatics has been characterised as the study of the relations between linguistic properties of utterances on the one hand, and aspects of the context in which a given utterance is used on the other hand (Bunt and Black, 2000). Computational pragmatics is pragmatics with computational means, which include corpus data, context models, and algorithms for context-dependent utterance generation and interpretation.

Relations between linguistic form and context of use are apparent for example in the definiteness of a noun phrase, encoding a speaker's assumption about beliefs shared with the hearer (presuppositions*); in the way the sentence structure reflects the speaker's intention to structure information into a part believed to be new for the hearer and a part assumed to be known (given-new* structuring); or in a particular word order indicating which information the speaker wants the hearer to focus on (topic-focus* marking).

The linguistic side of the relations that are studied in pragmatics is formed by utterances in a conversation or sentences in a written text. The context side consists in the case of written text of the surrounding text and the setting in which the text is meant to function. In dialogue, the context of an utterance is likewise formed by what has been said before and the interactive setting, but additionally also by perceptual information, e.g. which objects and events are visible for the participants; by social obligations, such as the pressure to return a greeting; and by the epistemic context, e.g. what do the participants know about the topic of the conversation and about each other. Much of this information is dynamic, as it changes during a dialogue and, more importantly, *as a result of* the dialogue, since the participants in a conversation influence each other's state of information. Dialogue contexts are thus updated continuously as an effect of communication.

The use of large corpora, supported by software tools for searching and querying the data, has the effect that studies in pragmatics can be informed more systematically by empirical observations. This is evidently beneficial for developing pragmatic theories with better coverage and validity. Such developments may well go beyond minor adjustments to existing theories; they may lead to a rethinking of pragmatic concepts and theoretical frameworks. An example is the development of the theory of dialogue acts, originating from speech act theory, under the influence of spoken dialogue analysis through corpus collection and computational modelling. Modern dialogue act theory deals with very different and richer sets of action types then traditional speech act theory, and uses a different, more complex and articulate notion of communicative action.

Pragmatics has also been characterised as being concerned with making explicit what a speaker leaves implicit. Conversational implicatures* and indirect speech acts* are two

well-known cases where the speaker means something else or something more than what he says.

For example, suppose I am meeting a student to discuss an essay that she had handed in, and that I praise the student for the attractive lay-out, for the catchy title, for its length, for the pretty diagrams that it contains, for the nice colours in the screen shots, for the good quality of the paper on which it is printed,…. and the students starts to worry: by praising peripheral aspects of the essay, an implicature emerges (from the apparent violation of the Gricean maxim of relevance*) that I'm not very positive about the contents. Note that this implicature is specific for the context in which conversation takes place; if I would say similar things to an assistant in a copyshop, then no such implicature would arise.

Indirect speech acts, as illustrated by the classical example *"Can you pass me the salt?",* also illustrate how a speaker can mean more ("Please pass me the salt") than he says, by relying on the addressee to infer the intended meaning in the given context. Again, this indirect interpretation arises only in certain contexts; in the context of a rehabilitation clinic where passing the salt is a common physical exercise, that interpretation would be unlikely to occur.

Conversational implicatures and indirect speech acts illustrate that speakers can convey something more or something else that what they say by relying on the addressee's ability to infer the intended interpretation by combining information from the utterance with information from the context, using Grice's general cooperative principle* and the more specific maxims* to guide the inference process.

Context-based inference is needed not only to understand an indirect speech act or a conversational implicature, but in general to understand *any* utterance in a conversation. A simple illustration is formed by utterances of *"Yes",* which can be a positive answer to a propositional question, or a positive feedback act, or the acceptance of an offer, or an expression of agreement, or several other things. Without taking the preceding dialogue into account there is no way of interpreting the meaning of this utterance. An interesting, more intricate example can be found in the following fragment of a dialogue between the information service at Amsterdam Airport (I) and a client (C). The utterances 3 and 4 have different meanings in spite of being identical; their correct understanding clearly depends crucially on information which is not in the utterances themselves but in the context.

(1) 1. C: what departure times do you have for flights to Munich early in the morning?

2. I: I have Lufthansa at 07:15, KLM at 07.25, again KLM at 07:50, and another Lufthansa at 08:20

3. C: and that's on Saturday too

4. I: and that's on Saturday too

Utterance 1.3 is an example of what has been called a *declarative question*, i.e. a question expressed by a declarative sentence. Beun (1989) found that in a corpus of spoken information-seeking dialogues about 20% of the questions had this form, and that the participants, as well as subjects participating in recognition experiments, recognised

these utterances perfectly well as questions (more specifically, as check questions), even though like statements. The understanding of an utterance like 1.3 as a question can be explained as follows.

1. A participant in an information-seeking dialogue may have several reasons for performing a communicative action with content *p*: he wants to know whether *p*; or he wants the other participant to know that *p*; or he believes that the other participant would like to know whether *p*.

2. In a dialogue setting of this kind, the client believes the information agent to be an expert concerning the domain, so for every proposition *p* about the discourse domain *D*, *C* believes that *I* knows whether *p* (and that *I* knows this 'better' than *C*). It therefore makes no sense to interpret *C*'s utterance as a statement.

3. Likewise, it cannot be the case that *C* contributes the utterance because he believes that *I* wants to know whether *p*, hence it makes no sense to interpret *C*'s utterance as an answer.

4. Looking for an interpretation that does make sense (in view of the maxim of relevance), a remaining possibility it is that *C* contributes this utterance because he wants to know whether *p*. This means that the essential condition for *C* to ask a question about *p* is fulfilled.

5. As already noted in 2, *C* believes that *I* knows whether *p*. The most important supporting condition for *C* to ask *I* whether *p* is thus also fulfilled. Therefore the utterance can be interpreted as the question whether *p*.

Note that interpretation not only of utterance 1.3 as a (check) question but also of utterance 4 as a confirmation requires inferencing, even though none of these utterances exemplify indirect speech acts, conversational implicatures, or other phenomena that are well known to require inferencing. In fact, the interpretation of utterances in a conversation or in a running text *in general* involves inferencing; it is the rule, rather than the exception, that interpretation and inferencing are interlocking processes.

19.2 Inferencing in computational pragmatics

Inferencing can take a variety of forms, which can be divided into d*eductive, abductive,* and *inductive* reasoning. *Deductive* reasoning is the form of inferencing that has been defined by logicians for combining facts and hypotheses in order to construct proofs of their logical consequences. This form of reasoning is important in science, but apart from Sherlock Holmes no mortals seem to apply this form of reasoning much in their daily lives.

Abduction works in the reverse direction: given an observation, abductive inference yields hypotheses that could explain the observation. For example, given the observation that the street is wet and the general fact that streets get wet when it rains, abduction yields as a possible explanation the hypothesis that it rains. Different from the consequences derived by deduction, the hypotheses derived by abduction are not logically valid - they may be wrong. For example, the street might be wet because a

water pipe has burst, or because the fire brigade has held an exercise. Abduction is a form of inferencing that people seem to apply all the time in order to interpret and explain what they see and hear.

Induction is inferencing by generalising from examples. For instance, from encountering lots of examples of white swans, and never seeing a swan which is not white, one may induce that swans are white. Of course, one will revise one's opinion upon seeing a black swan; like the results of abductive inference, the results of inductive inference may be wrong; they are not logically valid. And like abduction, induction is a form of inferencing that people commonly use. Induction is especially important for constructing models of the world, while abduction is important for the interpretation of observations, in particular for the understanding of communicative behaviour.

