The Semantics of dialogue

Jonathan Ginzburg

Cambridge Handbook of Semantics

Maria Aloni and Paul Dekker

This guide was compiled using
cambridge7A.cls 2010/09/09, v2.10

The latest version can be downloaded from:
https://authornet.cambridge.org/information/productionguide/
LaTeX_files/cambridge7A.zip

1

The Semantics of Dialogue

Jonathan Ginzburg

1.1 Introduction

The semantics of dialogue is a fundamental topic for a number of reasons. First, dialogue is the primary medium of language use, phylogenetically and ontogenetically. Second, studying dialogue forces one to a particularly careful study of the nature of context. The context has a role to play in determining *what* one can or should say at a given point and also *how* to say it. Conversely, it affords the interlocutors a very impressive economy of expression—there is much subtlety that can be achieved with relatively little effort drawing simply on material that is *in* the context.

Consequently, two themes will drive this article, relating to two fundamental problems a semantic analysis in dialogue has to tackle:

(1) a. **Conversational relevance**: given that a conversation is in a certain state, what utterances can be produced coherently by each conversational participant in that state?

b. **Conversational meaning**: what conversational states are appropriate for a given word/construction and what import will that word have in such a state?

Conversational Relevance is without doubt a very fundamental and difficult problem. It is closely connected to the Turing test and solving this problem in full generality is, as Turing pointed out, a possible basis for understanding the nature of intelligence (Turing (1950), see also Łupkowski and Wiśniewski (2011)). But is it a *semantic*, as opposed to a *pragmatic* problem or one connected to generalized notions of cognition? We will not deal with this issue of territorial demarcation in anything but passing (see also the article on the semantics/pragmatics in this

volume). Nonetheless, we will offer a detailed empirical and theoretical analysis of the components of conversational relevance.

Conversational meaning is *semantic* enough, but the obvious question one might ask is—why *conversational*? Why should we consider meaning in the context of conversations? One might turn the question on its head and ask, noting, as I already have above, that conversation is the primary linguistic medium in which language evolved and existed for millenia and that it is the setting in which language is acquired: how could we *not* take conversation as the basic setting for semantic theory? I think this perspective is highly attractive—I have offered some evidence for it in Ginzburg (2012) and establishing this point—more generally that *grammar* should be viewed as characterizing talk in interaction—is the main topic of Ginzburg and Poesio (2014). For current purposes, it will suffice to point to the existence of various words and constructions whose import is tied to a conversational setting. These in themselves justify the need for at least some of semantics to be conversationally oriented. In section 1.5 we will also encounter phenomena from dialogue that motivate an incremental view of semantic composition.

The structure of this article is as follows: I start by pointing to work in the philosophy of language, pragmatics, cognitive psychology, and conversational analysis, which provides current formal work on dialogue with key phenomena and concepts. I then point to a number of core phenomena that linguistic theories of dialogue need to account for. I subsequently sketch an account of certain of these phenomena in contemporary dialogue frameworks. Finally, I point to a number of additional phenomena that seem to require significant modifications of our view of semantics, context, and language.

1.2 Antecedents of Formal Dialogue Theory

1.2.1 Language is or could be dialogical

Wittgenstein (1953) introduced the notion of a *language game*. Now the notion of *language* game is no more definite than the notion of a game *simpliciter*, which Wittgenstein put forward as an example of a concept difficult to characterize in hard and fast terms. Still a language game can be viewed as a type of interaction involving language use within a more or less restricted set of associated actions. For dialogue researchers it is important for at least two reasons. First, in discussing the hypo-

thetical 'builder's language' Wittgenstein illustrates that within a specific domain 'non-canonical' utterances, particularly non-sentential ones such as 'Slab!' or 'Beam!', can be as canonical or even more natural than fully spelled out, sentential utterances. Indeed Wittgenstein provides one of the first challenges to the 'sententialist' view of ellipsis that will subsequently, from the 1960s onwards, predominate in generative grammar:

> But what about this: is the call "Slab I" in example (2) a sentence or a word? If a word, surely it has not the same meaning as the like- sounding word of our ordinary language, for in [section 2] it is a call. But if a sentence, it is surely not the elliptical sentence: "Slab" of our language. As far as the first question goes you can call "Slabl" a word and also a sentence; perhaps it could be appropriately called a 'degenerate sentence' (as one speaks of a degenerate hyperbola); in fact it is our 'elliptical' sentence. But that is surely only a shortened form of the sentence "Bring me a slab", and there is no such sentence in example (2). But why should I not on the contrary have called the sentence "Bring me a slab" a lengthening of the sentence "Slab"? Because if you shout "Slab!" you really mean: "Bring me a slab". But how do you do this: how do you mean that while you say "Slabl"? Do you say the unshortened sentence to yourself? ... (Wittgenstein (1953), section 19)

Tying utterance interpretation to facts characteristic of specific domains provides a potential way of dealing with a variety of actually occurring non sentential utterances (NSUs) in various domains:

(2) a. [A advances to bar, addresses barman] A: A Franziskaner and a Duwel.
b. (1) 'Your name?' asked Holmes.
(2) 'Patrick Cairns.'
(3) 'Harpooner?'
(4) 'Yes, sir. (5) Twenty six voyages.'
(6) 'Dundee, I suppose?'
(7) 'Yes, sir.' ('Black Peter', Sir Arthur Conan-Doyle).

More generally, it opens up the way to talk about domain specificity of language. *Variation* is a big, perhaps one of the biggest issues in contemporary sociolinguistics (Tagliamonte (2006)). This view of language lies at present in big contrast to the domain independent view almost universally assumed in formal grammar of most stripes. However, workers on speech recognition assume such a view via the notion of *language model* (Chelba (2010)). Given that one of the main concerns of dialogue is characterizing relevance and that, as demonstrated in (2), this is clearly

domain relative, the importance of a language game perspective is evident.

In fact, there is a long tradition predating Wittgenstein by centuries of logic games or formal dialectic. This tradition, expunged from mainstream attention in post-Fregean logic until recently, is documented in Hamblin (1970) and has been active in argumentation theory (Walton (2011)). This perspective is quite narrower than the Wittgensteinian, with a normative intent, but more directly tractable. Thus, Hamblin's *Why-Because System with Questions* (Hamblin (1970) pp. 265–276 can be viewed as one of the first modern attempts to provide a formal description of two person discussion: moves are characterized in terms of formulas of propositional logic which can affect a set of *commitments*. In addition to moves corresponding to assertion and the posing of alternative questions, the system provides for moves that elicit the justification of an assertion, retraction, and a request to resolve an inconsistency.

1.2.2 Meaning from rationality

From our current perspective, Gricean pragmatics (Grice (1989)) is very much focussed on the issue of Conversational Relevance: Grice's maxims constitute a quasi-calculus for establishing what utterance to select at a given point in a conversation. And, consequently, what inferences to draw in case this selection does not get realized in practice.

The big gaping hole in Grice's account, one which he was fully aware of, was the lack of substance concerning the maxim of Relevance. Two crucial ingredients are missing: an explicit notion of *current conversational state* and a means for generating the range of *potentially relevant* contributions. While there has been much insightful work in mainstream pragmatics since Grice, whether closely following Grice (e.g. Levinson (2000)) or radically changing his vision (e.g. Relevance Theory, Sperber and Wilson (1986))), it has not filled these lacunes.

On the other hand for workers on the semantics of dialogue, Grice has provided inspiration in significant ways. Groenendijk (2006) shows how using a relatively simple extension of first order logic that contains also questions can enable one to simultaneously define notions of Quality, Quantity and Relevance (his technical term is *pertinence*.) Dekker (2006), on the basis of a synthesis of dynamic semantics, Gricean pragmatics, and relevance theory, shows how to characterize the optimality of a discourse. This general strategy is taken a step further in *Inquisi-*

tive Semantics (Groenendijk (2009); Groenendijk and Roelofsen (2009)), discussed in detail in the article on Questions (this volume...)

As we discuss in section 1.4, accounts which emphasize a notion of Questions–Under–Discussion (QUD) as fundamental to context offer detailed explications of relevance, some of which emphasize Gricean rationality and cooperativeness (e.g. early Asher and Lacarides, Roberts), others somewhat more wary (e.g. Ginzburg, later Asher and Lacarides).

1.2.3 Structure from Interaction

A third tradition, Conversational Analysis (CA) (Sacks *et al.* (1974); Schegloff *et al.* (1977); Schegloff (1987)), puts interaction as the primary scene for linguistic use. CA's contributions to developing a theory of dialogue address primarily **move relevance**:

- **adjacency pairs**: CA provides extensive evidence for the existence of strong preference for a certain class of responses (*second part pairs*) as follow ups to a given class of *first part pairs*:

  (3) a. A:Who left; B: Bill (query/reply)
  b. A: Open the window please! B: Sure (command/acceptance)
  c. A: Hi! B: Hiya! (greeting/counter-greeting)

- **Repair**: perhaps an even more important contribution of CA is introducing the notion of *repair*:

  > By 'repair' we refer to efforts to deal with trouble in speaking, hearing, or understanding talk in interaction. 'Trouble' includes such occurrences as misarticulations, malapropisms, use of a "wrong" word, unavailability of a word when needed, failure to hear or to be heard, trouble on the part of the recipient in understanding, incorrect understandings by recipients, and various others. Because anything in talk can be a source of trouble, everything in conversation is, in principle, "repairable". (Schegloff (1987), p. 210)

  Schegloff *et al.* (1977) show that there are many commonalities between *self-repair* (A repairing her own utterance) and *other repair* (B repairing A's utterance.). This work was the first of many to show the regularity of repair, still very much neglected in generative and formal work, possibly under the influence of the competence/performance distinction, which consigns most self-repair to the performance dustbin. Both self-repair and other-repair have subsequently been the object of much study by researchers in other disciplines: as we discuss in sections 1.3.3 and 1.5.1 respectively, self-repair by psychologists and

speech scientists, whereas other-repair by HCI and dialogue system designers and by developmental psychologists.

- **Turn Taking**: a third crucial contribution of CA was to initiate the study of *turn taking* in Sacks *et al.* (1974). A remarkable feature of conversation is that there are relatively few overlaps but at the same time correspondingly few intervals of extended silence. Indeed gaps longer than 600 msec are understood to imply that a dispreferred response will be supplied and gaps longer than 1500 msec are relatively rare. Sacks, Jefferson, and Schegloff proposed a basic principle for determining how a next turn is assigned interactively, as a basis for explaining the efficacy of the turn taking system. This principle has withstood the ravages of time, serving as the basis for detailed psycholinguistic experimentation (see e.g. De Ruiter *et al.* (2006)) and for typological investigations (e.g. Stivers *et al.* (2009)), which provide some evidence for the account's cross-linguistic and cross-cultural validity.

1.3 Core Dialogue Phenomena

In this section I list some phenomena any theory of dialogue needs to account for or underpin other modules' accounts thereof, in line with the driving issues in (1).

1.3.1 'Direct' Relevance

The first such phenomenon, at the level of utterance content, is characterizing what one might call *direct relevance*. That is, the relationship that holds between moves m_1 and m_2 when m_2 constitutes a direct response to m_1. While the CA literature offers some data on this issue when adjacency pairs are discussed, this is an area that has been studied systematically primarily in the domain of question/answer relations (i.e. where m_1 is a query move and m_2 is assertoric.). Here notions such as *partial answerhood* (Groenendijk and Stokhof (1984)), Aboutness (Ginzburg (1995a); Ginzburg and Sag (2000)), and licensing (Groenendijk (2006)).

(4) a. Jo: When is the train leaving?
Carrie: At 2:58, 17.333398 seconds, according to our caesium clock./At 2:58./In about an hour./In a short while.

b. Chris: Did Merle leave?
Kim: Yes./Probably./It's not likely./No.

c. Sandy: Who will help the President?
Tracy: His close friends./Few people we know./Merle Africa or Merle Haggard.
(Ginzburg and Sag (2000)'s example (103))

More recently, Ginzburg (Ginzburg (1996, 2012)) suggested that the appropriate notion for 'direct' responses to queries (his notion of a q–specific utterance, relative to a question q) should encapsulate both answer responses together with query responses. The idea being that whenever an issue is being addressed there can be 'direct responses' of both assertoric and interrogative nature. A similar intuition seems to underlie the notion of *compliance* (Groenendijk and Roelofsen (2009)).

(5) a. A: Bo left. B: Right/Did he?/Isn't he still around?
b. A: Who shall we invite to the party? B: Jack and Jill./Who is available?

This raises in a particular the issue of how to characterize the range of queries that can coherently follow a query. A large proportion of these are clarification requests, discussed in section 1.3.3 and these do not address the query as such. However, results from a corpus study of the BNC (Łupkowski and Ginzburg (2014)), which offers a comprehensive characterization of such responses, show the existence of a number of other classes, including the following three:[a]

(6) a. **Dependent questions**: A: Do you want me to <*pause*> push it round?
B: Is it really disturbing you? [BNC, FM1, 679–680]
(cf. *Whether I want you to push it depends on whether it really disturbs you.)*
b. **Situationally pertinent questions**: A: Well do you wanna go down and have a look at that now?
<*pause*> While there's workmen there?
B: Why haven't they finished yet? [BNC, KCF, 617–619]

[a] Many attested examples in this article come from the British National Corpus (BNC, Aston and Barnard (1998)). I use the convention of referring to the file from which it is extracted, e.g., K7D and usually also the line number within that file.

c. **Rebound questions**: A: Yeah what was your answer?
B: What was yours? [BNC, KP3, 636–637]

Apart from clarification questions, dependent questions are by far the commonest class of query response—their existence was pointed out by (Carlson (1983)). The other two classes, somewhat less common, are not 'meta' as such —they, pretheoretically, address the subject matter associated with the query. Rebound questions, which do indeed strongly implicate reluctance to address the issue originally raised, go against any idea that reduces query response felicity to their being a means of finding an answer to the initial query, as in e.g. (Asher and Lascarides (2003)); indeed while their intended effect of imposing a distinct issue is easy to characterize, the range of such questions is difficult to characterize *a priori*. Situationally pertinent questions illustrate that the 'queried situation', not merely the 'queried predicate' can be significant in calculating a possible (query) response.

