Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 185/215

Grammars as Processes for Interactive Language Use: Incrementality and the Emergence of Joint Intentionality

Eleni Gregoromichelaki and Ruth Kempson

Abstract Recent research in the formal modelling of dialogue has led to the conclusion that bifurcations like language use versus language structure, competence versus performance, grammatical versus psycholinguistic/pragmatic modes of explanation are all based on an arbitrary and ultimately mistaken dichotomy, one that obscures the unitary nature of the phenomena because it insists on a view of *grammar* that ignores essential features of natural language (NL) processing. The subsequent radical shift towards a conception of NL grammars as procedures for enabling interaction in context (Kempson et al. 2009a, b) now raises a host of psychological and philosophical issues: The ability of dialogue participants to take on or hand over utterances mid-sentence raises doubts as to the constitutive status of Gricean intention-recognition as a fundamental mechanism in communication. Instead, the view that emerges, rather than relying on mind-reading and cognitive state metarepresentational capacities, entails a reconsideration of the notion of communication and a non-individualistic view on meaning. Coordination/alignment/intersubjectivity among dialogue participants is now seen as relying on low-level mechanisms like the grammar (appropriately conceived).

1 Introduction

Following Chomsky (1965), there has been a widespread perception, until recently, that formal accounts of natural language (NL) grammars must be grounded in the description of sentence-strings without any reflection of the

E. Gregoromichelaki (✉) · R. Kempson
King's College London, London, UK
e-mail: eleni.greg@blueyonder.co.uk

A. Capone et al. (eds.), *Perspectives on Linguistic Pragmatics*, Perspectives in Pragmatics, Philosophy & Psychology 2, DOI: 10.1007/978-3-319-01014-4_7,

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 186/215

dynamics of language performance. Departures from this anti-functionalist methodology were rejected on the basis that language use is often disfluent and disorderly, hence presumed to preclude rigorous systematization, a stance independently propounded by the antiformalist approach of Ordinary Language philosophy (Austin 1975) and followed up by many theoretical approaches to pragmatics. However, structural, formal accounts consistent with performance considerations are now being considered (see e.g. Newmeyer 2010), as witness the huge growth in context-modelling and information update in formal semantics since the development of DRT and related frameworks. However, when required to interface with standard grammar formalisms, these developments in formal semantics/pragmatics are now beginning to show that the standard methodological dichotomies, e.g. language use versus language structure, competence versus performance, grammatical versus psycholinguistic/pragmatic modes of explanation seem problematic. This is because all phenomena of NL context-dependency are explainable only by bifurcating them into grammar-internal versus grammar-external/discourse processes. This is because NL grammars are, on the one hand, taken to be limited to phenomena occurring within sentence boundaries but, on the other, unable to reflect the incremental word-by-word comprehension and production at the subsentential domain. However, context-dependency phenomena—anaphora, ellipsis, tense-construal, quantification, etc.—all allow unified ways of resolving how they are to be understood within and across sentence boundaries and even across distinct interlocutor turns in dialogue (Purver et al. 2009; Gregoromichelaki et al. 2011). And these update mechanisms are constrained at all levels by the incremental nature of processing. Hence, in this chapter, we suggest that these bifurcations -language use versus language structure, competence versus performance, grammatical versus psycholinguistic/pragmatic modes of explanation- are all based on an arbitrary and ultimately mistaken dichotomy of phenomena, one that obscures their unitary nature because it insists on a view of grammar that ignores essential features of NL processing like incremental update.

As a response to such considerations, grammatical models have recently begun to appear that reflect aspects of performance to varying degrees (e.g. Purver 2006; Fernandez 2006; Ginzburg 2012; Gregoromichelaki et al. 2011; Hawkins 2004; Phillips 1996; Sturt and Lombardo 2005; Ginzburg and Cooper 2004; Kempson et al. 2001; Cann et al. 2005). One such model, *Dynamic Syntax* (DS), has the distinctive characteristic of taking a fundamental feature of real-time processing—the concept of underspecification and incremental goal-directed update—as the basis for grammar formulation. This shift of perspective has enabled the modelling of core syntactic phenomena as well as phenomena at the syntax-semantics-pragmatics interface in a unified and hence explanatory way (see e.g. Kempson et al. 2001; Cann et al. 2005; Kempson et al. 2011b). Moreover, instead of ignoring dialogue data as beyond the remit of grammars, DS takes the view that joint-construal of meaning in dialogue is fundamentally based on the same mechanisms underlying language structure: structure is built through incremental procedures, that integrate context in every step, and this provides principled explanations for the syntactic properties of linguistic signals; but, in addition, since