Human inference in everyday situations is shallow rather than deep, employs vast amounts of general as well as situation-specific knowledge, and aims at plausible and useful rather than logically valid results. A fundamental issue in the construction of plausible interpretations is the question what makes one interpretation more plausible than another. Consider again the observation that the street is wet (O_1), which would be entailed both by the circumstance that it is raining (P_1) and by the circumstance that a water pipe has burst (P_2). According to abductive inference, both P_1 and P_2 are possible explanations of O_1. Most people would say that P_1 is more plausible, due to the fact that it rains more often than that a water pipe bursts. Note that this is only the case in the absence of further information; if the observation O_1 occurred in a context where it only rarely rains, but where bursting water pipes are a notorious problem, then P_2 could be the more plausible explanation. So frequency of occurrence apparently has an influence on the plausibility of interpretations.

Stickel (1988) implemented a form of abduction called *weighted abduction*, which uses numerical plausibilities built into the inference process. Propositions involved in inferences are given a *cost*. For example, when a rule is applied of the form $P \Rightarrow Q$ then Q will cost more than P, so shorter proofs cost less than longer ones. Proofs with lower cost are considered to provide more plausible explanations than proofs with higher cost. The framework called *'Interpretation as Abduction'*, developed by Hobbs and associates (Hobbs, 1990; Hobbs et al., 1993) applies weighted abduction to explain a variety of context-dependent semantic and pragmatic phenomena.

As an illustration of what this form of abduction can do, Hobbs et al. (1993) use the example (3), which displays three phenomena that require reasoning with context information in order to arrive at a good understanding: (1) the use of a definite article; (2) the unspecified relation in the nominal compound *"Boston office";* and (3) the metonymical relation connecting an office and the act of making a phone call.

(2) The Boston office called.

Concerning the interpretation of the nominal compound, compositional semantic analysis may produce a logical form of the following kind, where the predicate *NN* represents the unknown relation between Boston and the office, and the variables *x* and *y* should be understood as existentially quantified:

(3) *boston'(x) & office'(y) & NN(y,x)*

The knowledge base against which this interpretation takes place contains among other facts the existence of a certain office *B1*, located in Boston, i.e. it contains the following facts:

(4) *office'(B1), located-in(B1,B), boston'(B)*

It also contains the general knowledge that location is a possible implicit relation between the elements of a nominal compound, i.e. it contains the fact *located-in(x,y)* ⇒ *NN(x,y).* From these facts, abductive inference does indeed construct the interpretation that the office mentioned in (3) can be understood as the office *B1*. There may of course be other offices in Boston than *B1*, and other known offices with other relations to Boston, which would provide alternative interpretations of (3). This is where the weights of different possible interpretations come into play, and determine which interpretation will come out as most plausible.

This example illustrates that the inferencing in language understanding systems typically combines situation-specific and general world knowledge. Ovchinnikova et al. (2014) have presented a system based on weighted abduction, called mini-TACITUS (re-implementing Hobbs' TACITUS system, Hobbs, 1986) which incorporates a knowledge base extracted from the large-scale resources WordNet (Fellbaum 1998) and FrameNet (see Ruppenhofer et al., 2006). The '*synsets*' of which WordNet is made up correspond to word senses; a lexeme can participate in several synsets. For every word sense, its frequency in the WordNet annotated corpora is indicated. Ovchinnikova et al. use the lexeme-synset mapping for generating axioms (with the frequencies of word senses converted into axiom weights). For example, the verb *compose* is mapped into its sense *synset-X* (in WordNet version 3.0), which represents one of its senses, as shown in (5).

(5) $synset\text{-}X(s,e) \Rightarrow compose(e, x_1, x_2)$

FrameNet represents the lexical meaning of predicates in terms of frames that describe prototypical situations. Every frame contains a set of roles corresponding to the participants in the situations that may be described. Moreover, syntactic patterns show the surface realisation of verbs and their arguments and contain information about their frequency in the FrameNet annotated corpora. From these patterns and their frequencies, weighted axioms are derived, which together with the axioms derived from WordNet form a large base of common-sense knowledge about the world as we speak about it. Preliminary experiments with this approach show promising possibilities for making clear exactly how knowledge about the world and abductive inference can explain such phenomena as conversational implicatures, understanding of nominal compounds, and metonymy.

The mini-TACITUS system and its knowledge base illustrate the three most important kinds of tools that computational pragmatics brings to general pragmatics: (1) algorithms, in this case for making inferences; (2) data, notably digital corpora with utterances annotated with pragmatically relevant properties, in this case WordNet and FrameNet corpora; and (3) computational representations of knowledge and context, in this case the general knowledge base derived from these corpora.

19.3. Language as action in context

19.3.1 Speech acts and dialogue acts

A fundamental contribution from the study of spoken language is the insight that utterances are best viewed as actions. When we talk, we *greet, ask questions, apologise, answer, make promises,* and so on – we perform *speech acts.* While specific syntactic structures, lexical items, and prosodic forms can be used to encode some of these action types, in general the communicative functions of an utterance cannot unambiguously be derived from its linguistic form, but requires taking the context into account. This point was illustrated by utterances 3 and 4 in example (1) above.

In computational pragmatics the notion of a *dialogue act* (Bunt, 1979; 1989) has become popular for modelling the use of language as the performance of actions. While the term 'dialogue act' is sometimes understood in the sense of 'speech act used in dialogue', the distinction between the concepts of speech act and dialogue act is more fundamental. Where speech act theory is an approach to meaning within the philosophy of language (Austin, 1962; Searle, 1969), the theory of dialogue acts is an empirically-based approach to the modeling of linguistic, nonverbal, and multimodal communicative behaviour, and considers dialogue acts as acts with an articulate internal structure, related to its functions in a dialogue. A formal definition of the dialogue act concept, as used in ISO standard 24617-2 for dialogue annotation, is as follows:

(6) *A dialogue act is a stretch of communicative activity of a dialogue participant, interpreted as having a certain communicative function and a semantic content, and which may additionally have certain functional dependence relations, rhetorical relations, and feedback dependence relations.* (ISO 24617-2:2012; cf. Bunt et al., 2010).

Dialogue acts offer a way of characterising the intended meaning of communicative behaviour computationally in terms of update operations on the information states of dialogue participants; this approach is known as the 'information-state update' or 'context-change' approach (Bunt, 1989; 2000; Traum and Larsson, 2000; 2003). For example, when an addressee understands the utterance "*Do you know what time it is?"* as a question about the time, then the addressee's information state is updated to contain (among other things) the information that the speaker does not know what time it is and would like to know it. If, by contrast, it is understood that the speaker is *reproaching* the addressee for being late, then the addressee's information state is updated to include (among other things) the information that the speaker *does* know what time it is. Distinctions such as that between a question and a reproach concern the *communicative function* of a dialogue act, which is one of its two main components. The other main component is its *semantic content*, which describes the objects, properties, relations, situations, actions or events that the dialogue act is about. The communicative function of a dialogue act specifies how an addressee updates his information state with the information expressed in the semantic content, when he understands the dialogue act.

The major differences between speech act theory and dialogue act theory are the following:

a. speech act theory is concerned with verbal behaviour; dialogue act theory applies also to nonverbal and multimodal behaviour;

b. speech act theory assumes that every utterance encodes one speech act; dialogue act theory assumes utterances to be multifunctional;

c. dialogue acts have a computational semantics defined in terms of update operations on dialogue participants' information states;

d. dialogue acts are not considered in semantic isolation, but as dependent on other dialogue acts, to which they mave have semantic and pragmatic relations.