1.3.2 Interjections and Non Sentential Utterances

One characteristic of conversation is the prevalence of utterances without an overt predicate. A semantics for dialogue needs to explain the meaning (= felicity and import) of such utterances. From among these one can mention initially *sentential fragments*—predicateless utterances whose content is propositional or interrogatory. Semantically oriented taxonomies for this class and corpus studies based on them can be found in (Fernández and Ginzburg (2002); Schlangen (2003)). Common types of NSUs are exemplified in bold face in (7):

(7)

a. Ann: Can you hear the birds singing? Listen. James: Er (pause) **yeah.** Ann: Can you hear? Bryony: I hear birds singing. Ann: You can hear the birds singing. **Yes.** (BNC, KB8)

b. Ann: Well put it on the draining board and I'll wash it and then put it back (pause) James: Right, I'll see ya tonight Ann: **Mhm, mhm** (pause) James: Tarrah Ann: mm, **bye** [conversation ends] (BNC, KB6)

c. A: Who's that?
B: My Aunty Peggy **[last or full name]**. My dad's sister. [BNC, G58, 33–35]

d. Tim: Those pink things that af after we had our lunch. Dorothy: **Pink things?** Tim: Yeah. Er those things in that bottle. Dorothy: Oh I know what you mean. For your throat? (BNC, KBW)

e. Cherrilyn: Are you still (pause) erm (pause) going to Bristol (pause) on Monday? Fiona: Dunno. Cherrilyn: **No?** Fiona: I dunno. Doubt it, **why?** (=*Why do you ask if I'm going to Bristol?*) (BNC, KBL)

A second class of predicate-less utterances are *interjections* (for descriptive work see e.g. Schelfhout *et al.* (2005); Norrick (2009), for formal treatments of some case studies see Kaplan (1999); Jayez (2004); Potts (2007); McCready (2008)), which as (8) makes clear fulfill a variety of functions, including conversation initiation and closing, expressing emotional attitudes towards events, and attention-getting devices:

(8) a. Mark: that's poison oak. Christa: **damn**. I don't want that on my body. Longman Grammar of Spoken American English Corpus (LSWEC-AC) Norrick (2009), p. 869

b. A: because they go at six fifty a pop. c: ((laughs)) b: **god** I know. London Lund Corpus (Svartvik and Quirk (1980), 1-10) example from Norrick (2009), p. 877

c. Cooper: can I have a bite of that cookie? Sara: **hey** they're low calorie. you can have the whole cookie. Cooper: thank you. LSWEC-AC, Norrick (2009), p. 881

Sentential fragments and interjections pose various challenges both to the semantics and to the organization of the grammar. The most basic challenge is to ensure that context is organized so as to enable resolution of the content and its relevance in the context. There is the pervasive problem that in many cases deciding what the content on a given use is not straightforward and how many distinct senses to reify, as already noted by Wittgenstein. I will exemplify this below with meta-communicative utterances. For now, I illustrate the issue of relevance specification in a case where the content resolution is straightforward. The Arabic word 'marxabteyn' is used exclusively as a response to an initial greeting, indeed typically where the initial greeting had the form 'marxaba' ('marxabteyn' is the dual form of 'marxaba').. Thus, in its lexical specification we need to find a way to encode this information.

This already hints that the notion of context required must be rich and structured, intrinsically more than the dynamic notions introduced to explicate textual meaning in the 1990s (see the articles on theories

of discourse, dynamic semantics). Indeed, NSUs exhibit varying degrees of structural parallelism between the source and the target. For example, both short answers and sluicing cross-linguistically requires cross-categorial concord between the antecedent *wh*-phrase and the target (as first pointed out, respectively, by Morgan (1973); Ross (1969)), whereas a reprise fragment on its *intended content* reading requires segmental phonological identity between source and target (Ginzburg and Cooper (2004)):

(9) a. A: lemi hixmeta? B: #moti/lemoti.
To-who flattered-2nd-sg? moti/to-moti
A: Who did you flatter? B: Moti.

b. A: et mi šibaxt? B: et moti/#lemoti.
def-acc who praised-2nd-sg? def-acc moti/to-moti
A: Who did you praise? B: Moti.

c. A: Did Bo leave? B: Bo? (Intended content reading: **Who are you referring to?** or **Who do you mean?**) /Max? (lacks intended content reading; can only mean: **Are you referring to Max?**)

Since such parallelism can be carried across a number of turns, particularly in multi-party dialogue (Ginzburg and Fernández (2005)), this means that information needed to ensure parallelism is satisfied needs to be projected into context.

1.3.3 Other repair

The pervasiveness of NSUs and interjections in dialogue is a symptom of the richness of context available to the interlocuters. There is another aspect which is absolutely fundamental to dialogue and concomitantly still missing from most formal approaches to semantics—metacommunicative interaction.

If all goes well with an utterance, we're rarely aware of the communicative process, though it's always there ticking in the background, evinced by the constant stream of backchanelling utterances and gestures produced by the participants of a conversation (for empirical work on back channels see e.g. Novick and Sutton (1994); Muller and Prevot (2003); Nakano *et al.* (2003) . Switch that off and bizarreness ensues. This process, establishing that the most recent move has been understood to the satisfaction of the conversationalists, has come to be known as *grounding*, following extensive empirical work by Herb Clark and his collaborators (Clark and Schaefer (1989); Clark and Wilkes-Gibbs (1986); Clark

(1996)). A particularly detailed semantic theory of grounding has been developed in the PTT (not an acronym!) framework, discussed further below in section 1.4.3. One concrete task for a dialogical theory is to account for the potential for and meaning of acknowledgement phrases, as in (10), either once the the utterance is completed, as in (10a), or concurrently with the utterance as in (10b):

(10) a. Tommy: So Dalmally I should safely say was my first schooling. Even though I was about eight and a half. Anon 1: Mm. Now your father was the the stocker at Tormore is that right ? (BNC, K7D)

b. A: Move the train ...
B: Aha
A:... from Avon ...
B: Right
A:... to Danville. (Adapted from the Trains corpus, Allen *et al.* (1995))

An additional task is to characterize the range of (potential) presuppositions emerging in the aftermath of an utterance, whose subject matter concerns both content and form:

(11) a. A: Did Mark send you a love letter?

b. B: No, though it's interesting **that you refer to Mark/my brother/our friend**

c. B: No, though it's interesting **that you mention sending**

d. B: No, though it's interesting **that you ask a question containing seven words.**

e. B: No, though it's interesting **that the final two words you just uttered start with 'l'**

Developing a semantic theory that can fully accommodate the challenges of grounding is far from straightforward. A more radical challenge, nonetheless, is to explicate what goes on when an addressee cannot ground her interlocutor's utterance.

More radical because this ultimately leads to seemingly radical conclusions of an intrinsic *semantic indeterminacy* (Davidson (1986)): in such a situation the public context is no longer identical for the interlocutors—the original speaker can carry on, blissfully unaware that a problem exists, utilizing a 'grounded context', whereas if the original addressee takes over the context is shifted to one which underwrites a *clarification request*. This potential context–splitting is illustrated in (12), originally

discussed in (Ginzburg (1997)). (12) illustrates that the contextual possibilities for resolving the fragment 'Bo?' are distinct for the original speaker A and the original addressee B. Whereas there is one common possibility, the short answer reading, only B has the two clarification request readings, whereas only A has a self-correction reading, albeit one that probably requires an additional elaboratory follow up:

(12)

a. A: Who does Bo admire? B: Bo?
   Reading 1 (**short answer**): Does Bo admire Bo?
   Reading 2 (**clausal confirmation**): Are you asking who BO (of all people) admires?;
   Reading 2 (**intended content**): Who do you mean 'Bo'?)
b. A: Who does Bo admire? Bo?
   Reading 1 (**short answer**): Does Bo admire Bo?
   Reading 2 (**self correction**): Did I say 'Bo'?

CRs can take many forms, as illustrated in (13):

(13)

a. A: Did Bo leave?
b. **Wot**: B: Eh? / What? / Pardon?
c. **Explicit (exp)** : B: What did you say? / Did you say 'Bo' /What do you mean 'leave'?
d. **Literal reprise** : B: Did BO leave? / Did Bo LEAVE?
e. **Wh-substituted Reprise** : B: Did WHO leave? / Did Bo WHAT?
f. **Reprise sluice** : B: Who? / What? / Where?
g. **Reprise Fragments** : B: Bo? / Leave?
h. **Gap**: B: Did Bo . . . ?
i. **Filler** : A: Did Bo . . . B: Win? (Table I from Purver (2006))

Now, as (14a) indicates, *a priori* ANY sub-utterance is clarifiable, including function words like 'the', as in (14c). While the potential for repetition-oriented CRification clearly applies to all utterances and their parts, it is an open question whether this is true for semantically/pragmatically oriented CRification. For empirical studies on this see (Healey *et al.* (2003); Purver *et al.* (2003, 2006)).

(14)

a. Who rearranged the plug behind the table?
b. Who? / rearranged?/ the plug? / behind? / the table?

c. A: Is that the shark? B: The? B: Well OK, *A*. (based on an example in the film *Jaws*.)

Moreover, *a priori* CRs can concern any aspect of the communicative process. Nonetheless, a key finding of recent corpus studies of CRs in both a general corpus (Purver *et al.* (2001)), as well as task oriented ones (Rodriguez and Schlangen (2004); Rieser and Moore (2005)), indicate that there are four main categories of CRs:

- **Repetition**: CRs that request the previous utterance to be repeated:

(15) a. Tim (1): Could I have one of those (unclear)?
Dorothy (2): Can you have what? (BNC, KW1)

b. s bust: Great memorial I think really isn't it?
e bust: Beg pardon?
s bust: Be a good appropriate memorial if we can afford it. (BNC, KM8)

- **Confirmation**: CRs that seek to confirm understanding of a prior utterance:

(16) a. Marsha: yeah that's it, this, she's got three rottweilers now and
Sarah: three? (=Are you saying she's got THREE rottweilers now?)
Marsha: yeah, one died so only got three now (BNC, KP2)

b. A: Is Georges here?
B: You're asking if Georges Sand is here.

- **Intended Content**: CRs that query the intended content of a prior utterance:

(17) a. Tim (5): Those pink things that af after we had our lunch.
Dorothy (6): Pink things?
Tim (7): Yeah. Er those things in that bottle.
Dorothy (8): Oh **I know what you mean.** For your throat? (BNC, KW1)

b. A: Have a laugh and joke with Dick.
B: Dick?
A: Have a laugh and joke with Dick.
B: Who's Dick? (BNC, KB7)

- **Intention recognition**: CRs that query the goal underlying a prior utterance:

(18) a. X: You know what, the conference might be downtown Seattle. So I may have to call you back on that.
PT: OK. Did you want me to wait for the hotel then? (Communicator corpus, National and Technology (2000))

b. Norrine: When is the barbecue, the twentieth? (pause) Something of June.
Chris: Thirtieth.
Norrine: A Sunday.
Chris: Sunday.
Norrine: Mm.
Chris: Why? (= *Why do you ask when the barbecue is*)
Norrine: Becau Because I forgot (pause) That was the day I was thinking of having a proper lunch party but I won't do it if you're going out. (BNC, KBK)

As ever, these data impose a requirement on a theory of dialogue to characterize the relevance of such utterances and to be able to describe precisely how the content of utterances such as the CRs in (15–18) emerge. An additional twist, exemplified in (19), is that the context required for this task has to be linguistically rich:

1. **Hyperintensionality**: 'lawyer' and 'attorney' are synonymous terms but give rise to distinct CRification conditions:

   (19) a. Ariadne: Jo is a lawyer. Bora: A lawyer?/What do you mean a lawyer?/#What do you mean an advocate?/#What do you mean an attorney?

   b. Ariadne: Jo is an advocate. Bora: #What do you mean a lawyer?/An advocate?/What do you mean an advocate?/#What do you mean an attorney?

2. **Parallelism**: the existence of syntactic and phonological parallelism conditions on certain interpretations of clarification requests (CRs) (9c) above and for detailed discussion see (Ginzburg and Cooper (2004); Ginzburg (2012)).
3. **Utterance tokens**: It must underwrite reference to utterance tokens, given that they are constituents of the content of CRs, though in fact this is a more general requirement concerning quotative acts in dialogue:

   (20) a. A: Max is leaving. B: leaving? (=What does 'leaving' mean in the A's sub-utterance, NOT in general.)

   b. A: Did Bo leave? B: Who is Bo?

c. A: We're fed up. B: Who is we? (= *Who is we* **in the sub-utterance needing clarification**)

We briefly discuss a tie-in of these considerations with work on quotation in section 45.

1.3.4 Genre sensitivity

The extent to which one can make linguistic generalizations in a way that *evades* genre–relativity is perhaps an open question; it is of course also very much a question for those concerned with demarcating the semantics/pragmatics boundary. Nonetheless, for anyone engaged in designing systems that can interact with humans the importance of domain specificity is taken for granted.