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 187/215

the grammar licenses partial, incrementally constructed structures, speakers can start an utterance without a fully formed intention/plan as to how it will develop relying on feedback from the hearer to shape their utterance and its construal and this provides the basis for the joint derivation of structures, meaning and action in dialogue. Thus, with grammar mechanisms defined as inducing growth of information and sustaining interactivity, the availability of derivations for genuine dialogue phenomena from within the grammar shows how core dialogue activities can take place without any other-party meta-representation at all. From this point of view then, communication is not definitionally the full-blooded intention-recognising activity presumed by Gricean and post-Gricean accounts. This then leads to questions regarding fundamental notions in philosophy and pragmatics, namely, the status of notions like intentions, common ground and linguistic versus extra-linguistic knowledge and their role in communication. We turn to examine those questions next.

2 Rethinking Intentionalism[1] in Communication

2.1 Intentions, Common Ground and Communication

The noted discrepancies between the representations delivered by the *grammar*, i.e. syntax/semantics mappings ('sentence meaning' or encoded content), and 'speaker meaning' (conveyed content) led to Grice's account of $meaning_{NN}$, (Grice 1975) to become the point of departure for many subsequent pragmatic models (see Levinson 1983; Bach 1997; Bach and Harnish 1982; Cohen et al. 1990, Searle 1969, 1983 a.o.).[2] From this point of view, it has been seen as necessary that, beyond some modular linguistic knowledge, communication should essentially involve notions of rationality and cooperation. In certain versions, this is displayed by the requirement that interpretation must be guided by reasoning about mental states: speaker's meaning, whose recovery is elevated as the fundamental criterion for successful communication, involves the speaker, at minimum, (a) having the intention of producing a response (e.g. belief) in the addressee (i.e. having a thought about the addressee's thoughts) and (b) also having a higher order intention regarding the addressee's belief about the speaker's second order thought (in order to capture the presumed fulfilment of the communicative intention by means of its recognition). Under this definition, speakers must, in order to

[1] The term is from Levinson (1995: 228) denoting the view that any kind of interaction involves an attribution of meaning or intention to the other.

[2] Note that our arguments here do not necessarily concern Grice's philosophical account, in so far as it is seen by some as just normative, but its employment in subsequent (psychological/computational) models of communication/pragmatics.

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 188/215

communicate, have (at least) fourth order thoughts and hearers must recover the speaker's meaning through reasoning about these thoughts.

Millikan (1984: Chap. 3, 2005) argues that the standard Gricean view, with its heavy emphasis on mind-reading (see Cummings, this volume) over-intellectualises communication. Unlike the Gricean conception of meaning$_{NN}$ which rules out causal effects on the audience, e.g. involuntary responses in the hearer, Millikan's account, to the contrary, examines language and communication on the basis of phenomena studied by evolutionary biology, with linguistic understanding seen as analogous to direct perception rather than reasoning (see also McDowell 1980)[3]: Objects of ordinary perception, e.g. vision, are no less abstract than linguistic meanings, both requiring contextual enrichment through processing of the incoming data in order to be comprehended. Yet, in the case of ordinary perception, this processing does not require any consideration of someone else's intention. An analogous assumption can then be made as regards linguistic understanding, so that the resolution of underspecified input in context does not require considering interlocutors' mental states as a necessary ingredient. Millikan then provides an account of linguistic meaning in a continuum with natural meaning based on the *function* that linguistic devices have been selected to perform (their survival value). These functions are defined through what linguistic entities are supposed to do (not what they normally do or are disposed to do) so that "function", in Millikan's sense, becomes a normative notion. Norms of language, "conventions", are uses that had survival value, and meaning is thus equated with function. In contrast then to accounts of intentional action which see the structures involved as distinctive of rational agents, distinguishing them from entities exhibiting merely purposive behaviour (see, e.g. Bratman 1999: 5), in Millikan's naturalistic perspective, function, i.e. meaning, does not depend upon speaker intentions. Nonetheless, speakers indeed can be conceived as behaving purposefully in producing tokens of linguistic devices (as hearts and kidneys behave purposefully) but without representing hearers' mental states or having intentions about hearers' mental states (see also Csibra and Gergely 1998; Csibra 2008). Similarly, hearers understand speech through direct perception of what the speech is about without necessary reflection on speaker intentions.[4,5]

[3] The strict dichotomy between "meaning$_{NN}$" and "showing" has also been disputed within Relevance Theory (see, e.g., Wharton 2003).