These differences are discussed in the remainder of this section.

19.3.2 Dialogue segmentation

A spoken dialogue is naturally segmented into turns, defined as stretches of speech produced by one participant and bounded by periods of silence of that participant. The 'turn' notion is closely related to that of having 'the floor' (Sacks et al. 1974), or the 'speaker role', which has been defined as *role of a participant who has temporary control over the dialogue and speaks for some time* (DAMSL Revised Manual, 1997).

Turns can have a complex structure, as the following example illustrates (Allwood, 1992):

(8) A: Yes! Come tomorrow. Go to the church. Bill will be there. OK?
B: The church, OK.

A's turn contains sequentially the five functions *feedback giving, request, request, statement and response elicitation*. This shows that dialogue acts often correspond to segments that are smaller than turns.[1]

Utterances are mostly understood to be contiguous stretches of linguistic behaviour which satisfy some well-formedness constraints, e.g. being a grammatical unit such as a sentence, clause, or phrase. Syntactic and prosodic features are often used as indicators of utterance endings, but the detection of utterance boundaries is very hard in general (see e.g. Shriberg et al., 1998; Stolcke et al., 2000; Nöth et al., 2002). Using grammatically and/or prosodically defined units of segmentation runs into the problems that (a) not every grammatical unit expresses a dialogue act, and (b) not every functionally relevant segment forms a grammatical unit, since such segments are not always grammatically well-formed, may stretch over more than one turn, are not always contiguous, and may contain parts contributed by different speakers, as illustrated below.

In order to deal with these phenomena, the notion of a *functional segment* has been introduced (Geertzen et al., 2007), which is defined as follows:

(9) *A functional segment is a minimal stretch of communicative behaviour that has a communicative function (and possibly more than one); it is minimal in the sense of not having parts that are irrelevant for expressing its communicative function(s).*

Example (10) shows that a functional segment may be discontinuous:

[1] Confusingly, in the literature the term 'utterance' is sometimes used to refer to everything said within a single turn, and sometimes to refer to smaller segments that have a communicative function.

(10) A: Do you know what time the next train leaves?
B: The next train is ... *let me see*... at 7.48.

The discontinuous stretch "*The next train is […] at 7.48"* is a functional segment with the communicative function *Answer*; due to the minimality condition in (9), this functional segment does not include the stretch "... *let me see*...," which is itself a functional segment with the communicative function *Stalling*.

Example (11) shows that a functional segment may stretch over more than one turn:

(11) 1. A: and what departure times do you have on Saturday?
2. B:on Saturday I have a Lufthansa flight in the morning leaving at 08:15,
3. A: yes,
4. B: and a KLM flight at 08:50,
5. A: yes,
6. B: and a Garoeda flight at 10:30.

The utterances 2, 4, and 6 in (11) together constitute a multi-turn answer to the question in utterance 1. The intervening "*yes"* utterances are functional segments with two communicative functions: they provide positive feedback, indicating that the preceding utterance was well understood, and the give the speaking turn back to the previous speaker.

Example (12) shows that a functional segment may contain parts contributed by different speakers:

(12) A: and then, *what is* the, ehm, *the branch office admin*, ehm, ...
B: *booking code* - the admin booking code for this office is 14 2600

In (12), participant A is struggling to formulate a question; B jumps in to help and completes the question (and subsequently answers it). The functional segment expressing A's question is the discontinuous stretch "*what is (..) the branch office admin (..) booking code"*, of which B contributed the last two words.

Example (13) shows that a functional segment is not always grammatically well-formed:

(13) A: what time did he say the meeting will resume?
B: uhm, two-thirty I think

In (13), B's contribution "*two-thirty I think"* is not a well-formed sentence, clause, phrase, or other grammatical unit; yet it is clearly a functional segment, expressing an answer to A's question. The utterance-initial "*uhm"* is a separate functional segment with the communicative functions of taking the turn and stalling for time.

The use of functional segments, as defined in (9), solves many problems in the segmentation of dialogue into meaningful units, taking into account the occurrence of interruptions, of multiple speakers, of overlapping talk, and of simultaneous verbal and nonverbal dialogue acts.

19.3.3 Multifunctionality in utterances

As noted, classical speech act theory assumes that speakers perform one speech act with each utterance. In a critical examination of this view, Allwood (1992) distinguished two

cases where a speaker performs multiple speech acts, *sequential* and *simultaneous multifunctionality,* illustrated both in dialogue fragment (8) above, repeated here:

(8) A: Yes! Come tomorrow. Go to the church. Bill will be there. OK?
B: The church, OK.

The sequence of five dialogue acts performed by A illustrates phenomenon of sequential multifunctionality; moreover, A's contribution illustrates simultaneous multifunctionality in that *Bill will be there* can be taken to be both a statement and a promise.

Sequential multifunctionality disappears when sufficiently small units of segmentation are used; this evidently happens when we segment A's turn in (8) into five segments. However, the size of segments is not what really matters. Bunt (2010) notes that multifunctionality also occurs in more complex forms, and additionally distinguishes *overlapping, discontinuous* and *interleaved* forms of multifunctionality. The latter is illustrated in (14):

(14) I think twenty five euros for a remote... *is that locally something like fifteen pounds*... is too much money to buy an extra remote or a replacement one ...*or is it even more?*

This stretch of speech cannot be cut up into a sequence of functionally meaningful contiguous sub-utterances, since the part *"I think twenty five euros for a remote"* does not express a dialogue act, and neither does *"is too much money to buy an extra remote or a replacement one".* By contrast, using functional segments works perfectly well; it segments (14) into the discontinuous segments *"I think twenty five euros for a remote* (...) *is too much money to buy an extra remote or a replacement one"*} and *"is that locally something like fifteen pounds (...) or is it even more?"*.

Similarly, the occurrence of nested discontinuities in segments corresponding to dialogue acts can be handled well by means of functional segments, as example (15) shows.[2] Four functional segments can be identified here: (1) the turn-initial *"and",* which is a *Turn Keeping* act, indicating that the speaker wants to keep the speaker role; (2) the discontinuous segment *"and so* (...) *we started from an empty lot",* expressing an *Inform* act; (2) the embedded discontinuous segment *"we started in* (...) *we started from*"}, which expresses a *Self-Correction*; and (3) the embedded segment *"uh,...,"* which forms a *Stalling* as well as a *Turn Keeping* act.

(15) A: ***and*** **so** [we started in [*uh,...,*] **we started from**] **an empty lot**

The use of functional segments has important advantages for accurately analyzing dialogue in terms of its functional units, doing justice to the many forms of multifunctionality. Functional segments by definition have no sequential, interleaved, overlapping, or discontinuous multifunctionality, but they do allow simultaneous multifunctionality: a single functional segment may express more than one dialogue act.

[2] From the Switchboard corpus (http://stripe.colorado.edu/~jurafsky/manual.august.html).

19.3.3 Dialogue act interpretation

Two important differences between dialogue acts and the acts of traditional speech act theory is that dialogue acts have a more articulate internal structure, described below, and a computational interpretation in terms of information state updates. The details of such a semantics depend on the precise definition of information states. Poesio and Traum (1998) describe an axiomatic semantics for dialogue acts using an enriched form of discourse representation structures (Kamp and Reyle, 1993) to model information states. Other proposals include the use of Constructive Type Theory (CTT, see Ahn, 2001) of Type Theory with Records (RTT, see Cooper, 1998; 2000), of minimal partial models (Bunt, 2000), and of typed feature structures (Petukhova, 2011).