Genre specificity is one of the principal factors in determining, to use Herb Clark's phrase, *what drives the dialogue*—the basic intentions underlying participation in a particular conversation. That is, the range of topics that can be relevantly introduced for discussion: e.g. discussing the price of bread is reasonable in a bakery, but not—putting aside specialized circumstances—in a courtroom. In certain cases a genre determines the existence of special moves e.g. the opening and closing moves in an English court ('The court is in session.', 'The court will rise.') or even their form (e.g. the fact that in addressing a judge, one needs to end the utterance with the vocative 'm'lud'.).

Let us distinguish this notion of genre specificity (*interaction determining genre specificity*) from *language determining genre specificity*, where a genre determines its own language—viz its own lexicon and constructions. The existence of this kind of lexical genre specificity is familiar to us via the notion of 'jargon', the latter less so, but its reality should be clear—different genres have distinct language models and constructions appropriate for some domains are less so for others; for discussion of how to combine this with formal grammar, see (Cooper and Ranta (2008)). I will restrict attention here to *interaction determining genre specificity*, though any serious theory of dialogue will require integration with the other notion as well.

In addition to Conversational Relevance, genre specificity is also clearly implicated in Conversational Meaning, as we noted already in (2). NSUs can appear as initiating moves (i.e. without a prior linguistic antecedent or segment initially.). These seem to require a rather sterotypical inter-

actional setting (buying tickets at a train station, querying for directions in a taxi etc).

(21) a. Buying a train ticket:
Client: A return to Newcastle please. (=I want a return ..., please give me a return ..., ...)
b. Driver to passenger in a taxi: Where to?

Explicating how such NSUs get resolved is a basic requirement for a theory of dialogue.

1.3.5 Core Dialogue Phenomena as theoretical benchmarks

In this section I have compiled a collection of dialogue phenomena which can be used as benchmarks for theories of conversational relevance and meaning.

Taken as a whole they make clear the need for a significantly different view of context and its evolution from the influential Stalnakerian view, essentially that the effects of utterances on the state of the conversation can be derived entirely from truth-conditional meaning (suitably enriched dynamically) in interaction with general pragmatic principles.

For a start there is the fact that in dialogue one cannot appeal in general to a single context, given *context branching phenomena* discussed with respect to clarification interaction; there is the evidence that 'direct relevance' is parametrized by agent–specific intentional factors, as illustrated by rebound questions; there is the need to maintain non–content oriented information in context due to parallelism constraints on non-sentential utterances and the hyperintensionality of clarification potential.

1.4 Dialogue Frameworks

In the first part of the paper, I suggested two problems a theory of dialogue should strive to solve, Conversational Relevance and Conversational Meaning. In this section I will indicate how some existing frameworks tackle these problems and, in particular, the empirical challenges discussed in the previous section ; for a thoughtful comparison of recent dialogue frameworks, see (Poesio and Rieser (2010)). Perhaps the central development enabling progress is the emergence of a formally well

defined notion of the structure of a *conversational state* and, correspondingly, notions of how such states change.

1.4.1 KOS

Dialogue Gameboards

I start by considering the framework KoS (Ginzburg (1994); Ginzburg and Cooper (2004); Larsson (2002); Purver (2006); Fernández (2006); Ginzburg and Fernández (2010); Ginzburg (2012)). KoS—a toponym, not an acronym—is a theory that combines an approach to semantics inspired by situation semantics and dynamic semantics with a view of interaction influenced by CA. On the approach developed in KoS, there is actually no single context — instead of a single context, analysis is formulated at a level of information states, one per conversational participant. Each information state consists of two 'parts', a private part and the dialogue gameboard that represents information that arises from publicized interactions. For recent psycholinguistic evidence supporting this partition see (Brown-Schmidt *et al.* (2008)), an issue we return to in section 1.5.3.

The type definition—in a formal sense we will shortly elucidate— of an information state is given in (22a). We defer until somewhat later (brief) discussion of the structure of the private part, which typically constitutes a 'hidden variable' of dialogue analysis. For now we focus on the dialogue gameboard. Its structure is given in (22b) — the *spkr,addr* fields allow one to track turn ownership, *Facts* represents conversationally shared assumptions, *Moves* represents the contents of moves that have been grounded, *QUD* tracks the questions currently under discussion:

(22) a. $\text{TotalInformationState (TIS)} =_{def} \begin{bmatrix} \text{dialoguegameboard} : \text{DGBtype} \\ \text{private} : \text{Private} \end{bmatrix}$

b. $\text{DGBType (provisional definition)} =_{def}$

$$\begin{bmatrix} \text{spkr} : \text{Ind} \\ \text{addr} : \text{Ind} \\ \text{utt-time} : \text{Time} \\ \text{c-utt} : \text{addressing(spkr,addr,utt-time)} \\ \text{Facts} : \text{Set(Proposition)} \\ \text{Moves} : \text{list(illocutionaryProposition)} \\ \text{QUD} : \text{poset(Question)} \end{bmatrix}$$

To understand better the specification in (22), we need to make a

brief digression concerning the logical underpinnings of KoS. KoS is formulated within the framework of Type Theory with Records (Cooper (2005, 2012); Cooper and Ginzburg (2013)), a model–theoretic descendant of Martin-Löf Type Theory (Ranta (1994)) and situation semantics (Barwise and Perry (1983); Cooper and Poesio (1994); Seligman and Moss (1997); Ginzburg and Sag (2000)). TTR enables one to develop a semantic ontology, including entities such as events, propositions, and questions, whence types characterizing questions and propositions, in (22). With the same means TTR enables the construction of a *grammatical* ontology consisting of utterance types and tokens and of an interactional domain in which agents utilize utterances to talk about the semantic universe. What makes TTR advantageous for our dialogical aims is that it provides access to both types and tokens at the object level. This plays a key role in developing metacommunicative interaction, as we shall see below, in that it enables simultaneous reference to both utterances and utterance types. I propose that this constitutes a fundamental requirement on logical frameworks that aim to characterize dialogue.

For current purposes, the key notions of TTR are the notion of a *judgement* and the notion of a *record*.

- **The typing** *judgement*: $a : T$ classifying an object a as being of type T.
- **Records**: A record is a set of fields assigning entities to labels of the form (23a), partially ordered by a notion of *dependence* between the fields—dependent fields must follow fields on which their values depend. A concrete instance is exemplified in (23b). Records are used here to model events and states, including utterances, and dialogue gameboards.[b]

(23)

a. $\begin{bmatrix} l_1 = val_1 \\ l_2 = val_2 \\ \dots \\ l_n = val_n \end{bmatrix}$

[b] Cooper and Ginzburg (2013) suggest that for events with even a modicum of internal structure, one can enrich the type theory using the *string theory* developed by Tim Fernando (e.g. Fernando (2007)).

b. $\begin{bmatrix} \text{x} & = & \text{-28} \\ \text{e-time} & = & \text{2AM, Feb 17, 2011} \\ \text{e-loc} & = & \text{Nome} \\ \text{c}_{temp-at-in} & = & \text{o1} \end{bmatrix}$

- **Record Types**: a record type is simply a record where each field represents a judgement rather than an assignment, as in (24a).

(24) $\begin{bmatrix} l_1 : T_1 \\ l_2 : T_2 \\ \dots \\ l_n : T_n \end{bmatrix}$

The basic relationship between records and record types is that a record r is of type RT if each value in r assigned to a given label l_i satisfies the typing constraints imposed by RT on l_i. More precisely,

(25) The record
$\begin{bmatrix} l_1 & = & a_1 \\ l_2 & = & a_2 \\ \dots & & \\ l_n & = & a_n \end{bmatrix}$ is of type: $\begin{bmatrix} l_1 & : & T_1 \\ l_2 & : & T_2 \\ \dots & & \\ l_n & : & T_n \end{bmatrix}$
iff $a_1 : T_1, a_2 : T_2, \dots, a_n : T_n$

To exemplify this, (26b) is a possible type for (23b), assuming the conditions in (26c) hold. Records types are used to model utterance types (aka as *signs*) and to express rules of conversational interaction.

(26) a. $\begin{bmatrix} \text{x} & : & \text{Ind} \\ \text{e-time} & : & \text{Time} \\ \text{e-loc} & : & \text{Loc} \\ \text{c}_{temp-at-in} & : & \text{temp\_at\_in(e-time,e-location,x)} \end{bmatrix}$

b. -28 : Ind; 3:45AM, Feb 17, 2011 : Time; Nome : Loc; o1 : temp_at_in(3:45AM, Feb 17, 2011, Nome, -28)

The final logical notion we introduce is the situation semantics notion of an Austinian proposition Barwise and Etchemendy (1987). These are records of the form (27a). The type of Austinian propositions is the record type (27b),where the type $RecType^{\dagger}$ is a basic type which denotes the type of (*non-dependent*) record types closed under meet, join and negation. Truth conditions for Austinian propositions are defined in (27c):

(27) a. $\begin{bmatrix} \text{sit} & = & s \\ \text{sit-type} & = & T \end{bmatrix}$

b. $AustProp =_{def} \begin{bmatrix} \text{sit} & : & Rec \\ \text{sit-type} & : & RecType^{\dagger} \end{bmatrix}$

c. A proposition $p = \begin{bmatrix} \text{sit} & = & s_0 \\ \text{sit-type} & = & ST_0 \end{bmatrix}$ is true iff $s_0 : ST_0$

One important subtype of *AustProp* is the type of *illocutionary propositions* (IllocProp). These are the content of conversational moves. An example of such a proposition is in (28a), where in addition to the two standard fields for a proposition, fields exist for speaker, addressee, and utterance time and descriptive content. (28b) is the general characterization of the type IllocProp. For notational economy I will typically abbreviate (28a) as (28c), omitting the fields, retaining only the illocutionary predicate type component:

(28)

a. $\begin{bmatrix} \text{sit} & = & \text{u1} \\ \text{spkr} & = & \text{x} \\ \text{addr} & = & \text{y} \\ \text{utt-time} & = & \text{t} \\ \text{a} & = & \text{p} \\ \text{R} & = & \text{Assert} \\ \text{sit-type} & = & \left[\text{c1} : \text{R(x,y,t,p)}\right] \end{bmatrix}$

b. $\text{IllocProp} =_{def} \begin{bmatrix} \text{sit} & : & \text{Record} \\ \text{spkr} & : & \text{Ind} \\ \text{addr} & : & \text{Ind} \\ \text{utt-time} & : & \text{Time} \\ \text{a} & : & \text{AbstSemObj} \\ \text{R} & : & \text{IllocRel} \\ \text{sit-type} = \left[\text{c1} : \text{R(spkr,addr,utt-time,a)}\right] & : & \text{RecType} \end{bmatrix}$

c. R(spkr,addr,utt-time,p)

Armed with these basic logical notions, let us return to characterizing conversational states. A conversational state $c1$ will be a record $r1$ such that (29a) holds; in other words r1 shd have the make up in (29b) and the constraints in (29b) need to be met:

(29) a. r1 : DGBType

b. $\begin{bmatrix} \text{spkr} = \text{A} \\ \text{addr} = \text{B} \\ \text{utt-time} = \text{t1} \\ \text{c-utt} = \text{p}_{utt(A,B,t1)} \\ \text{Facts} = \text{cg1} \\ \text{Moves} = \langle \text{m1},\ldots,\text{mk} \rangle \\ \text{QUD} = \text{Q} \end{bmatrix}$

c. A: Ind, B: IND, t1: TIME, $\text{p}_{utt(A,B,t1)}$: addressing(A,B,t1), cg1: Set(Proposition), $\langle \text{m1},\ldots,\text{mk} \rangle$: list(illocutionaryProposition), Q : poset(Question)

Our job as dialogue analysts is to construct a theory that will explain how conversational interactions lead to observed conversational states. Let us consider how an initial conversational state looks like: initially no moves have been made and no issues introduced, so a dialogue gameboard will have the form in (30):

(30) $\begin{bmatrix} \text{spkr} & = & \text{A} \\ \text{addr} & = & \text{B} \\ \text{utt-time} & = & \text{t1} \\ \text{Moves} & = & \langle\rangle \\ \text{qud} & = & \{\} \\ \text{facts} & = & \text{cg1} \end{bmatrix}$

This allows us to write a lexical entry for a greeting word such as 'hi', as in (31), whose context—specified via the field 'dgb-params'—is supposed to be the initial state of a conversation. For a justification of this analysis of 'hi' and, more generally, the integration of illocutionary information in the grammar, see Ginzburg *et al.* (2003); Ginzburg (2012).

(31)
$$\begin{bmatrix} \text{phon} : \texttt{hi} \\ \text{cat.head} = \text{interj} : \text{syncat} \\ \text{dgb-params} : \begin{bmatrix} \text{spkr} : \text{IND} \\ \text{addr} : \text{IND} \\ \text{utt-time} : \text{TIME} \\ \text{Moves} = \langle\rangle : \text{list(LocProp)} \\ \text{qud} = \{\} : \text{set(Question)} \end{bmatrix} \\ \text{cont} = \text{Greet(spkr,ind,utt-time)} : \text{IllocProp} \end{bmatrix}$$

How do we specify the effect of a conversational move? The basic units of change are mappings between dialogue gameboards that specify how one gameboard configuration can be modified into another on the basis of dialogue moves. We call a mapping between DGB types a *conversational rule*. The types specifying its domain and its range we dub, respectively, the *pre(conditions)* and the *effects*, both of which are subtypes of DGBType. A conversational rule that enables us to explain the effect a greeting, the initial conversational move, has on the DGB is given in (32). The preconditions state that both Moves and QUD need to be empty; the sole effect is to initialize Moves with the illocutionary proposition $greet(A, B)$, A the speaker, B the addressee.