[4] Of course, adults can, and often do, use reflections about the interlocutor's mental states; but the point is that this is not a necessary ingredient for meaningful interaction. Gricean mechanisms, that is, can be invoked but only as derivative or in cases of failure of the normal functioning of the primary mechanisms involved in the recovery of meaning, such as deception, specialised domains of discourse etc.

[5] An alternative account of communication combining Gricean and Millikanesque perspectives is that of Recanati (2004), which makes Gricean higher-order intention recognition a prerequisite only for implicature reconstruction. For what he terms "primary processes", on the other hand, Recanati adopts Millikan's account of understanding-as-direct-perception for the pragmatic processes that are involved in the determination of the truth-conditional content of an underspecified linguistic signal. These processes are blind and mechanical relying on

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 189/215

Early on, philosophers like Strawson (1964) and Schiffer (1972) severally presented scenarios where the criterion of higher-order intention recognition was satisfied even though this still was not sufficient for the cases to be characterised as instances of "communication" (as opposed to covert manipulation, "sneaky intentions" etc.). This led to the postulation of successively higher levels of intention recognition as a prerequisite for communication, and an attendant concept of "mutual knowledge" of speaker's intentions, both of which were recognised as facing a charge of infinite regress (see e.g. Sperber and Wilson 1995: 256–77). Although in applications of this account in psychological implementations it is not necessary to assume that explicit reasoning takes place online, nevertheless, an inferentially-driven account of communication on this basis has to provide a model that explicates the concept of 'understanding' as effectively analysed through an inferential system that implements these assumptions (see e.g. Allott 2005). So, even though such a system can be based on heuristics that short-circuit complex chains of inference (Grice 2001: 17), the logical structure of the derivation of an output has to be transparent if the implementation of that model is to be appropriately faithful (see e.g. Grice 1981: 187 on the 'calculability' of implicatures). Agents that are not capable of grasping this logical structure independently cannot be taken to be motivated by such computations, except as an idealisation pending a more explicit account. On the other hand, ignoring in principle the actual mechanisms that implement such a system as a competence/performance issue, or an issue involving Marr's (Marr 1982) computational versus the algorithmic and implementational levels of analysis (see e.g. Stone 2005, 2004; Geurts 2010) does not shield one from charges of psychological implausibility: if the same effects can be accounted for with standard psychological mechanisms, without appeal to the complex model, then, by Occam's razor, such an account would be preferable, especially if subtle divergent predictions can be uncovered (as in e.g. Horton and Gerrig 2005).

In this respect then, a range of psycholinguistic research suggests that recognition of intentions is an unduly strong psychological condition to impose as a prerequisite to effective communication. First, there is the problem of autism and related disorders. Autism, despite being reliably associated with inability (or at least markedly reduced capacity) to envisage other people's mental states, is not a syndrome precluding first-language learning in high-functioning individuals (Glüer and Pagin 2003). Secondly, language acquisition across children is established well before the onset of ability to recognise higher-order intentions (Wellman et al. 2001), as evidenced by the so-called 'false-belief task' which necessitates the child distinguishing what they believe from what others believe (Perner 1991). Given that language-learning takes place very largely through the

(Footnote 5 continued)
'accessibility' so that no inference or reflection of speaker's intentions and beliefs is required. It is only at a second stage, for the derivation of implicatures, that genuine reasoning about mental states comes into play.

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 190/215

medium of conversational dialogue, these results appear to show that at least communication with and by children cannot rely on higher-order intention recognition.

Such evidence has led to a move within Relevance Theory (RT) (Sperber and Wilson 1995) weakening further its Gricean assumptions (Breheny 2006). The RT view of communication is that the content of an utterance is established by a hearer relative to what the speaker could have intended (relative also to a concept of 'mutual manifestness' of background assumptions). This explanation involves meta-representation of other people's thoughts, but the process of understanding is effected by a mental module enabling hypothesis construction about speaker intentions. As noted by RT researchers, along with the communicated propositions, the context for interpretation falls under the speaker's communicative intention and the hearer selects it (in the form of a set of conceptual representations) on this basis. So, even though, unlike common ground, mutual manifestness of assumptions is in principle computable by conversational participants, and the interpretation process is not a "rational" one in the sense of Grice (cf. Allott 2008), it still remains the case that speaker meaning and intention are the guiding interpretive criteria which are implemented on mechanisms that have evolved to effect mind-reading. For this reason, Breheny argues that children in the initial stages of language acquisition communicate relative to a weaker 'naive-optimism' strategy in which some context-established interpretation is simply presumed to match the speaker's intention, only coming to communicate in the full sense substantially later (see also Tomasello 2008). In effect, this presents a non-unitary view of communication, which, based on the occasional sophistication that adult communicators exhibit, radically separates the abilities of adult communicators from those of children and high-functioning autistic adults.