The definition of DiAML, the Dialogue Act Markup Language which is part of the ISO annotation standard 24617-2, includes a semantics for dialogue acts that makes no assumptions about the representation formalism used in information state modelling. The only assumption of this semantics is that an information state has a part called the 'dialogue history', which records the contributions that have been made to the dialogue, and a part called the 'pending context' where update information is buffered that needs to be checked for consistency before being added to the consolidated part of the information state (Bunt, 2011; 2014). The semantics exploits the hierarchical structure of taxonomies of dialogue act types defined in annotation schemes such as DIT^{++} and the ISO 24617-2 scheme by defining the update operations of dialogue acts as joins of certain elementary update operations. The following example illustrates this.

The semantics of the communicative function *Propositional Question* is defined as (16), where V_a is the evaluation function for dialogue act annotation structures in DiAML:

(16) V_a*(Propositional Question)* = $\lambda X.\, \lambda Y.\, \lambda Y_i.\, \lambda p.\, U_{10}(X,Y,Y_p, p) \cup U_{11}(X,Y,Y_p, p)$

The variables in the right-hand side have the following significance; X and Y are the sender and addressee of the dialogue act; Y_p is the addressee's pending context; and *p* is the propositional content of the question. The elementary update operations U_{10} and U_{11} are defined as follows:

(17) $U_{10}(X,Y,Y_i, p)$: add to Y's pending context the information that participant X wants to know whether p

$U_{11}(X,Y,Y_i, p)$: add to Y's pending context the information that participant X assumes participant Y to know whether p

Applied to two participants and a proposition, the join (16) is a function that updates the addressee's pending context.

This approach to dialogue act semantics accounts for inferences among dialogue acts, such as a *Confirm* act entailing an *Answer* act, because the interpretation of a confirmation is an update operation which differs from that of an answer in that it causes the additional update saying that the sender believes the addressee *expected* the information supplied in the confirmation. Similarly, a *Threat* can be shown to entail an *Inform* act, and the acceptance of an offer to entail a request. See Bunt (2014) for details.

109.3.4 Dialogue Acts and Grounding

Stalnaker (1978) introduced the notion *"common ground"* as *"what the participants in a conversation treat as their common or mutual knowledge. Two people's common ground is the sum of their mutual, common, or joint knowledge, beliefs, and suppositions".* The participants in a conversation must have a certain common ground at the start of the interaction, in order to be able to understand each other and to contribute utterances that can be understood by others. Clark (1996) situates the notion of common ground squarely within pragmatics, when he says: *"Common ground is important to any account of language that appeals to 'context'"*.

Grounding in dialogue is the phenomenon that the participants in a conversation update their common ground, in particular adding elements to the perceived common ground.[3] In Clark and Schaefer's classical model of grounding (Clark and Schaefer, 1989), participants in a dialogue try to establish for each utterance the mutual belief that the addressees have understood what the speaker meant. This is accomplished by the collective actions of the current contributor and the partner, in units called '*contributions'*. Contributions are divided into an acceptance and a presentation phase, giving every contribution the role of accepting the previous contribution, except for those that express negative evidence. Computational studies based on this model includes its extension to human-computer interaction by Brennan and collaborators (Brennan, 1998; Cahn and Brennan, 1999), and the formal theory of grounding by Paek and Horvitz (2000).

Traum (1994) provides a computational model of how conversants reach a state of mutual understanding of what was intended by the speaker of an utterance. His model relies on the distinction of so-called *'grounding acts'*, and the use of protocols which can determine, for any sequence of grounding acts, whether the content expressed by the utterances comprising the acts is grounded. Matheson et al. (2000) use elements of Traum's model in their treatment of grounding from the Information State Update perspective. They represent grounded and ungrounded discourse units in the information state, and change their status from ungrounded to grounded through grounding acts. *Acknowledgement* is the only type of grounding act that is implemented; its main effect is to merge the information in the acknowledged discourse unit into the grounded information. The model keeps only the last two utterances in the information state, so it is not quite clear what would happen if the utterance to be grounded occurs more than two utterances back.

Bunt et al. (2007) describe a computational model of grounding based on the DIT theory of dialogue acts. In this theory, whenever a speaker performs a dialogue act intending to make the addressee aware that a certain proposition *p* holds, in the absence of information to the contrary the speaker expects the addressee to become aware that *p.* Moreover, the addressee knows that speakers operate this way. And the speaker knows this. And addressees know that speakers have this knowledge. And so on, and so on. This infinite iteration can be summed up in a finite way as '*speaker and addressee mutually believe that the speaker expects that the addressee becomes aware that p'.*

[3] The analysis of the relation between symbols and the entities that they refer to involves the notion of 'symbol grounding', which is not directly related to 'grounding' in the sense of common ground construction.

On this approach, the performance of a single dialogue act creates an element in their common ground, albeit a belief of an uncertain nature, an 'expectation', which would need to be strengthened to become a firm belief of the kind that is usually thought to form the common ground. It is argued on empirical grounds that such a strengthening tends to occur after two rounds of positive feedback, making use of the *'Feedback Chaining Principle'*:

(18) *if you receive positive feedback on your last contribution to the dialogue, then that is evidence for you that the current speaker believes that you successfully processed his preceding contribution*.

The following dialogue fragment illustrates this principle.

(19) 1. U: Where should I insert the paper?
2. S: In the feeder.
3. U: Should I put it in the bottom front tray?
4. S: No, in the open tray on top.
5. U: OK thanks.

In utterance 5, participant U gives positive feedback on S's utterance 4, indicating that S's answer was well understood and was a useful answer to the question in 3. From this, S may infer that he correctly understood that question.

In other words, positive feedback on the last contribution implies positive feedback on the contribution before that. Bunt et al. (2007) show that this model is backed up by evidence from dialogue corpora, and that it can be effectively implemented in the information-state update approach.

19.4 Relations in dialogue and discourse

Some dialogue acts are inherently dependent for their meaning on one or more dialogue acts that occurred earlier in the dialogue. This is for example the case for answers, whose meaning is partly determined by the question that is responded to, and for the acceptance or rejection of offers, suggestions, requests, and apologies.

The following example illustrates this, where the meaning of (20.1) clearly depends on whether it is a response to the question (20.2) or to the question (20.3).

(20) 1. A: I'm expecting Jan, Alex, Claudia, and David, and maybe Olga and Andrei.
2. B: Do you know who's coming tonight?
3. B: Which of the people from the project will be there tonight?

As an answer to (20.2), it says that no other people are expected to come than the ones that are mentioned, but as an answer to (20.3) it leaves open the possibility that others will come, who are not 'from the project'.

For dialogue acts which have such a dependence on other dialogue acts, due to their responsive nature, the marking up of the links to their 'antecedent' dialogue acts allows the annotation to express not just that an utterance is an answer, for example, but also to which question it is an answer. This type of relation is called a *functional dependence relation*.

Feedback acts also relate to what happened earlier in the dialogue, but in a different way; they signal something about the processing of what was said before - such as its perception or its interpretation. The following example illustrates this.

(21) 1. A: Is this flight also available on Thursday?
2. B: On Thursday you said?