(32) Greet
$$\begin{bmatrix} \text{pre} : \begin{bmatrix} \text{spkr} & : & \text{Ind} \\ \text{addr} & : & \text{Ind} \\ \text{utt-time} : \text{TIME} & & \\ \text{moves} = \langle\rangle & : & \text{list(IllocProp)} \\ \text{qud} = \{\} & : & \text{poset(Question)} \\ \text{facts} & : & \text{Prop} \end{bmatrix} \\ \text{effects} : \begin{bmatrix} \text{spkr} = \text{pre.spkr} & : & \text{Ind} \\ \text{addr} = \text{pre.addr} & : & \text{Ind} \\ \text{utt-time} = \text{pre.utt-time} : \text{TIME} & & \\ \text{Moves} = \langle \text{Greet(spkr,addr,utt-time)} \rangle & : & \text{list(IllocProp)} \\ \text{qud} = \text{pre.qud} & : & \text{poset(Question)} \\ \text{facts} = \text{pre.facts} & : & \text{Prop} \end{bmatrix} \end{bmatrix}$$

In the sequel I employ a number of abbreviatory conventions. First, instead of specifying the full value of the list Moves, we usually record merely its first member, which we call 'LatestMove'. Second the preconditions can be written as a *merge* of two record types $DGBType^{-} \wedge_{merge} PreCondSpec$, one of which $DGBType^{-}$ is a strict supertype of DGBType and therefore represents *predictable information common to all conversational rule*s; $PreCondSpec$ represents information specific to the preconditions of this particular interaction type. Similarly,the effects can be written as a merge of two record types $DGBType^{0} \wedge_{merge} ChangePrecondSpec$, where $DGBType^{0}$ is a supertype of the preconditions and $ChangePrecondSpec$ represents those aspects of the preconditions that have changed. So we can *abbreviate* conversational rules as in (33a). For example, the abbreviated version of (32) Greet would be as in (33b).

(33) a. $\begin{bmatrix} \text{pre} & : & \text{PreCondSpec} \\ \text{effects} & : & \text{ChangePrecondSpec} \end{bmatrix}$

b. $\begin{bmatrix} \text{pre} & : & \begin{bmatrix} \text{moves} = \langle\rangle & : & \text{list(IllocProp)} \\ \text{qud} = \{\} & : & \text{poset(Question)} \end{bmatrix} \\ \text{effects} & : & \left[\text{LatestMove} = \text{Greet(spkr,addr,utt-time)} \quad : \quad \text{IllocProp}\right] \end{bmatrix}$

One final remark on this score relating to the specification of turn ownership via the labels 'spkr', 'addr'. In 2-person interaction there are three basic possibilities: the turn remains the same (32a), the turn necessarily changes (exemplified in (35) and occurring in cases such as assertion acceptance, where one cannot accept one's own assertion), or the turn is underspecified, as specified in (34a)—which picks the speaker from among the conversationalists and the addressee as the distinct other element in this set. —it is up for grabs for either conversationalist, as in (34c), *No turn change* gets abbreviated away in the notational convention introduced in (33). The other two cases will get notated explicitly—turn change by specifying the effects, whereas turn underspecification by means of a merge type conjunction, as in (34b), exemplified below in (37b) and many other cases:

(34)

a. Turnholder-Underspecified $=_{def}$

$$\left[\begin{array}{lll} \text{pre} & : & \left[\begin{array}{lll} \text{spkr} & : & \text{Ind} \\ \text{addr} & : & \text{Ind} \end{array}\right] \\ \text{effects} & : & \left[\begin{array}{lll} \text{PrevAud} = \left\{\text{pre.spkr,pre.addr}\right\} & : & \text{Set(Ind)} \\ \text{spkr} & : & \text{Ind} \\ \text{c1} & : & \text{member(spkr, PrevAud)} \\ \text{addr} & : & \text{Ind} \\ \text{c2} & : & \text{member(addr, PrevAud)} \\ & & \wedge\ \text{addr} \neq \text{spkr} \end{array}\right] \end{array}\right]$$

b. effects : TurnUnderspec $\wedge_{merge}$ $\left[\ldots\right]$

As a first illustration, we provide a lexical entry for a word like 'marx-abteyn' discussed in section 1.3.2: this entry assumes that the word presupposes that the its contextual background requires the LatestMove to be a greeting and it expresses a countergreeting; strengthening the presupposition to require the *form* of the antecedent greeting to be 'marx-aba' is possible within the refinement to KoS sketched in section 45.

(35)
$$\left[\begin{array}{l} \text{phon : } \texttt{marxabteyn} \\ \text{cat.head = interj : syncat} \\ \text{dgb-params : } \left[\begin{array}{l} \text{spkr : IND} \\ \text{addr : IND} \\ \text{utt-time : TIME} \\ \text{utt-time' : TIME} \\ \text{Moves} = \left\langle \text{Greet(addr,spkr,utt-time)} \right\rangle \text{: list(IllocProp)} \\ \text{qud} = \left\{\right\} \text{: poset(Question)} \end{array}\right] \\ \text{cont = CounterGreet(spkr,ind,utt-time') : IllocProp} \end{array}\right]$$

Parting, and concomitantly the termination of a conversation can be specified in quite similar terms, though as Ginzburg (2012) shows, they involve quite subtle presuppositions that seem absent from greeting interaction.

Querying and Assertion in Dialogue

Before considering how the actual dialogical interaction gets regulated, it is worth saying something about the "direct", coherent responses to

a question q. As we discussed above, there seem to be two types of such responses—propositional answers and queries. With respect to the latter, I assume that dependence is the central coherence relation; see (Łupkowski and Ginzburg (2014)) for account that shows how to capture the coherence of other types of question responses. Given that, I introduce the notion of a q–specific utterance:

(36) Given a question q, an utterance u is q–specific iff either:

1. $u.cont$ = p : Prop and About(p,q)
2. $u.cont$ = q1 : Question and Depend(q,q1)

(37a) says that given a question q and ASK(A,B,q) being the LatestMove, one can update QUD with q as QUD–maximal. QSPEC can be thought of as KoS' 'relevance maxim': it characterizes the contextual background of reactive queries and assertions. (37b) says that if q is QUD–maximal, then subsequent to this either conversational participant may make a move constrained to be q–specific (37c):

(37) a. Ask QUD–incrementation

$$\left[\begin{array}{l} \text{pre}: \left[\begin{array}{l} \text{q}: \quad \text{Question} \\ \text{LatestMove} = \text{Ask(spkr,addr,utt-time,q)}: \quad \text{IllocProp} \end{array}\right] \\ \text{effects}: \quad \left[\text{qud} = \left\langle \text{q,pre.qud} \right\rangle : \quad \text{poset(Question)}\right] \end{array}\right]$$

b. QSPEC

$$\left[\begin{array}{l} \text{pre}: \left[\text{qud} = \left\langle \text{q, Q} \right\rangle : \text{poset(Question)}\right] \\ \text{effects}: \text{TurnUnderspec} \wedge_{merge} \\ \left[\begin{array}{l} \text{r}: \text{Prop} \vee \text{Question} \\ \text{R: IllocRel} \\ \text{LatestMove} = \text{R(spkr,addr,utt-time,r)}: \text{IllocProp} \\ \text{c1}: \text{Qspecific(r,q)} \end{array}\right] \end{array}\right]$$

QSPEC involves factoring out turn taking from the assumption that A asking q means B answering it. In other words, the fact that A has asked q leaves underspecified who is to address q (first or at all). This is justified by data such as that in (38a,b), where the querier can or indeed needs to keep the turn, as well as multi-party cases such as (38c) where the turn is multiply distributed. It is also crucial for describing parent/infant interaction, where invariably the parent answers their own questions (Ginzburg and Moradlou (2013)):

(38)
a. Vicki: When is, when is Easter? March, April? (BNC, KC2)
b. Brian: you could encourage, what's his name? Neil. (BNC, KSR)
c. A: Who should we invite? B: Perhaps Noam. C: Martinu. D: Bedrich.

Considering assertion from a dialogical perspective means above all taking seriously the fact that in 2 person interaction A's asserting p requires B to react, either by producing an explicit acceptance utterance or gesture Clark (1996); Nakano *et al.* (2003) or by discussing whether p is indeed the case.

(39) illustrates some interesting patterns involving post-assertoric contexts: (39a,b) indicate that 'yes' can occur both in such a context and in one following polar query. (39c,d) illustrates that acceptance utterances need to be supplied by the addressee of the assertion, however such utterances are not acceptable if the asserter asks for explicit commitment, as in (39e). (39f) shows that the confirmation particle 'Right?' is only available to the original asserter, whereas the dubitative particle 'Really?' is only available to the addressee, while a same polarity tag, as in (39g,h) is available to both speakers:[c]

(39) a. A: Bo is leaving. B: Yes.
b. A: Is Bo leaving? B: Yes.
c. A: Bo is leaving. B: I see.
d. A: Bo is leaving, ♯I see.
e. A: Bo is leaving, right?. B: Right/Yes /♯I see.
f. A: Bo is leaving. B: Really?/♯Right?
g. A: Bo is leaving, ♯really?
h. A: Bo is leaving, is he?
i. A: Bo is leaving. B: Is he?

A full treatment of this pattern would take us too far afield and is a matter of vigorous current debate (see Beyssade and Marandin (2009); Farkas and Bruce (2010); Malamud and Stephenson (2011); Ginzburg (2012))—one issue being in what cases to assume that the move ('pure assertion', 'confirmation' etc) involves introducing p? into QUD, using a rule analogous to (37a). However this debate gets resolved, one thing this data emphasizes is how **grammar needs to make reference to the fine structure of conversational context** to explicate subtle differences such as those between 'Right?' and 'Really?. Possible lexical entries for these particles are sketched in (40):

[c] Thanks to Elisabet Engdahl for pointing out to me the contrast (39f,g).

(40) a.
$$\begin{bmatrix} \text{phon} : \texttt{right} \\ \text{cat.head} = \text{interj} : \text{syncat} \\ \text{dgb-params} : \begin{bmatrix} \text{spkr} : \text{IND} \\ \text{addr} : \text{IND} \\ \text{utt-time} : \text{TIME} \\ \text{utt-time'} : \text{TIME} \\ \text{LatestMove.content} = \text{Assert(spkr,addr,utt-time,p)} : \text{IllocProp} \end{bmatrix} \\ \text{cont} = \text{Check(spkr,addr,utt-time',p?)} : \text{IllocProp} \end{bmatrix}$$

b.
$$\begin{bmatrix} \text{phon} : \texttt{really} \\ \text{cat.head} = \text{interj} : \text{syncat} \\ \text{dgb-params} : \begin{bmatrix} \text{spkr} : \text{IND} \\ \text{addr} : \text{IND} \\ \text{utt-time} : \text{TIME} \\ \text{utt-time'} : \text{TIME} \\ \text{LatestMove.content} = \text{Assert(addr,spkr,utt-time,p)} : \text{IllocProp} \end{bmatrix} \\ \text{cont} = \text{Doubt(spkr,addr,utt-time',p?)} : \text{IllocProp} \end{bmatrix}$$

FACTS and QUD are coupled: a question q can be introduced only in so far FACTS does not contain information resolving q. Hence, updating FACTS involves a simultaneous downdate of QUD. This is formulated in (41): given an acceptance or confirmation of p by B, p can be unioned into FACTS, whereas QUD is modified by the function NonResolve. NonResolve is a function that maps a partially ordered set of questions $poset(q)$ and a set of propositions P to a partially ordered set of questions $poset'(q)$ which is identical to $poset(q)$ modulo those questions in $poset(q)$ resolved by members of P.

(41) Fact Update/ QUD Downdate $=_{def}$
$$\begin{bmatrix} \text{pre} : \begin{bmatrix} \text{p} & : & \text{Prop} \\ \text{LatestMove} = \text{Accept(spkr,addr,utt-time,p)} \vee & & \\ \text{Confirm(spkr,addr,utt-time,p)} & : & \text{IllocProp} \\ \text{qud} = \left\langle \text{p?,pre.qud} \right\rangle & : & \text{poset(Question)} \end{bmatrix} \\ \text{effects} : \begin{bmatrix} \text{facts} = \text{pre.facts} \cup \left\{ \text{p} \right\} & : & \text{Set(Prop)} \\ \text{qud} = \text{NonResolve(pre.qud,facts)} & : & \text{poset(Question)} \end{bmatrix} \end{bmatrix}$$

The notion of q–specificity and the conversational rules sketched above provide us with an initial characterization of query relevance and how it emerges in interaction. The approach to relevance and contextual structure and change discussed up to this point resembles closely the approach developed in work by Roberts (Roberts (1996, 2011a,b)). Roberts' framework differs from KoS with respect to certain assumptions and in terms of its basic methodological aims. Roberts abstracts away from meta-communicative interaction, hence she assumes only the existence of a single, communal context which tracks fields corresponding to FACTS, MOVES, and QUD. One of her principal aims is to offer a precise fleshing out of the Gricean program using this view of context, identifying for instance as the primary intention at any time to be the intention to address the agreed-upon (maximal element) of QUD. She uses this framework to develop detailed accounts of intonational focus, definiteness, and presupposition/implicature.

We need to consider how to incorporate genre specificity and meta-communicative interaction into our account of relevance.