But there is also very considerable independent evidence that even though adults are able to think about other people's perspectives, they are significantly influenced by their own point of view (*egocentrism*) (Keysar 2007). This suggests that the complex hypotheses required by Gricean reasoning in communication may not reliably be constructed by adults either.[6] This is corroborated by an increasingly large body of research demonstrating that Gricean "common ground" is not a necessary building block in achieving coordinative communicative success: speakers regularly violate shared knowledge at first pass in the use of anaphoric and referential expressions which supposedly demonstrate the necessity of established common ground (Keysar 2007, a.o.).[7] Given this type of observation, checking in parsing or producing utterances that information is jointly held by the dialogue participants—the perceived *common ground* (see Allan, this volume)—

[6] Indeed, it is useful to note that even adults fail the false belief task, if it is a bit more complex (Birch and Bloom 2007).

[7] Though 'audience design' and coordination effects are regularly observed in experiments (see e.g. Hanna et al. 2003), these can be shown to result from general memory-retrieval mechanisms rather than as based on some common ground calculation based on metarepresentation or reasoning (see Horton and Gerrig 2005; Pickering and Garrod 2004).

cannot be a necessary condition on such activities. And there is psycholinguistic evidence that such neglect of common ground does not significantly impede successful communication and is not even detected by participants (Engelhardt et al. 2006, a.o.). Moreover, if such data are set aside as exceptional or unsuccessful acts of communication, one is left without an account of how people manage to understand what each other has said in these cases. But it is now well-documented that "miscommunication" phenomena not only provide vital insights as to how language and communication operate (Schegloff 1979), but also facilitate coordination: as Healey (2008) shows, the local processes involved in the detection and resolution of misalignments during interaction lead to significantly more positive effects on measures of successful interactional outcomes (see also Brennan and Schober 2001; Barr 1998). In addition, these localised procedures lead to more gradual, group-level modifications, which in turn account for language change. It seems then from this perspective that the Gricean and neo-Gricean focus on detecting speaker meaning as the sole criterion of communicative success misrepresents the goals of human interaction: miscommunication (which is an inevitable ingredient in the interaction of interlocutors that do not share a priori common ground) and the specialised repair procedures made available by the structured linguistic and interactional resources available are the main means that can guarantee intersubjectivity and coordination; and, as Saxton (1997) shows, in addition, such mechanisms, in the form of negative evidence and embedded repairs (see also Clark and Lappin 2011), crucially mediate language acquisition (see also Goodwin 1981: 170–171).

2.2 *Joint Intentions, Planning and Dialogue Modelling*

More recently, work in philosophy has started exploring notions of joint agency/joint action/joint intentions (see e.g. Searle 1990, 1995; Bratman 1990, 1992, 1993, 1999; Gilbert 1996, 2003; Tuomela 1995, 2005, 2007 a.o.). As the Gricean individualistic view of speaker's intention being the sole determinant of meaning underestimates the role of the hearer, current dialogue models have turned to Bratman's account of *joint intentions* to model participant coordination. The controversial notion of 'intention' as a psychological state has been explicated in terms of hierarchical planning structures (Bratman 1990), a view generally adopted in AI models of communication (see, e.g. Cohen et al. 1990). In this type of account, collective intentions are reduced to individual intentions and a network of mutual beliefs. A similar style of analysis features prominently in H. Clark's model: dialogue involves joint actions built on the coordination of (intention-driven) individual actions based on shared beliefs (*common ground*):

> What makes an action a joint one, ultimately, is the coordination of individual actions by two of more people (Clark 1996: 59).

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 192/215

In this respect, a strong Gricean element underlies the psycholinguistic and computational modelling of dialogue reflecting reasoning about speakers' intentions even though now supported by an account in terms of joint action and conversational structure. Thus, within psycholinguistics and (computational) semantics, the move from individualistic accounts of action, planning and intention to joint action and coordination in dialogue has seen the latter as derivative.

However, joint action seems to involve a number of lower-level cognitive phenomena that cannot be easily explicated in Gricean terms. We should distinguish here between the terms 'coordination' and 'cooperation': cooperation is taken as involving a defined shared goal between interlocutors whereas coordination is the dynamically matched behaviour of two or more agents so that it might appear that there is a joint purpose, whether there is one or not (see also Allott 2008: 15). In this respect, psycholinguistic studies on dialogue have demonstrated that when individuals engage in a joint activity, such as conversation, they become "aligned", i.e. they (unconsciously) synchronise their behaviour at a variety of different levels, e.g. bodily movements, speech patterns etc. These coordinations draw on subpersonal, synchronised mechanisms (Pickering and Garrod 2004) or emotional, sensory-motor practices that are, crucially, nonconceptual (Gallagher 2001: 81; Hutto 2004).