With his utterance, B checks whether he heard correctly what A said. This is a response to A's *utterance*, rather than to the dialogue act that the utterance expresses. This type of dependence relation is called a *feedback dependence* relation.

Note that positive feedback, signalling correct hearing, understanding, or agreement with what was said, is often expressed nonverbally, for instance by nodding, or in a multimodal way by head movements in combination with vocal backchannels like "*mm*". Nodding can also be used to perform responsive dialogue acts with a positive semantic content, such as a positive answer to a propositional question, or an acceptance of an offer. These are some examples of the interpretation of nonlinguistic communicative behaviour in terms of dialogue acts.

Dialogue acts may also be semantically related through other relations, as in (22)[4]:

(22) 1. A: it ties you on in terms of the technology and the complexity that you want
2. A: like for example voice recognition
3. A: because you might need to power a microphone and other things

Of the three dialogue acts that are contributed sequentially by the same speaker, the one expressed in (22.2) is related to the one in (22.1) through an *Exemplification* relation, and (22.3) is related to the one in (21.1) through an *Explanation* relation. Such relations, known alternatively as '*rhetorical relations*', '*coherence relations*', or '*discourse relations*', have been studied most as relations between units of written text. A wide diversity of taxonomies and classifications of such relations have been proposed and discussed in the literature; e.g. by Mann and Thompson in the framework of Rhetorical Structure Theory (RST, Mann and Thompson 1988); Hovy and Maier (1993); Sanders et al. (1992); PDTB {Prasad et al., 2008). These studies are all focused on discourse relations in written discourse; Petukhova et al. (2011), Tonelli et al. (2010) and Lascarides and Asher (2007) study such relations in dialogue, where they may occur either between the events or propositions that form the semantic contents of dialogue acts, or between the dialogue act themselves, as illustrated in (23) and (24), respectively.

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

(24) 1. A: How much would people be willing to pay for a remote, max?
2. A: I'm afraid we tend to forget that when we consider all these features.

In (23) the semantic content of the dialogue act contributed by B is rhetorically related through a *Cause* relation to that of A's contribution; in (24), by contrast, the dialogue act expressed in the second utterance provides a *Motivation* for asking the question expressed in the first utterance.

[4] From the AMI corpus.

The three types of relations considered here, functional dependence relations, feedback dependence relations, and rhetorical relations, all contribute to the meanings of the related dialogue acts, in particular of the 'dependent' acts. They are therefore taken into account in the theory of dialogue acts and in the more advanced dialogue act annotation schemes, such as DIT^{++} and ISO 24717-2.

19.5 Computational resources for discourse and dialogue pragmatics

19.5.1 Corpora, standards, and other tools

Large-scale annotated corpora, especially those with with (morpho-)syntactic, semantic, and pragmatic annotations are the basis of modern research in linguistics, both in developing empirically valid theories and in empirical testing of linguistic theories. And not only that: such corpora are also the basis for developing modules in natural language processing systems, such as syntactic and semantic parsers, through the application of machine learning techniques, as well as for evaluating such modules by means of quantitative measurements of their performance.

Large-scale annotated corpora of written, spoken, and multimodal discourse in which *pragmatic* phenomena have been marked up, are still scarce, however. Typically, they either contain annotations of communicative functions (in transcribed spoken dialogue) or annotations of coherence relations in written text. The scarcity of pragmatic resources is partly due to the lack of sufficiently well-developed and generally accepted theoretical accounts of pragmatic phenomena, that would be needed as a basis for widely accepted annotation schemes.

Generally speaking, the creation and use of large annotated corpora depend on the availability of two kinds of computational and methodological resources:

- well-founded and widely accepted annotation schemes, ideally in the form of annotation standards, including comprehensive sets of carefully defined annotation concepts and interoperable representation formats;
- software tools to support the use of such schemes in the creation of consistently annotated corpora and in their querying.

Existing resources of these kinds are discussed in the rest of this section.

19.5.2 Annotation schemes

Of the pragmatic phenomena that have received attention in the construction of annotated corpora, the use of language to express dialogue acts in spoken interaction stands out. In the 1980s and '90s a variety of dialogue act annotation schemes was developed, including those of the TRAINS project in the US (Allen et al., 1994), the Map Task studies in the UK (Carletta et al. 1996), and the Verbmobil project in Germany (Alexandersson et al, 1998). These schemes were all designed for a particular purpose and a specific application domain; they made use of different but overlapping sets of dialogue act types, and used often mutually inconsistent terminologies. In the 1990s a group of researchers gathered in the 'Discourse Research Initiative', and drafted a general-purpose schema for multidimensional dialogue act annotation called DAMSL:

Dialogue Act Markup using Several Layers (Allen and Core, 1997; Core and Allen, 1997). With its focus on multidimensionality and domain-independence, DAMSL represented an important step forward compared to earlier dialogue annotation schemes, even though its design was left in an unfinished state. Several variations and extensions of DAMSL have been designed for specific annotation task, such as COCONUT (Di Eugenio et al., 1998) and Switchboard-DAMSL (Jurafsky, et al. 1997).

While more advanced than most other schemes at the time, DAMSL and its derivatives have serious shortcomings. The communicative functions in DAMSL lack precise definitions; the choice of its dimensions is not well-motivated; and its inventory of communicative functions is incomplete. Although intended to support multidimensional annotation, DAMSL is not based on an analysis of multidimensionality; the term 'dimension' is used informally to denote a cluster of intuitively similar functions. The design of the DIT^{++} annotation scheme is based on the observation that participants in a dialogue do not just act in order to achieve a certain goal or perform a certain activity, but also provide and elicit feedback; pause and stall for time in order to avoid unexplained silences; take turns; help each other in expressing themselves; correct themselves and each other when detecting a speech error, and so on. A dimension is defined in DIT^{++} as follows:

(25) *A dimension is a class of dialogue acts that are concerned with a particular aspect of communication, corresponding to a particular category of semantic content.*

Aspects of communication include advancing the task or activity that motivates the dialogue; monitoring attention and understanding; taking turns; managing the use of time; editing one's own speech or that of another speaker; opening and closing topics; and dealing with social obligations like thanking and apologising. The categories of semantic content that correspond with these activities are task-related information; the success of processing previous utterances; the allocation of the speaker role; the time requirements of contributing to the dialogue; speech disfluencies; topic progression; and social obligations in conversation.

Using this notion of dimension, the DIT^{++} taxonomy of was developed by establishing criteria for distinguishing dimensions and communicative functions, and incorporating communicative functions defined in various schemes (including DAMSL and its derivatives, AMI, DIT, ICSI-MRDA and Vermobil) into a single comprehensive scheme with precise definitions. This scheme served as the basis for the ISO 24617-2 standard for dialogue act annotation, developed in a collaborative effort involving an international team of experts (see Bunt et al., 2010). The ISO 24617-2 standard includes the definition of the Dialogue Act Markup Language DiAML, with the information-state update semantics mentioned above (see Bunt, 2009).

The research involved in constructing the Penn Discourse Treebank (Prasad et al., 2008), where discourse relations have been added to the syntactic trees in the Penn Treebank has recently been taken as the starting point for an ISO effort aiming to establish a standard for the annotation of spoken and written discourses with coherence relations; see Bunt, Prasad and Joshi, 2012). This effort combines forces with a new European initiative (TextLink, see http://textlinkcost.wix.com/textlink) to define an annotation scheme for discourse relations applicable to all the languages of Europe, including languages of immigrant groups such as Arabic and Chinese.