Incorporating genre specificity

Psycholinguistic labs aside, all meaningful interaction occurs within a particular conversational genre / activity type / language game. As I emphasized above, this is the fundamental entity determining *what drives a dialogue*—what issues can and need to be discussed and also *how* such discussion will take place (in terms of constructions and lexica), though this latter aspect I cannot consider here.

A very basic issue is—how to classify a conversation into a genre? The approach I describe here originates with Larsson (2002), building on earlier work in AI, and subsequently somewhat recast in Ginzburg (2012). The essential idea is to provide a description of an information state of a conversational participant who has successfully completed such a conversation. This is a reasonable way of doing things as long as the final state does not lose "significant" information concerning what took place during the conversation. On the view of contextual evolution described here the final state of a conversational participant will be a DGB of the type in (42):

(42)

$$\begin{bmatrix} \text{facts} & : & \text{Set(Prop)} \\ \text{qud} = \{\} & : & \text{poset(Question)} \\ \text{moves} & : & \text{list(IllocProp)} \end{bmatrix}$$

At the end of a conversation QUD is empty,[d] but the issues that have been discussed during the conversation can be recovered by examining FACTS: we can introduce a simple refinement of FACTS update/QUD downdate such that every acceptance leads to the recording not solely of the facts added but also which questions were resolved by these facts and downdated. We can track such question using a field labelled QNUD (Questions No longer Under Discussion).[e] Final states of a conversation will then be records of type T for T a subtype of DGBType_{fin}. I rename this latter to *GenreType* since it may be identified as the general type of all conversational genres:

(43) GenreType $=_{def}$

$$\begin{bmatrix} \text{facts} & : & \text{Set(Prop)} \\ \text{qnud} & : & \text{set(question)} \\ \text{moves} & : & \text{list(IllocProp)} \end{bmatrix}$$

Let us consider two toy examples. Casual conversation among acquaintances seems to be governed by a convention that an initial greeting optionally raises as MaxQUD an issue glossable as $\lambda P.P(A)$ ('How is A'), $\lambda P.P(B)$ ('How is B'), A and B being the conversational participants. QSpec then licenses assertions such as 'You look well/as young as ever/pale etc'. In contrast, interaction in a bakery is more specifically delimited: the client needs to indicate what baked goods are desired, whereas the vendor needs to indicate how much needs to be paid. Specifications for both types are in (44):

(44)

a. CasualChat $=_{def}$

[d] Or rather, in order to end a conversation CPs need to ensure QUD is empty.

[e] Such information is also what underwrites presuppositions of *resolvedness*. Such presuppositions characterize the complements of fact embedding verbs that combine with interrogative clauses, as discussed in Ginzburg (1995b).

$$\begin{bmatrix} \text{A} & : & \text{Ind} \\ \text{B} & : & \text{Ind} \\ \text{utt-time} & : & \text{TimeInterval} \\ \text{c1} & : & \text{Speak(A,t)} \vee \text{Speak(B,t)} \\ \text{facts} & : & \text{Set(Prop)} \\ \text{qnud} & : & \text{list(question)} \\ \text{c2} & : & \left\{\lambda P.P(A), \lambda P.P(B)\right\} \subset \text{qnud} \\ \text{moves} & : & \text{list(IllocProp)} \end{bmatrix}$$

b. BakeryBuy $=_{def}$

$$\begin{bmatrix} \text{A} & : & \text{Ind} \\ \text{B} & : & \text{Ind} \\ \text{utt-time} & : & \text{TimeInterval} \\ \text{c1} & : & \text{Speak(A,t)} \vee \text{Speak(B,t)} \\ \text{facts} & : & \text{Set(Prop)} \\ \text{qnud} & : & \text{list(question)} \\ \text{c2} & : & \left\{\begin{matrix} \lambda x.\text{InShopBuy(A,x)}, \\ \lambda x.\text{Pay(A,x)} \end{matrix}\right\} \subset \text{qnud} \\ \text{moves} & : & \text{list(IllocProp)} \end{bmatrix}$$

Differing effects—e.g. whether questions needs to be discussed in a certain order— can also be described depending on whether we assume QNUD to be an unordered set, partially ordered, or a list. The appeal to genres is most crucial in explicating topic choice in *initiating* moves, where without some such notion one could in principle address any issue whatever. For reasons of space, I will not enter into the issue of initiating move specification and merely hint at how this can be done. A genre type provide a specification of how a conversation can unfold. Assuming that information states that encode this information, one can then express in precise terms the constraints a genre sets on moves in the following terms:

(45) $m0$ is felicitous relative to the current DGB $dgb0$, and $G0$, the genre one assumes the conversation to be, if and only if one believes that updating $dgb0$ with $m0$ results in an final state $dgb1$ which is a conversation of type G0.

Metacommunicative Interaction

Integrating metacommunicative interaction into the DGB involves two modifications to the picture we have had so far, one minor and one major. The minor modification, drawing on an early insight of CA, is that repair can involve 'putting aside' an utterance for a while, a while during which the utterance is repaired. That in itself can be effected without further ado by adding further structure to the DGB, specifically an additional field we will call *Pending*.

'Putting the utterance aside' raises the issue of *what is it that we are 'putting aside'*. In other words, how do we represent the utterance? Most work on (dialogue) context to date involves reasoning and representation solely on a semantic/logical level. But if we wish to explicate MCI, we already saw at the end of section 1.3.3 data from hyperintensionality, parallelism, and utterance tokens reasons indicating that we cannot limit ourselves in this way.

These considerations lead Ginzburg (2012) to conclude that the type of **Pending** must combine tokens of the utterance, its parts, and of the constituents of the content with the utterance type associated with the utterance. An entity that fits this specification is the *locutionary proposition* defined by the utterance: in the immediate aftermath of a speech event u, **Pending** gets updated with a record of the form $\begin{bmatrix}\text{sit} = \text{u} \\ \text{sit-type} = \text{T}_u\end{bmatrix}$ (of type *locutionary proposition* (LocProp)). Here T_u is a grammatical type for classifying u that emerges during the process of parsing u. In other words, an entity such as the *sign* in the sense of sign-based grammars such as Head Driven Phrase Structure Grammar (HPSG), Categorial Grammar (see e.g. Calder *et al.* (1988); Moortgat (1997)), or in versions of Lexical Functional Grammar (see e.g. Muskens (2001)). The relationship between u and T_u—describable in terms of the proposition $p_u = \begin{bmatrix}\text{sit} = \text{u} \\ \text{sit-type} = \text{T}_u\end{bmatrix}$— can be utilized in providing an analysis of grounding/CRification conditions:

(46)

a. Grounding: p_u is true: the utterance type fully classifies the utterance token.
b. CRification: p_u is false, either because T_u is weak (e.g. incomplete word recognition) or because u is incompletely specified (e.g. incomplete contextual resolution—problems with reference resolution or sense disambiguation).

This means that—incorporating also the assumption regarding non-resolvedness of questions in QUD, mentioned above—the DGB now has the type definition in (47); there is an interesting theoretical and empirical issue as to what type to associate with the elements of Moves: the LatestMove also needs to be a locutionary proposition, not merely a content, at least within a system such as that developed in (Ginzburg (2012)): speakers are assumed to update the DGB with the content of their utterances as soon as the utterance is completed. However, given the potential need to engage in clarificatory discussion concerning the utterance, backtracking to the locutionary proposition needs to be possible. Whether this applies to other moves remains to be addressed experimentally.

(47) DGBType (final version) $=_{def}$

$$\begin{bmatrix} \text{spkr} & : & \text{Ind} \\ \text{addr} & : & \text{Ind} \\ \text{utt-time} & : & \text{Time} \\ \text{c-utt} & : & \text{addressing(spkr,addr,utt-time)} \\ \text{Pending} & : & \text{list(LocProp)} \\ \text{Moves} & : & \text{list(LocProp)} \\ \text{Facts} & : & \text{Set(Prop)} \\ \text{QUD} & : & \text{poset(Question)} \\ \text{non-resolve-cond} & : & \forall q \in QUD[\neg Resolve(FACTS, q)] \end{bmatrix}$$

In principle one could have a theory of CRification based on generating all available CRs an utterance could give rise to. But in practice, as the data in (14) showed us, there are simply too many to be associated in a 'precompiled' form with a given utterance type.

Instead, repetition and meaning–oriented CRs can be specified by means of a uniform class of conversational rules, dubbed *Clarification Context Update Rules (CCURs)* in (Ginzburg (2012)). Each CCUR specifies an accommodated MaxQUD built up from a sub-utterance u1 of the target utterance, the maximal element of Pending (*MaxPending*). Common to all CCURs is a license to follow up *MaxPending* with an utterance which is *co-propositional* with MaxQud. (48) is a simplified formulation of one CCUR, Parameter identification, which allows B to raise the issue about A's sub-utterance $u0$: *what did A mean by u0?*:

(48) Parameter identification:

$$\begin{bmatrix} \text{pre} & : & \begin{bmatrix} \text{Spkr} : \text{Ind} \\ \text{MaxPending} : \text{LocProp} \\ \text{u0} \in \text{MaxPending.sit.constits} \end{bmatrix} \\ \text{effects} & : & \begin{bmatrix} \text{MaxQUD} = \begin{bmatrix} \text{q} = \lambda x \text{Mean(A,u0,x)} \\ \text{fec} = \text{u0} \end{bmatrix} : \text{InfoStruc} \\ \text{LatestMove} : \text{LocProp} \\ \text{c1: CoProp(LatestMove.cont,MaxQUD.q)} \end{bmatrix} \end{bmatrix}$$

Parameter Identification (48) underpins CRs such as (49b–49c) as follow-ups to (49a). We can also deal with corrections, as in (49d). B's corrective utterance is co-propositional with λxMean(A,u0,x), and hence allowed by the specification:

(49) a. A: Is Bo here?
b. B: Who do you mean 'Bo'?
c. B: Bo? (= Who is 'Bo'?)
d. B: You mean Jo.

As a final point, I should add a remark concerning the considerable philosophical literature on the various kinds of quotation (see e.g. Cappelen and Lepore (2012) for a survey) and a growing linguistic literature (see e.g. De Brabanter (2010) for a survey). On the whole, this work has not had much tie in with dialogue, though works such as Clark and Gerrig (1990); Recanati (2010) propose a view of (direct) quotation as involving a demonstration to a previous speech or other sound producing event. Ginzburg and Cooper (2014) argue that by utilizing utterance types and locutionary propositions as denotations for quotative constructions many of the seemingly mysterious aspects of quotation disappear.

Combining Relevance

Sofar we have characterized dialogical relevance in a modular fashion—indicating various sources of relevance—illocutionary, genre–specific, and metacommunicative. This modularity, while helpful methodologically, is clearly a theoretical artefact. One might ponder whether we actually, in practice, need a means of combining the various strands into a single, unified notion of relevance.

One argument for the need derive from cases where relevance is manifestly absent. One such case was pointed out originally by Grice and

relates to cases where an explicitly irrelevant response is provided in order to communicate the lack of wish to address a previous utterance, as in (50a,b). A related case are utterances which trigger CRs about their relevance, exemplified by (50c):

(50)

a. A: Horrible talk by Rozzo. B: It's very hot in here.

b. Harry: Is that you James? Stella: What? No, it isn't. Who is it? Harry: Where's James? Stella: He's out. Harry: Out? Oh, well, all right. I'll be straight round. Stella: What are you talking about? Who are you? (Harold Pinter 'The Collection'. In: *Plays Two*, p. 133)

c. Marjorie: Don't touch that cos she hasn't had it yet. Dorothy: Does she eat anything? Marjorie: What do you mean? (British National Corpus (BNC))

To the extent we wish to write rules that capture these inferences and/or preconditions, we need to have an *IrRelevance* predicate, directly relatable to a *Relevance* predicate.

What then does Relevance amount to? Pretheoretically, Relevance relates an utterance u to an information state I just in case there is a way to successfully update I with u. Let us restrict attention to the case where the input context is a query. Given a set of conversational rules $\mathcal{C}$, a grammar $\mathcal{G}$ and an information state $I_0 : TIS$, an utterance u is **U(tterance)**$_{\mathcal{C},\mathcal{G}}{}^{I_0}$**-relevant** iff there exist $c_1, \ldots, c_{k+1} \in \mathcal{C}, T_u \in \mathcal{G}, k \geq 0$ such that $c_1(I_0) = I_1, \ldots, c_{k+1}(I_k) = I_{k+1}$, where C's information state I_0 satisfies (51a); where by means of a sequence of updates the locutionary proposition $p_u = prop(u, T_u)$ becomes the value of LatestMove (condition (51b); and the final element of the sequence of updates I_{k+1} is such that one of the conditions in (51c-f) is satisfied—u is either q–specific, an appropriate CR, relates to the issue of willingness to discuss q, or is genre–relevant:

(51) a. I_0.DGB.LatestMove = v; v.content = Ask(A,q),

b. $I_{k+1}.DGB.LatestMove = p_u$

c. $p_u.content$ is q–specific relative to I.DGB, Or

d. $p_u.content$ is CoPropositional with some question q0 that satisfies q0 = CCUR1.effects.
maxqud($I_0.DGB.MaxPending$) for some Clarification Context Update Rule CCUR1, Or

e. $p_u.content$ is q_0–specific, where q_0 is the question ?WishDiscuss(B,q), Or

f. One of C's beliefs in I_0 is that: for some G0 there exists dgb1 such that $(I_0.DGB \bigoplus p_u) \sqsubset dgb1$, and such that dgb1 : G0

A number of remarks can be made about (51), primarily about the relata of this notion.