19.5.3 Software tools

Software tools that have been developed for supporting the annotation of dialogue data include DialogueView (Heeman et al., 2002), ACT (Yang et al., 2002), the GATE tools (Cunningham, 2002), the NITE XML toolkit (Carletta et al., 2009), ANVIL (Kipp, 2001), ELAN (Wittenburg et al., 2006), and MMAX2 (Stede and Heintze, 2004).

The ANVIL tool, for example, offers a graphical user interface for creating annotation elements on as many user-defined tiers as desired for a particular purpose or annotation scheme. Tiered representations (also used in some other tools) are convenient for annotating multimodal dialogue, allowing the use of different tiers for different modalities (e.g. one tier for speech, one for gaze direction, one for body posture, one for facial expression,…), and also convenient for multidimensional annotation, using different tiers for different dimensions. ANVIL allows different tag sets to be imported, as well as annotations to be exported in a variety of formats including the DiAML format of the ISO 24617-2 standard (Kipp et al., 2012).

For the annotation of discourse relations, RSTTool (O'Donnell, 2000) is a popular tool for the segmentation of written text and the construction of tree structures according to Rhetorical Structure Theory. The Conano tool (Stede and Heintze, 2004) supports PDTB-style annotation of local coherence relations, detecting words that might function as discourse connectives and using syntactic information to guess the arguments of the relation.

Software tools for corpus searching and querying tend to be corpus-specific, applicable only to treebanks, for example. Moreover, most search engines are sentence-based, which makes them less useful for discourse phenomena that stretch over multiple utterances. The generalization of such tools in order to be able to handle discourse-level patterns across corpora and theoretical frameworks is a challenge task for specialists in computational pragmatics.

Another kind of software tool to support pragmatic studies is exemplified by the machine-learning based automatic dialogue act annotator developed by Petukhova (2011). This annotator has been applied successfully in the multidimensional annotation of spoken dialogues with DIT^{++} or ISO 24617-2 tags. The program performs both the identification of functional segments in raw speech and the assignment of communicative functions to each functional segment. Applied to data from the AMI corpus and the HCRC Map Task, accuracies have been achieved of up to 96%, indicating that it may be an excellent basis for semi-automatic annotation.

19.5.3 Annotated corpora

As mentioned above, existing pragmatically annotated corpora are concerned mainly with discourse relations in written text or dialogue act occurrences in spoken dialogue.

The Penn Discourse Treebank (PDTB) is a useful, fairly large corpus with discourse relation annotations added to the Penn Treebank corpus of syntactically annotated English texts from the Wall Street Journal. The annotation of discourse relations follows a lexically-grounded approach (see Prasad et al., 2008) and aims to be theory-neutral with respect to the nature of higher-level representation of discourse structure (as in

RST, for example), in order to allow the corpus to be usable within different theoretical frameworks. Corpora with (manually constructed) annotations for the occurrence of discourse relations exist for a range of languages including Chinese, Czech, Danish, Dutch, French, German, Hindi and Turkish, but these are all of modest size. The Potsdamer Commentary Corpus (version 2.0; see Stede and Neumann 2014) is a collection of 175 newspaper commentaries, annotated with nominal coreference relations, discourse connectives (similar to those in the PDTB), and rhetorical structures according to RST.

Useful dialogue corpora, annotated with dialogue act information, include the Switchboard corpus (Jurafsky et al., 1997), the ICSI-MRDA corpus (Dhillon et al., 2004); the AMI corpus (Ashby et al., 2005) and the related AMIDA corpus[5], and the HCRC Map Task corpus (Carletta et al., 1996). Unfortunately, the annotations in each of these corpora have been made using corpus-specific annotation schemes and frameworks, which make them hard to use in other theoretical frameworks. The creation of large-scale corpora with dialogue act annotation according to the ISO 24617-2 standard remains an important goal for the near future; initial steps in this direction have been made by Fang et al. (2011) in studies of the possibility to semi-automatically convert the annotations in the Switchboard corpus (which uses the SWBD-DAMSL variant of the DAMSL scheme) to ISO-24617-2 annotations.

19.6 Conclusions and Perspectives

In summary, computational pragmatics offers in the first place a number of tools and resources in support of research in pragmatics. The use of corpora, annotated for pragmatic phenomena, is especially beneficial for the coverage and empirical validity of accounts of these phenomena. Large annotated corpora are useful not only for linguistic research but also for building components of language processing systems through the application of machine learning techniques. The construction of such corpora depends crucially on widely agreed annotation schemes; efforts like the establishment of ISO annotation standards are therefore important.

Other computational resources, such as automatic reasoning programs like the abductive prover of the mini-Tacitus system, are interesting for pragmatic research since they allow to investigate in detail how inference processes that combine linguistic information with world knowledge can explain pragmatic phenomena such as conversational implicatures, and context-based interpretation more generally.

The fundamental challenge of pragmatics is to understand how language interacts with context and how inference interacts with the interpretation and generation of language. Computational pragmatics makes important contributions to the efforts to meet this challenge by providing computational models of interpretation, generation, inferencing and learning. Something which is still missing, however, is the construction and use of powerful context models. Much of the work that takes context information into account considers only the *linguistic* context, i.e. the preceding discourse. This is the only kind of context information that is available in corpora, and therefore for applying machine

[5] http://corpus.amidaproject.org.

learning techniques. This means that only a fraction of the relevant context information is taken into consideration. Richer context models have been introduced for describing the semantics of dialogue acts, and should also be considered for use in annotating dialogue and discourse corpora. Manual addition of this information to corpus annotations hardly seems feasible, in view of its complexity, therefore new computational methods will have to be developed to make such information available in annotations.

Bibliography

Ahn, René (2001) Agents, objects, and events: a computational approach to knowledge, observation, and communication. PhD Thesis, Eindhoven University of Technology.

Alexandersson, Jan, Bianka Buschbeck, Tsutomu Fujinami, Michael Kipp, Stephan Koch, Elisabeth Maier, Norbert Reithinger, Birte Schmitz and Melanie Siegel (1998) Dialogue acts in Verbmobil-2, Second edition. Verbmobil Report 226, DFKI, Saarbrücken.

Allen, James, Lenhart Schubert, George Ferguson, Peter Heeman, Chung Hee Hwang, Tsuneaki Kato, Marc Light, Nathaniel Martin, Bradford Miller, Massimo Poesio and David Traum (1994) The TRAINS Project: A case study in building a conversational planning agent. TRAINS Technical Note 94-3, Department of Computer Science, University of Rochester.

Allen, James and Mark Core (1997) Draft of DAMSL: Dialogue Act Markup in Several Layers. DRI: Discourse Research Initiative, University of Pennsylvania, Philadelphia.

Allwood, Jens (1992) On dialogue cohesion. *Gothenburg papers in Theoretical Linguistics* 65, Göteborgs Universitet, Department of Linguistics.

Ashby, Simone, Sebastien Bourban, Jean Carletta, Mike Flynn, Mael Guillemot, Thomas Hain, Jaroslav Kadlec, Vasilis Karaiskos, Wessel Kraaij, Melissa Kronenthal, Guillaume Lathoud, Mike Lincoln, Agnes Lisowska, Iain McCowan, Wilfried Post, Dennis Reidsma, and Pierre Wellner (2005) The AMI Meeting Corpus: A pre-announcement. In S. Renals and S. Bengio (eds.) *MLMI'05: Proceedings of the Workshop on Machine Learning for Multimodal Interaction.* Lecture Notes in Computer Science 3869. Berlin: Springer, pp. 28-39.