- The definition is relative to both the set of conversational rules and to a grammar from which the types T_u from which locutionary propositions originate.
- Relevance is, by and large, DGB oriented. Only (51f) explicitly involves reference to the entire information state.

Non Sentential Utterances

The detailed theory of context/relevance developed in previous sections enables the development of a grammar of the various types of sentential fragments discussed earlier. The basic strategy adopted in KoS to analyze NSUs is to specify construction types where the combinatorial operations integrate the (surface) denotata of the fragments with elements of the DGB. I sketch how this can be done with one such construction type, a detailed account of a wide variety of sentential fragments analyzed in such terms can be found in Fernández (2006); Ginzburg (2012).

B's utterance in (52) can receive a variety of contents, depending on the context in which it is uttered: it can be interpreted as a *short answer*, as in (52b); it can be interpreted without any prior utterance, as in (52c), though in such a case—as per Wittgenstein and Clark—the paraphrase provided here is only one of several possible; it can also be interpreted as the ('metalinguistic') correction in (52d). The different mechanisms underlying these resolutions can be uniformly described by the schema in (52e). This indicates that the content of the construction type *Declarative-fragment-clause* arises by predicating the propositional function constituted by the maximal element of QUD of the content of the bare fragment utterance, a generalization of a rule proposed already in Hausser and Zaefferer (1979). The particular content exhibited in (52b) could arise because the issue 'What did you buy in the bakery' is MaxQUD as a result of A's query; (52c) arises given that the issue 'What does the current customer want to buy' is a characteristic issue of the BakeryShopping genre (as it is of many related genres.); the content in (52d) could arise if B decided not to ground A's utterance, but using the parameter identification conversational rule to initiate repair interaction, accommodates the issue 'What did you mean by utterance 'four crescents'?' as MaxQUD.

(52) a. B: Four croissants.

b. (Context: A: What did you buy in the bakery?) Content: I bought four croissants in the bakery.
c. (Context: A: (smiles at B, who has become the next customer to be served at the bakery.)) Content: I would like to buy four croissants.
d. (Context: A: Dad bought four crescents.) Content: You mean that Dad bought four croissants.
e. *Declarative-fragment-clause*:
Cont = DGB.MaxQUD(u-frag.cont) : Prop

I have also emphasized that different NSU constructions exhibit morphosyntactic and/or phonological parallelism with their antecdents. In other words, not only the combinatorial semantics of NSU constructions integrates information from the DGB, but that this is also potentially true of the morphosyntactic and phonological specifications of such constructions. Given that parallelism, typically, exhibits a similar time course to the salience of the relevant entity of QUD, we can capture such effects by viewing QUD as tracking not simply questions qua semantic objects, but pairs of entities: a question and an antecedent subutterance. This latter entity provides a partial specification of the focal (sub)utterance, and hence it is dubbed the *focus establishing constituent* (FEC) (cf. *parallel element* in higher order unification–based approaches to ellipsis resolution e.g. Gardent and Kohlhase (1997).) Thus, the FEC in the QUD associated with a wh-query will be the wh-phrase utterance, the FEC in the QUD emerging from a quantificational utterance will be the QNP utterance, whereas the FEC in a QUD accommodated in a clarification context will be the sub-utterance under clarification. Hence the type of QUD is *InfoStruc*, as defined in (53):

(53) $$\text{Info-struc} = \begin{bmatrix} \text{q} : \text{Question} \\ \text{fec} : \text{set(LocProp)} \end{bmatrix}$$

In light of this, we can write a specification of *decl-frag-cl* as in (54). Categorially the construction is sentential, it has one dgb parameter—i.e. contextual parameter—the maximal element of QUD. Its content arises by functional application of MaxQUD to the entity denoted by the fragment and categorially the fragment has to match the categorial specification of the FEC:

(54) **declarative-fragment-clause**

$$\begin{bmatrix} \text{cat} = v : \text{syncat} \\ \text{dgb-params.max-qud} : \begin{bmatrix} \text{q} : \text{UnaryWhQuestion} \\ \text{fec} : \text{LocProp} \end{bmatrix} \\ \text{cont} = \text{max-qud.q(hd-dtr.cont.x)} : \text{Prop} \end{bmatrix}$$

$$\Big|$$

$$\text{hd-dtr:} \begin{bmatrix} \text{cat} = \text{max-qud.fec.cat} : \text{Syncat} \\ \text{cont} : \begin{bmatrix} \text{x} : \text{IND} \end{bmatrix} \end{bmatrix}$$

1.4.2 Segmented Discourse Representation Theory (SDRT)

SDRT, discussed in more detail in Asher's article on theories of discourse in this volume emerged from DRT by adding to DRT discourse relations inspired by theories of text structure such as Rhetorical Structure Theory (Mann and Thompson (1987))—the relations involved include *narration*, *explanation*, *elaboration*, *parallel*. Although SDRT was originally developed as a theory of coherence for texts (Asher (1993)), it was subsequently scaled up to apply to dialogue (Asher and Lascarides (1998, 2003)). This involves positing coherence relations across turns, including relations such as *QuestionAnswerPair (QAP)* and *Query-Elaboration*. These are used in an account of Query-Response coherence, indirect responses, discourse connectives, and non-sentential utterances (Schlangen (2003)).

In its initial formulations SDRT implicitly emphasized the continuity between text and dialogue coherence. Such continuity can indeed be recognized, to a first approximation, in that a single speaker's (uninterrupted) turn, for instance in an extended narrative, has many points in common with a text:

(55) A: Max had a great evening last night. He had a great meal. He ate salmon. He devoured lots of cheese. He then won a dancing competition.

Nonetheless, there is an intrinsic difference: the dialogical version must allow for various potential moves along the way by other interlocuters, possibilities that do not exist in a text.

(56)

A: Max had a great evening last night.
B: That's not what I heard.
A: He did though. He had a great meal.
B: Meaning what?
A: He ate salmon. He devoured lots of cheese.
B: Perhaps.
A: He then won a dancing competition.

In recent versions of SDRT intended for dialogue (e.g. Lascarides and Asher (2009)), the need to develop a distinct theory of relevance for dialogue has been emphasized, motivated in particular in work on *propositional grounding*—the assumptions dialogue participants all accept for conversational purposes. Dialogue SDRSs (DSDRSs) are defined by associating an SDRS for each participant at each turn, and accordingly the semantics of a dialogue turn is the product of the dynamic semantics for each constituent SDRS. As with its approach to text, SDRT for dialogue divides in two the task of semantico-pragmatic analysis: a glue-logic is developed to map syntactic structures to DSDRSs, whereas a defeasible cognitive logic explicates reasoning about agents' cognitive states in virtue of what they say (represented in the DSDRSs). The cognitive logic extends dynamic logics of public announcement (e.g. Baltag *et al.* (1998)) to provide default links between public announcements and cognitive attitudes. It also provides for links between game-theoretic principles and general axioms of rationality and cooperativity.

Similarly, whereas SDRT's analyses of dialogue have until recently been developed assuming cooperativity was maintained, a recent development (e.g. Asher and Lascarides (2012)) has been the abandonment of this assumption to deal with settings where this does not obtain.

1.4.3 PTT

PTT (Poesio and Traum (1997); Poesio and Rieser (2010, 2011)) shares certain commonalities with both KoS and with SDRT. Like KoS, it is an information–state based dialogue theory; it also draws inspiration from the Situation Semantics view of utterances. Indeed PTT pioneered

an emphasis on incrementality in dialogue, via the notion of *micro-conversational events* (MCEs), a notion inspired by Barwise and Perry's realist view of grammar (Barwise and Perry (1983), p. 122, recent discussion in Ginzburg (2011).). PTT shares with SDRT a DRT-inspired theoretical underpinning.

To date, work in PTT has focussed on providing highly detailed accounts of *grounding*, taking as its starting point Traum's computational model (Traum (1994)), of a variety of dialogue acts (Poesio and Traum (1997)), of collaborative utterances (Poesio and Rieser (2010)), of anaphora in dialogue, involving also visual antecedents (Poesio and Rieser (2011), gesture (Rieser and Poesio (2009)) and pointing (Rieser (2004)).

An information state is assumed to consist of three main parts:

- A private part, with information available to the participant, but not introduced in the dialogue.
- A public part consisting of the moves that are assumed by that agent to have become part of the common ground.
- A semi-public part, consisting of the information introduced with contributions that have not yet been acknowledged. This information is not yet grounded, but it is accessible.

Recent work in PTT has built on work in the philosophy of action (Bratman (1992); Tuomela (2000)) to offer a refined view of how intention recognition can drive interaction. Hence, the fact that one or more agents have a certain (possibly collective) intention, and that they are under certain obligations, may become part of the private, semi-private and public parts of an information state.

From a semantic point of view, PTT takes DRT as its starting point, specifically the compositional version from (Muskens (1996). However, the DRSs it uses to represent a discourse situation include not solely the standard DRS constituents (discourse referents for individuals, states and events), but also discourse referents for *conversational events* and associated conditions characterizing such events. This enables such events to be constituents of other conditions, e.g. acknowledgements, turn–control.

From its inception PTT has emphasized incremental interpretation, inspired by a wealth of psycholinguistic evidence (for extensive discussion and references see Rieser and Schlangen (2011).). The assumption being that the information state of a conversational participant is updated at frequent intervals—minimally, word-by-word. The term *micro-conversational event*(MCE) is used to refer to an event of uttering a sub-

sentential constituent. The update triggered by such an event involves characterizing the event on the phonological, syntactic, and meaning levels.

This incremental perspective is at the basis of PTT's analyses of various aspects of grounding, for instance back channel moves that can occur in parallel to another interlocuter's ongoing utterance. They are also a crucial component of one of PTT's signal achievements (see Poesio and Rieser (2010)), a highly explicit account of collaborative utterances such as utterance (1.2) in (57):

(57) 1.1 Instructor: So jetzt nimmst Du [pause]
well now you take
1.2 Constructor: eine Schraube
a screw
1.3 Inst: *eine* $< - >$ orangene mit einem Schlitz.
an $< - >$ *orange one with a slit*
1.4 Cnst: Ja
Yes (from The Bielefeld Toy Plane Corpus, cited in Poesio and Rieser (2010))

Completions generally involve guesswork by the addressee of the current speaker's intended next word/phrase. This guesswork becomes more justified in a collaborative task oriented setting, as in (57). Using MCEs and the Bratman/Tuomela theory of shared cooperative activity, PTT offers a detailed account of how completions can occur as a consequence of incremental intention recognition. This account is one of the most detailed existing analyses of interpretation in dialogue that integrates grammar, updates of information states in dialogue, and rich domain–specific information underpinning joint action.

1.5 Extensions: incrementality, learning, and entrainment

In this section I consider some dialogue phenomena that have not as yet been studied extensively, but that have far reaching consequences for our view of semantics.

1.5.1 Self repair

Dysfluencies are ubiquitous and observable in all but the briefest conversational interaction. Dysfluencies have been studied by researchers in CA (Schegloff *et al.* (1977)), in great detail by psycholinguists (e.g. Levelt (1983); Brennan and Schober (2001); Clark and FoxTree (2002); Bailey and Ferreira (2007)), and by computational linguists working on speech applications (e.g. Shriberg (1994); Heeman and Allen (1999)). To date, they have mostly been excluded from semantic analysis, primarily because they have been assumed to constitute low level 'noise', without semantic import. In fact, dysfluencies participate in semantic and pragmatic processes such as anaphora, conversational implicature, and discourse particles, as illustrated in (58). In all three cases, the semantic process is dependent on the *reparandum* (the phrase to be repaired) as the antecedent:

(58) a. Peter was + { well } he was] fired. (Example from Heeman and Allen (1999))
b. A: Because I, [[[any, + anyone,] + any friend,] + anyone] I give my number to is welcome to call me (Example from the Switchboard corpus) (implicature: 'It's not just her friends that are welcome to call her when A gives them her number')
c. From yellow down to brown - NO - that's red. (Example from Levelt (1983))

They also provide a particularly natural example of self-addressed queries, queries where the intended responder is the original querier:

(59) a. Carol: Well it's (pause) it's (pause) er (pause) what's his name? Bernard Matthews' turkey roast. (BNC, KBJ)
b. A: Here we are in this place, what's its name? Australia.

Since they can occur at just about any location in a given utterance and their effect is local, dysfluencies provide strong motivation for an incremental semantics, that is, a semantics calculated on a word-by-word, left to right fashion (see e.g. Steedman (1999); Morrill (2000); Kempson *et al.* (2000)). Moreover, they require the content construction process to be non-monotonic, since initial decisions can be overriden as a result of self-repair.

Ginzburg *et al.* (2014) show how, given an incremental dialogue semantics, accommodating dysfluencies is a straightforward extension of

the account discussed in section 45 for clarification interaction: the monitoring and update/clarification cycle is modified to happen at the end of each word utterance event, and in case of the need for repair, a repair question gets accommodated into QUD. Overt examples for such accommodation is exemplified in (59).

1.5.2 Learning

The lion's share of contemporary semantic work operates under the assumption that one is analyzing language qua static, shared entity, albeit one where a given form can convey radically distinct contents due to the constantly shifting context. The pervasive nature of metacommunicative interaction, discussed in sections 1.3.3, 45, and 1.5.1, indicate that we cannot maintain the simplifying assumption about a *shared* linguistic entity. For the child and the foreigner it is clear that the linguistic entity needs to be treated as dynamic. But of course the same is true for a mature speaker—in (60) Boris' lexicon changes, and just like the emergence of non-linguistic shared knowledge he can subsequently take the new knowledge for granted. Note that, in contrast to most (adult) clarification interaction, at issue here is information about a linguistic type, not token:

(60) Anja: My arrhythmia is causing me problems. Boris: arrhythmia? Anja: Erratic heart pace. Boris: Ah. (Later) Boris: Anja's arrhythmia is causing her problems. Cyprian: Oh.

The importance of a dialogical perspective for language acquisition has been emphasized in work by Clark (e.g. Clark (2012). Some formal semantic work on learning in a dialogue setting can be found in (Larsson and Cooper (2009)). Learning in a dialogue setting has been the focus of much recent computational work (e.g. Henderson *et al.* (2008)), though the paradigm in which such work is set (reinforcement learning) requires copious data, in contrast to (60) and the fast learning exhibited by children (Fernández *et al.* (2011)).

1.5.3 Conceptual Pacts and mutual adaptation

I have argued above that considerations of repair and of utterance–related presuppositions require dialogue participants to keep track of a very fine-grained record of utterances in their immediate aftermath. An obvious issue is how to construe 'immediate' and what aspects of

this fine grain get retained in the longer term. Consider (61)—Belinda's decision to use a different word from Alex could easily be construed as either corrective or as reference to a distinct entity:

(61) Alex: My sweater is really quite comfortable. Belinda: I've never seen this pullover.

There is a long standing debate in psycholinguistics concerning the extent to which common ground is actually used by participants in referential resolution. Some accounts have argued for an ego-centric perspective as a default (the Perspective-Adjustment model, Keysar *et al.* (2003)) or as a predominant factor (the Anticipation-Integration model, Barr (2008)). But there is now significant evidence, reviewed in (Brown-Schmidt (2009); Brennan *et al.* (2010)), that dialogue participants are aware of the history of reference acts to an entity—how the entity has been referred to and by whom, so that changes in the means of reference are taken to be significant. Thus, for instance (Metzing and Brennan (2003)) showed in an experiment incorporating interaction between confederate speakers and naïve addressees that the initial looks by the addressees to familiar target objects (that they had previously grounded during interaction with a speaker) were delayed by a few hundred milliseconds when the same speaker uttered an entirely new expression for the familiar object, but not when a new speaker uttered the same new expression. This is the basis for the notion of *conceptual pacts* (Brennan and Clark (1996)) between conversational participants, pacts that are not only partner-specific but also quite flexible: the first looks to the target by addressees are not delayed when a new speaker used a new expression.

Stent (2011) discusses how dialogue systems can exhibit behaviour of this kind in restricted tasks by means of keeping track of a restricted class of variables relating to an interaction (e.g. in a system that allows students to review their courses, successful adaptation can be achieved by having the system track the the form used in rating the course currently under discussion, the verb tense being used in this dialog; and the form used to refer to the instructor.).

1.6 Conclusions

Although dialogue is the primary medium of language use, phylogenetically and ontogenetically, it has yet to take centre stage in semantics.

Studying dialogue forces one to a particularly careful study of the nature of context, fundamental in much of human (inter)activity. I have focussed here on two essential tasks for a theory of dialogue, characterizing conversational relevance and conversational meaning. As far as conversational relevance goes I have sketched both empirical and theoretical characterizations, which underline it has at least three quite independent dimensions: (a) illocutionary, (b) metacommunicative, (c) genre-based. With respect to conversational meaning, I have exemplified the existence of various classes of words and constructions, notably interjections and non-sentential utterances, whose meaning is irreducibly tied to conversational contexts and can be strongly tied to their relevance.

All formal frameworks that have engaged in detailed description (including KoS, SDRT, and PTT) seem to share the assumption that semantic analysis in dialogue requires entities representing the publicized information, (a) one per conversational participant, (b) with significant internal structure, (c) tying in to a genre/task/language game, while (d) making intrinsic reference to non-semantic aspects of utterances.

I concluded by pointing to new challenges by phenomena such as dysfluencies, learning, and mutual adaptation, which suggest the need for an incremental semantics for language qua dynamic entity, with long-term, fine-grained memory of interactions.

The emergent common view concerning dialogical semantics is a very distinct picture of the nature of semantics from the standard discourse representation view (e.g. van Eijck and Kamp (1997)), or dynamic semantics view, described e.g. in (Dekker (2011)), let alone from more classical views of semantics.

1.6.1 Acknowledgements

This work was supported by the French Investissements d'Avenir–Labex EFL program (ANR-10-LABX-0083). For their very useful comments I would like to thank Jean-Marie Marandin, an anonymous reviewer for this volume, Maria Aloni, and Paul Dekker.

References

Allen, J. F., Schubert, L. K., Ferguson, G., Heeman, P., Hwang, C. H., Kato, T., Light, M., Martin, N. G., Miller, B. W., Poesio, M., and Traum, D. R. (1995). The trains project: A case study in building a conversational planning agent. *Journal of Experimental and Theoretical AI*, **7**, 7–48.

Asher, N. (1993). *Reference to Abstract Objects in English: a Philosophical Semantics for Natural Language Metaphysics*. Studies in Linguistics and Philosophy. Kluwer, Dordrecht.

Asher, N. and Lascarides, A. (1998). Questions in dialogue. *Linguistics and Philosophy*, **21**(3), 237–309.

Asher, N. and Lascarides, A. (2003). *Logics of Conversation*. Cambridge University Press, Cambridge.

Asher, N. and Lascarides, A. (2012). Agreement, disputes and commitments in dialogue. In *Proceedings of SemDial 2012 (SeineDial), the 16th Workshop on the Semantics and Pragmatics of Dialogue*, pages 31–40, Université Paris-Diderot.

Aston, G. and Barnard, L. (1998). *The BNC handbook*. Edinburgh University Press.

Bailey, K. G. D. and Ferreira, F. (2007). The processing of filled pause disfluencies in the visual world. In R. P. G. van Gompel, M. H. Fischer, W. S. Murray, and R. l. Hill, editors, *Eye Movements: A Window on Mind and Brain*, pages 485–500. Elsevier.

Baltag, A., Moss, L., and Solecki, S. (1998). The logic of public announcements, common knowledge, and private suspicions. In *Proceedings of the 7th conference on Theoretical aspects of rationality and knowledge*, pages 43–56. Morgan Kaufmann Publishers Inc.

Barr, D. (2008). Pragmatic expectations and linguistic evidence: Listeners anticipate but do not integrate common ground. *Cognition*, **109**(1), 18–40.

Barwise, J. and Etchemendy, J. (1987). *The Liar*. Oxford University Press, New York.

Barwise, J. and Perry, J. (1983). *Situations and Attitudes*. Bradford Books. MIT Press, Cambridge.

Beyssade, C. and Marandin, J.-M. (2009). Commitment : une attitude dialogique. *Langue française*, **162**, 89–108.

Bratman, M. (1992). Shared cooperative activity. *The Philosophical Review*, pages 327–341.

Brennan, S. and Clark, H. (1996). Conceptual pacts and lexical choice in conversation. *Learning, Memory*, **22**(6), 1482–1493.

Brennan, S., Galati, A., and Kuhlen, A. (2010). Two minds, one dialog: Coordinating speaking and understanding. *Psychology of Learning and Motivation*, **53**, 301–344.

Brennan, S. E. and Schober, M. F. (2001). How listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language*, **44**, 274–296.

Brown-Schmidt, S. (2009). The role of executive function in perspective taking during online language comprehension. *Psychonomic bulletin & review*, **16**(5), 893–900.

Brown-Schmidt, S., Gunlogson, C., and Tanenhaus, M. (2008). Addressees distinguish shared from private information when interpreting questions during interactive conversation. *Cognition*, **107**(3), 1122–1134.

Calder, J., Klein, E., and Zeevat, H. (1988). Unification categorial grammar: a concise, extendable grammar for natural language processing. In *Proceedings of the 12th conference on Computational linguistics - Volume 1*, COLING '88, pages 83–86, Stroudsburg, PA, USA. Association for Computational Linguistics.

Cappelen, H. and Lepore, E. (2012). Quotation. *Stanford Encyclopedia of Philosophy*.

Carlson, L. (1983). *Dialogue Games*. Synthese Language Library. D. Reidel, Dordrecht.

Chelba, C. (2010). Statistical language modelling. In A. Clark, C. Fox, and S. Lappin, editors, *Handbook of Computational Linguistics and Natural Language*, Oxford. Blackwell.

Clark, E. V. (2012). Children, conversation, and acquisition. In . K. M. M. Spivey, M. Joanisse, editor, *Cambridge handbook of psycholinguistics*. Cambridge University Press, Cambridge.

Clark, H. (1996). *Using Language*. Cambridge University Press, Cambridge.

Clark, H. and FoxTree, J. (2002). Using uh and um in spontaneous speech. *Cognition*, **84**, 73–111.

Clark, H. and Gerrig, R. (1990). Quotations as demonstrations. *Language*, **66**(4), 764–805.

Clark, H. and Schaefer, E. (1989). Contributing to discourse. *Cognitive science*, **13**(2), 259–294.

Clark, H. and Wilkes-Gibbs, D. (1986). Referring as a collaborative process. In *Arenas of Language Use*, pages 107–143. CSLI Publications, Stanford. Reprinted from a paper in Cognition.

Cooper, R. (2005). Austinian truth, attitudes and type theory. *Research on Language and Computation*, **3**(4), 333–362.

Cooper, R. (2012). Type theory and semantics in flux. In R. Kempson, N. Asher, and T. Fernando, editors, *Handbook of the Philosophy of Science*, volume 14: Philosophy of Linguistics. Elsevier, Amsterdam.

Cooper, R. and Ginzburg, J. (2013). Type theory with records for natural language semantics. In C. Fox and S. Lappin, editors, *Handbook of Contemporary Semantic Theory, second edition*, Oxford. Blackwell.

Cooper, R. and Poesio, M. (1994). Situation theory. In *Fracas Deliverable D8*. The Fracas Consortium, Centre for Cognitive Science, Edinburgh.

Cooper, R. and Ranta, A. (2008). Natural languages as collections of resources,. In R. Cooper and R. Kempson, editors, *Language in Flux: Dialogue Coordination, Language Variation, Change and Evolution*. College Publications, London.

Davidson, D. (1986). A nice derangement of epitaphs. *Philosophical grounds of rationality: Intentions, categories, ends*, **4**, 157.

De Brabanter, P. (2010). The semantics and pragmatics of hybrid quotations. *Language and Linguistics Compass*, **4**(2), 107–120.

De Ruiter, J., Mitterer, H., and Enfield, N. (2006). Projecting the end of a speaker's turn: A cognitive cornerstone of conversation. *Language*, **82**(3), 515–535.

Dekker, P. (2006). Optimal inquisitive discourse. In M. Aloni, A. Butler, and P. Dekker, editors, *Questions in Dynamic Semantics*, volume 17 of *Current Research in the Semantics/Pragmatics Interface*, pages 83–102. Elsevier, Amsterdam.

Dekker, P. (2011). Dynamic semantics. In K. von Heusinger, C. Maierborn, and P. Portner, editors, *The Handbook of Semantics*. Walter de Gruyter.

Farkas, D. and Bruce, K. (2010). On reacting to assertions and polar questions. *Journal of Semantics*, **27:1**, 81–118.

Fernández, R. (2006). *Non-Sentential Utterances in Dialogue: Classification, Resolution and Use*. Ph.D. thesis, King's College, London.

Fernández, R. and Ginzburg, J. (2002). Non-sentential utterances: A corpus study. *Traitement automatique des languages. Dialogue*, **43**(2), 13–42.

Fernández, R., Larsson, S., Cooper, R., Ginzburg, J., and Schlangen, D. (2011). Reciprocal learning via dialogue interaction: Challenges and prospects. In *Proceedings of the IJCAI 2011 Workshop on Agents Learning Interactively from Human Teachers (ALIHT)*.

Fernando, T. (2007). Observing events and situations in time. *Linguistics and Philosophy*, **30**(5), 527–550.

Gardent, C. and Kohlhase, M. (1997). Computing parallelism in discourse. In *IJCAI*, pages 1016–1021.

Ginzburg, J. (1994). An update semantics for dialogue. In H. Bunt, editor, *Proceedings of the 1st International Workshop on Computational Semantics*. ITK, Tilburg University, Tilburg.

Ginzburg, J. (1995a). Resolving questions, i. *Linguistics and Philosophy*, **18**, 459–527.

Ginzburg, J. (1995b). Resolving questions, ii. *Linguistics and Philosophy*, **18**, 567–609.

Ginzburg, J. (1996). Dynamics and the semantics of dialogue. In J. Seligman and D. Westerståhl, editors, *Logic, Language, and Computation*, pages 221–237. CSLI, Stanford, Ca.

Ginzburg, J. (1997). On some semantic consequences of turn taking. In P. Dekker, M. Stokhof, and Y. Venema, editors, *Proceedings of the 11th Amsterdam Colloquium on Formal Semantics and Logic*, pages 145–150. ILLC, Amsterdam.

Ginzburg, J. (2011). Situation semantics: from indexicality to metacommunicative interaction. In K. von Heusinger, C. Maierborn, and P. Portner, editors, *The Handbook of Semantics*. Walter de Gruyter.

Ginzburg, J. (2012). *The Interactive Stance: Meaning for Conversation*. Oxford University Press, Oxford.

Ginzburg, J. and Cooper, R. (2004). Clarification, ellipsis, and the nature of contextual updates. *Linguistics and Philosophy*, **27**(3), 297–366.

Ginzburg, J. and Cooper, R. (2014). Quotation via dialogical interaction. *Journal of Logic, Language, and Information*. Under review.