Austin, John L. (1962) *How to do things with words*. Clarendon Press, Oxford, UK.

Beun, Robbert-Jan 1989] *The recognition of declarative questions in information dialogues*. Ph.D. Thesis, Tilburg University.

Brennan, Susan (1998) The grounding problem in conversation with and through computers. In S. R. Fussell and R. J. Kreuz (Eds.), Social and cognitive psychological approaches to interpersonal communication. Hillsdale, NJ: Lawrence Erlbaum, pp. 201-225.

Bunt, Harry (1969) Conversational principles in question-answer dialogues. In Dieter Krallmann (ed.) *Zur Theorie der Frage*. Essen: Narr Verlag, pp. 119-141.

Bunt, Harry (1989) Information dialogues as communicative action in relation to partner modelling and information processing. In Martin Taylor, Françoise Néel and Don Bouwhuis (eds.) *The structure of multimodal dialogue*, Amsterdam: North-Holland, pp. 47-74.

Bunt, Harry (1995) Dynamic Interpretation and Dialogue Theory. in Martin Taylor, Don Bouwhuis and Françoise Néel (eds.) *The Structure of Multimodal Dialogue, Vol. 2*. Amsterdam: Benjamins, pp. 139-166.

Bunt, Harry (2000) Dialogue pragmatics and context specification. In Harry Bunt and William Black (eds.) *Abduction, Belief and Context in Dialogue; Studies in Computational Pragmatics*. Amsterdam: Benjamins, pp. 81-105

Bunt, Harry (2009) The DIT^{++} taxonomy for functional dialogue markup. In Dirk Heylen, Catherine Pelachaud, Roberta Catizone and David Traum (eds.) *Proceedings AAMAS 2009 Workshop 'Towards a Standard Markup Language for Embodied Dialogue Acts*, Budapest, pp. 13-23.

Bunt, Harry (2010). Multifunctionality in Dialogue. *Computers, Speech, and Language*, 22, 224-245.

Bunt, Harry (2011). The semantics of dialogue acts. *Proceedings 9th International Conference on Computational Semantics (IWCS 2011),* pp. 1-13.

Bunt, Harry, Jan Alexandersson, Jean Carletta, Joe.-Woong Chae, Alex Fang, Koiti Hasida, Kiyong Lee, Volha Petukhova, Andrei Popescu-Belis, Laurent Romary, Claudia Soria, and David Traum (2010). Towards an ISO standard for dialogue act annotation. In *Proceedings 7th International Conference on Language Resources and Systems (LREC 2010)*, Malta. Paris: ELRA.

Bunt, Harry and William Black (2000) The ABC of Computational Pragmatics. In Harry Bunt and William Black (eds.) *Abduction, Belief and Context in Dialogue. Studies in Computational Pragmatics*. Amsterdam: Benjamins, pp. 1-46.

Bunt, Harry, Simon Keizer and Roser Morante (2007) A computational model of grounding in dialogue. *Proceedings 8th SIGdial Workshop in Discourse and Dialogue*, Antwerp, pp. 591-598.

Bunt, Harry, Michael Kipp, and Volha Petukhova (2012) Using DiAML and ANVIL for multimodal dialogue annotation. In *Proceedings 8th International Conference on Language Resources and Systems (LREC 2012)*, Istanbul.

Bunt, Harry, Rashmi Prasad, and Aravind Joshi (2012) First steps toward an ISO standard for the annotation of discourse relations. *Procceedings 7th Joint ISO – ACL SIGSEM Workshop on Interoperable Semantic Annotation (ISA-7),* Istanbul, pp. 60-69.

Cahn, Janet, and Susan Brennan (1999). A psychological model of grounding and repair in dialog. *Proceedings, AAAI Fall Symposium on Psychological Models of Communication in Collaborative Systems*. North Falmouth, MA: American Association for Artificial Intelligence, pp. 25-33.

Carletta, Jean, Amy Isard, Stephen Isard, John Kowtko and Gwyneth Doherty-Sneddon (1996) HCRC dialogue structure coding manual. Technical Report HCRC/TR-82.

Carletta, Jean, Stefan Evert, Jonathan Kilgour, Craig Nicol, Dennis Reidsma, Judy Robertson, and Holger Voormann (2009) Documentation for the NITE XML Toolkit, version 0.3. School of Informatics, University of Edinburgh.

Clark, Herbert (1996) *Using language*. Cambridge University Press, Cambridge, UK.

Clark, Herbert and Edward Schaefer (1989) Contributing to discourse, *Cognitive Science* 13, 259-294.

Cooper, Robin (1998) Integrating diverse information resources into dialogue updates. *Proceedings Second International Workshop on Cooperative Multimodal Communication (CMC-98)*, Tilburg, pp. 72-82.

Cooper, Robin (2000) Information States, Attitudes and Dependent Record Types, In Peter Blackburn, Nick Braisby, Lawrence Cavedon, and Alsushi Shimojima (eds.), *Logic, Language and Computation, Vol 3*, Stanford: CSLI Publications, pp. 85--106.

Core, Mark and James Allen (1997) Coding Dialogs with the DAMSL Annotation schema. *Working Notes from the AAAI Fall Symposium on Communicative Action in Humans and Machines*, Cambridge, MA, pp. 28-35.

CISD Research Group (1997) DAMSL Revised Manual. Unpublished document, Computer Science Department, University of Rochester.
Available at http://www.cs.rochester/edu/research/cisd/damsl.

Cunningham, Hamish (2002) GATE: a General Architecture for Text Engineering. *Computers and the Humanities* 36, 223-254.

Fang, Alex, Harry Bunt, Jing Cao, and Xiaoyue. Li (2011) Relating the Semantics of Dialogue Acts to Linguistic Properties: A machine learning perspective through lexical cues. In *Proceedings 5th International Conference on Semantic Computing*, Stanford.

Di Eugenio, Barbara, Pamela W. Jordan and Lina Pylkkänen (1998) The COCONUT project: dialogue annotation manual. ISP Technical Report 98-1, University of Pittsburgh.

Fellbaum, Christiane (1998) *WordNet: An Electronic Lexical Database*. Cambridge, MA: MIT Press.

Geertzen, Jeroen (2009) *Dialogue act prediction and recognition*. Ph.D. Thesis, Tilburg University.

Geertzen, Jeroen, Volha Petukhova and Harry Bunt (2007) A multidimensional approach to utterance segmentation and dialogue act classification. *Proceedings 8th SIGdial Workshop in Discourse and Dialogue*, Antwerp, pp. 140-149.

Heeman, Peter, Fan Yang and Susan Strayer (2002) DialogueView: An Annotation Tool for Dialogue. *Proceedings 3rd SIGdial Workshop on Discourse and Dialogue*, pp. 50-59.

Hobbs, Jerry (1986) Overview of the TACITUS project. *Computational Linguistics* 12 (3).

Hobbs, Jerry (1990) An Integrated Abductive Framework for Discourse Interpretation. *Proceedings AAAI Spring Symposium on Abduction*, Stanford, pp. 10-12.

Hobbs, Jerry, Martin Stickel and Paul Martin (1993) Interpretation as abduction. *Artificial Intelligence* 63, 69-142.

Hovy, Edward and Elisabeth Maier (1993) Parsimonious or profligate: how many and which discourse structure relations? *ISI Research Report*, Information Sciences Institute, University of Southern California, Marina del Rey.