Ginzburg, J. and Fernández, R. (2005). Scaling up to multilogue: some benchmarks and principles. In *Proceedings of the 43rd Meeting of the Association for Computational Linguistics*, pages 231–238, Michigan.

Ginzburg, J. and Fernández, R. (2010). Computational models of dialogue. In A. Clark, C. Fox, and S. Lappin, editors, *Handbook of Computational Linguistics and Natural Language*, Oxford. Blackwell.

Ginzburg, J. and Moradlou, S. (2013). The earliest utterances in dialogue: towards a formal theory of parent/child talk in interaction. In R. Fernández and A. Isard, editors, *Proceedings of SemDial 2013 (DialDam)*, University of Amsterdam.

Ginzburg, J. and Poesio, M. (2014). Grammar is a system that characterizes types of talk in interaction. *Paris-Diderot and Essex Unpublished manuscript (in preparation)*.

Ginzburg, J. and Sag, I. A. (2000). *Interrogative Investigations: the form, meaning and use of English Interrogatives*. Number 123 in CSLI Lecture Notes. CSLI Publications, Stanford: California.

Ginzburg, J., Sag, I. A., and Purver, M. (2003). Integrating conversational move types in the grammar of conversation. In P. Kühnlein, H. Rieser, and H. Zeevat, editors, *Perspectives on Dialogue in the New Millennium*, number 114 in Pragmatics and Beyond new series, pages 25–42. John Benjamins Publishers.

Ginzburg, J., Fernández, R., and Schlangen, D. (2014). Disfluencies as intra-utterance dialogue moves. *Semantics and Pragmatics*.

Grice, H. P. (1989). *Studies in the Ways of Words*. Harvard University Press. Reprinted from a 1957 article.

Groenendijk, J. (2006). The logic of interrogation. In M. Aloni, A. Butler, and P. Dekker, editors, *Questions in Dynamic Semantics*, volume 17 of *Current Research in the Semantics/Pragmatics Interface*, pages 43–62. Elsevier, Amsterdam. An earlier version appeared in 1999 in the Proceedings of SALT 9 under the title 'The Logic of Interrogation. Classical version'.

Groenendijk, J. (2009). Inquisitive semantics: Two possibilities for disjunction. In P. Bosch, D. Gabelaia, and J. Lang, editors, *Proceedings of The Tbilisi Symposium on Logic, Language, and Computation 2007*, Lecture Notes in Artificial Intelligence, pages 80–94. Springer, Berlin.

Groenendijk, J. and Roelofsen, F. (2009). Inquisitive Semantics and Pragmatics. Presented at the Stanford workshop on Language, Communication and Rational Agency, May 30-31, 2009 .

Groenendijk, J. and Stokhof, M. (1984). *Studies on the Semantics of Questions and the Pragmatics ofAnswers*. Ph.D. thesis, University of Amsterdam.

Hamblin, C. L. (1970). *Fallacies*. Methuen, London.

Hausser, R. and Zaefferer, D. (1979). Questions and answers in a context dependent montague grammar. In F. Guenthner and M. Schmidt, editors, *Formal Semantics and Pragmatics for Natural Languages*. Reidel, Dordrecht.

Healey, P., Purver, M., King, J., Ginzburg, J., and Mills, G. (2003). Experimenting with clarification in dialogue. In R. Alterman and D. Kirsh, editors, *Proceedings of the 25th Annual Conference of the Cognitive Science Society*, pages 539–544. Mahwah, N.J.: LEA.

Heeman, P. A. and Allen, J. F. (1999). Speech repairs, intonational phrases and discourse markers: Modeling speakers' utterances in spoken dialogue. *Computational Linguistics*, **25**(4), 527–571.

Henderson, J., Lemon, O., and Georgila, K. (2008). Hybrid reinforcement / supervised learning of dialogue policies from fixed datasets. *Computational Linguistics*, **34**(4), 487–511.

Jayez, J. (2004). Bon : le mot de la fin. Handout from talk at the University of Geneva.

Kaplan, D. (1999). The meaning of ouch and oops: Explorations in the theory of meaning as use. *Manuscript, UCLA*.

Kempson, R., Meyer-Viol, W., and Gabbay, D. (2000). *Dynamic Syntax: The Flow of Language Understanding*. Blackwell, Oxford.

Keysar, B., Lin, S., and Barr, D. (2003). Limits on theory of mind use in adults. *Cognition*, **89**(1), 25–41.

Larsson, S. (2002). *Issue based Dialogue Management*. Ph.D. thesis, Gothenburg University.

Larsson, S. and Cooper, R. (2009). Towards a formal view of corrective feedback. In *Proceedings of the EACL 2009 Workshop on Cognitive Aspects of Computational Language Acquisition*, Athens.

Lascarides, A. and Asher, N. (2009). Agreement, disputes and commitments in dialogue. *Journal of Semantics*, **26**(2), 109–158.

Levelt, W. J. (1983). Monitoring and self-repair in speech. *Cognition*, **14**(4), 41–104.

Levinson, S. C. (2000). *Presumptive meanings: the theory of generalized conversational implicature*. MIT Press, Cambridge.

Łupkowski, P. and Ginzburg, J. (2014). Question answers. *Computational Linguistics (under review)*.

Łupkowski, P. and Wiśniewski, A. (2011). Turing interrogative games. *Minds and Machines*, pages 1–14.

Malamud, S. and Stephenson, T. (2011). Three ways to avoid commitments: Declarative force modifiers in the conversational scoreboard. pages 74–83.

Mann, W. and Thompson, S. (1987). Rhetorical structure theory: A framework for the analysis of texts. Cognitive Science Research Papers RS-87-185.

McCready, E. (2008). What man does. *Linguistics and Philosophy*, **31**(6), 671–724.

Metzing, C. and Brennan, S. (2003). When conceptual pacts are broken: Partner-specific effects on the comprehension of referring expressions. *Journal of Memory and Language*, **49**(2), 201–213.

Moortgat, M. (1997). Categorial grammar. In J. van benthem and A. ter Meulen, editors, *Handbook of Logic and Linguistics*. North Holland, Amsterdam.

Morgan, J. (1973). Sentence fragments and the notion 'sentence'. In B. Kachru, editor, *Issues in Linguistics: papers in honour of Henry and Rene Kahane*. UIP.

Morrill, G. (2000). Incremental processing and acceptability. *Computational Linguistics*, **26**(3), 319–338.

Muller, P. and Prevot, L. (2003). An empirical study of acknowledgement structures. *Proceedings of DiaBruck*, **3**.

Muskens, R. (1996). Combining montague semantics and discourse representation. *Linguistics and Philosophy*, **19**(2), 143–186.

Muskens, R. (2001). Categorial grammar and lexical-functional grammar. In *Proceedings of the LFG01 Conference, University of Hong Kong*, pages 259–279.

Nakano, Y., Reinstein, G., Stocky, T., and Cassell, J. (2003). Towards a model of face-to-face grounding. In *Proceedings of the 41st Annual Meeting of the Association for Computational Linguistics*, pages 553–561. Association for Computational Linguistics.

National and Technology (2000). DARPA communicator travel reservation corpus – June 2000 evaluation. Technical report, National Institute of Standards and Technology, Gaithersburg, MD.

Norrick, N. (2009). Interjections as pragmatic markers. *Journal of pragmatics*, **41**(5), 866–891.

Novick, D. and Sutton, S. (1994). An empirical model of acknowledgment for spoken-language systems. In *Proceedings of the 32nd Annual Meeting of the Association for Computational Linguistics*, pages 96–101, Las Cruces.

Poesio, M. and Rieser, H. (2010). (prolegomena to a theory of) completions, continuations, and coordination in dialogue. *Dialogue and Discourse*, **1**, 1–89.

Poesio, M. and Rieser, H. (2011). An incremental model of anaphora and reference resolution based on resource situations. *Dialogue and Discourse*, **1**, 235–277.

Poesio, M. and Traum, D. (1997). Conversational actions and discourse situations. *Computational Intelligence*, **13**, 309–347.

Potts, C. (2007). The expressive dimension. *Theoretical linguistics*, **33**(2), 165–198.

Purver, M. (2006). Clarie: Handling clarification requests in a dialogue system. *Research on Language & Computation*, **4**(2), 259–288.

Purver, M., Ginzburg, J., and Healey, P. (2001). On the means for clarification in dialogue. In J. van Kuppevelt and R. Smith, editors, *Current and New Directions in Discourse and Dialogue*, pages 235–256. Kluwer.

Purver, M., Healey, P. G. T., King, J., Ginzburg, J., and Mills, G. J. (2003). Answering clarification questions. In *Proceedings of the 4th SIGdial Workshop on Discourse and Dialogue*, Sapporo. ACL.

Purver, M., Ginzburg, J., and Healey, P. (2006). Lexical categories and clarificational potential. Revised version under review.

Ranta, A. (1994). *Type Theoretical Grammar*. Oxford University Press, Oxford.

Recanati, F. (2010). *Truth-Conditional Pragmatics*. Clarendon Press Oxford.

Rieser, H. (2004). Pointing in dialogue. In J. Ginzburg and E. Vallduvi, editors, *Proceedings of Catalog'04, The 8th Workshop on the Semantics and Pragmatics of Dialogue*, Universitat Pompeu Fabra, Barcelona.

Rieser, H. and Poesio, M. (2009). Interactive gesture in dialogue: a ptt model. In P. G. T. Healey, R. Pieraccini, D. K. Byron, S. Young, and M. Purver, editors, *Proceedings of SIGDIAL 2009*, pages 87–96. The Association for Computational Linguistics.

Rieser, H. and Schlangen, D. (2011). Special issue on incremental processing in dialogue. *Dialogue and Discourse*, **2**.

Rieser, V. and Moore, J. (2005). Implications for generating clarification requests in task-oriented dialogues. In *Proceedings of the 43rd Meeting of the Association for Computational Linguistics*, Michigan.

Roberts, C. (1996). Information structure in discourse: Towards an integrated formal theory of pragmatics. *Working Papers in Linguistics-Ohio State University Department of Linguistics*, pages 91–136. Reprinted in Semantics and Pragmatics, 2012.

Roberts, C. (2011a). Information structure: Afterword. *Semantics and Pragmatics*.

Roberts, C. (2011b). Solving for interpretation. *Ms., OSU*.

Rodriguez, K. and Schlangen, D. (2004). Form, intonation and function of clarification requests in german task-oriented spoken dialogues. In J. Ginzburg and E. Vallduvi, editors, *Proceedings of Catalog'04, The 8th Workshop on the Semantics and Pragmatics of Dialogue*, Universitat Pompeu Fabra, Barcelona.

Ross, J. (1969). Guess who. In *Proceedings of the 5th Annual Meeting of the Chicago Linguistic Society*, pages 252–286. CLS.

Sacks, H., Schegloff, E., and Jefferson, G. (1974). A simplest systematics for the organization of turn-taking for conversation. *Language*, **50**(4), 696–735.

Schegloff, E. (1987). Some sources of misunderstanding in talk-in-interaction. *Linguistics*, **25**, 201–218.

Schegloff, E., Jefferson, G., and Sacks, H. (1977). The preference for self-correction in the organization of repair in conversation. *Language*, **53**, 361–382.

Schelfhout, C., Coppen, P., Oostdijk, N., and van der Slik, F. (2005). Interjections in dutch: a corpus-based approach. *NEERLANDISTIEK. NL*, **5**.

Schlangen, D. (2003). *A Coherence-Based Approach to the Interpretation of Non-Sentential Utterances in Dialogue*. Ph.D. thesis, University of Edinburgh, Edinburgh.

Seligman, J. and Moss, L. (1997). Situation Theory. In J. van Benthem and A. ter Meulen, editors, *Handbook of Logic and Linguistics*. North Holland, Amsterdam.

Shriberg, E. E. (1994). *Preliminaries to a theory of speech disfluencies*. Ph.D. thesis, University of California at Berkeley, Berkeley, USA.

Sperber, D. and Wilson, D. (1986). *Relevance*. Basil Blackwell, Oxford. A second edition appeared in 1995.

Steedman, M. (1999). *The Syntactic Process*. Linguistic Inquiry Monographs. MIT Press, Cambridge.

Stent, A. (2011). Shared experiences, shared representations, and the implications for applied natural language processing. In *Proceedings of FLAIRS*.

Stivers, T., Enfield, N., Brown, P., Englert, C., Hayashi, M., Heinemann, T., Hoymann, G., Rossano, F., De Ruiter, J., Yoon, K., *et al.* (2009). Universals and cultural variation in turn-taking in conversation. *Proceedings of the National Academy of Sciences*, **106**(26), 10587.

Svartvik, J. and Quirk, R. (1980). *A Corpus of English Conversation*. Liber Laromedel Lund, Lund.

Tagliamonte, S. (2006). *Analysing sociolinguistic variation*. Cambridge University Press.

Traum, D. (1994). *A Computational Theory of Grounding in Natural Language Conversations*. Ph.D. thesis, University of Rochester.

Tuomela, R. (2000). *Cooperation: A philosophical study*, volume 82. Springer.

Turing, A. (1950). Computing machinery and intelligence. *Mind*, **59**(236), 433–460.

van Eijck, J. and Kamp, H. (1997). Representing discourse in context.

Walton, D. (2011). Burden of proof in a modified hamblin dialogue system. *Informal Logic*, **31**(4).

Wittgenstein, L. (1953). *Philosophical Investigations*. Basil Blackwell, Oxford. Citations from second edition, 1988 reprint.