ISO (2012) ISO 24617-2 Language resource management - Semantic annotation framework (SemAF), Part 2: Dialogue acts. International Standard, ISO, Geneva.

Jurafsky, Dan, Elisabeth Schriberg, and Debra Biasca (1997) Switchboard SWBD-DAMSL Shall-Discourse-Function Annotation Coders Manual.
Available at http://stripe.colorado.edu/~jurafsky/manual.august.html.

Kamp, Hans and Uwe Reyle (1993) *From Discourse to Logic*. Dordrecht: Kluwer.

Kipp, Michael (2001) ANVIL - A Generic Annotation Tool for Multimodal Dialogue. In *Proceedings of Eurospeech 2001*, 1367--1370.

Lascarides, Alex and Nicholas Asher (2007) Segmented Discourse Representation Theory: Dynamic Semantics with Discourse Structure. In Harry Bunt and Reinhard Muskens (eds.) *Computing Meaning, Vol. 3*. Dordrecht: Springer, pp. 87-124

Mann, William and Sandra Thompson (1988) *Rhetorical structure theory: toward a functional theory of text organisation*, Cambridge, MA: MIT Press,

Matheson, Colin, Massimo Poesio, and David Traum (2000) Modelling grounding and discourse obligations using update rules. In *Proceedings of the 1st Annual Meeting of the North American Association for Computational Linguistics (NAACL 2000)*.

Muller, Christoph and Michael Strube (2003) Multi-Level annotation of linguistic data with MMAX2. *Proceedings 4th SIGdial Workshop on Discourse and Dialogue*, Sapporo, pp. 198-207.

Nöth, Elmar. Anton Batliner, Volker Warnke, Jürgen Haas, Manuela Boros, Jan Buckow, Richard Huber, Florian Gallwitz, Matthias Nutt, and Heinrich Niemann (2002), On the use of prosody in automatic dialogue understanding. *Speech Communication* 36 (1-2), 45-62.

Ovchinnikova, Ekaterina, Nicofar Montazeri, Theodore Alexandrov, Jerry Hobbs, Michcael McCord, and Rutu Mulkar-Mehta (2011) Abductive Reasoning with a Large Knowledge Base for Discourse Processing. *Proceedings 10th International Conference on Computational Semantics (IWCS 2011),* Oxford, pp. 225-234.

Paek, Tim and Ed Horvitz (1999) Uncertainty, utility, and misunderstanding: A decision-theoretic perspective on grounding in conversational systems. *AAAI Fall Symposium on Psychological Models of Communication in Collaborative Systems*, Cape Cod, MA.

Petukhova, Volha (2011) *Multidimensional Dialogue Modelling*. PhD Thesis, Tilburg University.

Petukhova, Volha, Laurent Prévot, and Harry Bunt (2011) Multilevel discourse relations between dialogue units. *Proceedings ISA-6, 6th Joint ISO – ACL SIGSEM Workshop on Interoperable Semantic Annotation*, Oxford, pp. 18-27.

Petukhova, Volha and Harry Bunt (2011) Incremental dialogue act understanding. *Proceedings 9th International Conference on Computational Semantics (IWCS 2011),* Oxford, pp. 235-244.

Poesio, Massimo and David Traum (1998) Towards an axiomatization of dialogue acts. *Proceedings Twente Workshop on the Formal Semantics and Pragmatics of Dialogue*, Enschede, pp. 207-222.

Prasad, Rashmi, Nikhil Dinesh, Alan Lee, Eleni Miltsakaki, Livio Robaldo, Aravind Joshi, and Bonnie Webber (2008) The Penn Discourse TreeBank 2.0. In *Proceedings 6th International Conference on Language Resources and Systems (LREC 2008),* Marrakech.

Ruppenhofer, Josef, Michael Ellwsorth, Miriam Petruck, Christopher Johnson and Jan Scheffczyk (2006) *FrameNet II: Extended Theory and Practice*. International Computer Science Institute, Berkeley, CA.

Sacks, Harvey, Emmanuel Schegloff and Gail Jefferson (1974) A simplest systematics for the organization of turn-taking for conversation. *Language* 53:696--735.

Sanders, Ted, Wilbert Spooren, and Leo Noordman, Toward a taxonomy of coherence relations, *Discourse Processes* 15, 1-35.

Searle, John (1969) *Speech Acts*. Cambridge, UK: Cambridge University Press.

Shannon, Claude (1948) *A mathematical theory of communication*. Chicago: University of Illinois Press.

Shriberg, Elisabeth, Rebecca Bates, Andreas Stolcke, Paul Taylor, Dan Jurafsky, Klaus Ries, Noah Coccaro, Rachel Martin, Marie Meteer, and Carol van Ess-Dykema, Can prosody aid the automatic classification of dialog acts in conversational speech? *Language and Speech* 41 (3-4), 439-487.

Shriberg, E.lisabeth, Raj Dhillon, Sonali Bhagat, Jeremy Ang, and Hannah Carvey (2004) The ICSI meeting recorder dialog act (MRDA) corpus. In *Proceedings 6th SIGdial Workshop in Discourse and Dialogue*, Boston, pp. 97-100.

Stalnaker, Robert (2002) Common Ground, *Linguistics and Philosophy* 25, 701-721.

Stede, Manfred and Silvan Heintze (2004) Machine-assisted rhetorical structure annotation. In *Proceedings 20th Internationl Conference on Computational Linguistics (COLING 2001)*, pp. 425-431, Geneva.

Stede, Manfred and Arne Neumann (2014) Potsdam Commentary Corpus 2.0: Annotation for Discourse Research. *Proceedings 9th International Conference of Linguistic Resources and Evaluation (LREC 2014)*. Reykjavik, pp. 925-929.

Stickel, Martin (1988) A Prolog-like inference system for computing minimum-cost abductive explanations in natural language interpretation. Technical Note 451, SRI International, Menlo Park.

Stolcke, Andreas, Klaus Ries, Noah Coccaro, Elisabeth Shriberg, Rebecca Bates, Dan Jurafsky, Paul Taylor, Rachel Martin, Marie Meteer and Carol van Ess-Dykema (2000) Dialogue act modeling for automatic tagging and recognition of conversational speech, *Computational Linguistics* 26 (3), 339-373.

Tonelli, Sara, Giuseppe Riccardi, Rashmi Prasad, and Aravind Joshi (2010) Annotation of Discourse Relations for Conversational Spoken Dialogs. In *Proceedings 7th International Conference of Linguistic Resources and Evaluation (LREC 2010)*, Malta.

Traum, David (1994) *A computational theory of grounding in natural language conversation.* Ph.D. Thesis, University of Rochester.

Traum, David and Steffan Larsson (2003) The Information State Approach to Dialogue Act Management. In Jan van Kuppevelt and R. Smith (eds.) *Current and new Directions in Discourse and Dialogue*, Dordrecht: Kluwer.

Wittenburg, Peter, Hennie Brugman, Albert Russel, Alexander Klassmann, and Han Sloetjes (2006) ELAN: A professional framework for multimodality research. In *Proceedings 5th International Conference on Language Resources and Evaluation (LREC 2006),* pp. 1556-1559.

Yang, F., S. Strayer and Peter Heeman (2002) ACT: A Graphical Dialogue Annotation Comparison Tool. In *Proceedings ICSLP 2002: 7th International Conference on Spoken Language Processing (INTERSPEECH 2002),* Denver, Colorado.