From this perspective, taking the individualistic conception of intention in, e.g. Bratman's analysis as the basis of conversational dialogue seems either conceptually or cognitively implausible (Tollefsen 2005; Becchio and Bertone 2004). In this connection, the Schiffer and Strawson scenarios mentioned earlier that led to a more complicated picture of utterance meaning seem to show, in fact, that Gricean assumptions are on the wrong footing as a foundation for accounts of communication: The method of generalising from these elaborate cases to cases of ordinary conversation makes it inevitable that paradoxes will be generated, e.g. the *mutual knowledge paradox* (Clark and Marshall 1981), according to which, interlocutors have to compute an infinite series of beliefs in finite time. The dilemma here is that there is plenty of evidence for *audience design* in language production, a type of (seemingly) cooperative, coordinative behaviour, posing the problem of how to model the interlocutors' abilities allowing them to achieve this during online processing. But the solution to such problems, ideally, should not replicate the problematic structure involved (as in, e.g. Clark and Marshall 1981, who assume that interlocutors carry around detailed models of the people they know which they consult when they come to interact with them). Replacing such accounts with a psychological perspective that focuses on the lower-level mechanisms involved can undercut the intractability of such solutions by invoking independently established memory mechanisms that provide explanation of how people appear to achieve "audience designed" productions without in fact constructing explicit models of the interlocutor or metarepresentations. In this respect, Horton and Gerrig (2005) show, through subtle experimental manipulations, that the ordinary retrieval of episodic memory traces during interaction predicts much better both participants' conformity but also, and more crucially, their deviations from the assumptions derived from the "common ground" idealisation.

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 193/215

In the same spirit, empirical Conversational Analysis (CA) accounts of the sequential coherence of conversations emphasise the importance of the turn-by-turn organisation of dialogue which allows juxtaposition of displays of participant understandings and provides structures for organised repair (see e.g. Schegloff 2007). Rather than interlocutors having to figure out each other's mental states and plans through metarepresentational means, conversational organisation provides the requisite structure for coordination through repair procedures and routines. Accordingly, as Garrod and Anderson (1987) observe, in task-oriented dialogue experiments, explicit negotiation is neither a preferential nor an effective means of coordination, as would be expected to be if reasoning about speaker plans and common ground were the primary means of coordination. Explicit negotiation, if it occurs at all, usually happens after participants have already developed some familiarity with the task. Hence, the Interactive Alignment model developed by Pickering and Garrod (2004) emphasizes the importance of tacit alignment mechanisms and implicit common ground as the primary means of coordination. The establishment of routines and the significance of repair as externalised inference are also noted by Pickering and Garrod. Further psycholinguistic experiments reported in Mills and Gregoromichelaki (2008, 2010) and Mills (2011) suggest that, by probing the process of coordination in task-oriented dialogue, it can be demonstrated that notions of joint intentions and plans emerge gradually in a regular manner, rather than guiding utterance production and interpretation throughout. The hypothesis that these implicit means, rather than intention recognition, are the primary method of coordination is probed in these experiments by inserting artificial clarifications regarding intentions (*why?*) and observing the responses they receive at initial and later stages of rounds of games. At early stages, individuals display little recognition of specific intentions/plans underpinning their own utterances and explicit negotiation is either ignored or more likely to impede (see also Mills 2007; Healey 1997). This is because participants have not yet figured out the structure of the task, hence they do not have yet developed a metalanguage involving plan and intention attribution in order to explicitly negotiate their purposes. As CA research indicates, this then implies that discursive constructs such as "intentions" need to emerge, even in such task-oriented joint projects. Initially, participants seem to follow trial-and-error strategies to figure out what the task involves and coordinate their responses. These strategies and the routines participants develop lead, at later stages of the games, to highly coordinated, efficient interaction and, at this stage, issues of "intention/plan" can be raised. These results appear to undermine both accounts of coordination that rely on an a *priori* notion of (joint) intentions and plans (e.g. Bratman 1990) and also accounts which rely on some kind of strategic negotiation/agreement to mediate coordination. This is because it seems that, even in such task-specific situations, joint intentionality is not guaranteed ab initio but rather has to evolve incrementally with the increasing expertise.

These observations seem consonant with an alternative approach to planning and intention-recognition according to which forming and recognising such constructs is a subordinated activity to the more basic processes that underlie people's

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 194/215

performance (see e.g. Suchman 1987/2007; Agre and Chapman 1990). Given the known intractability of notions like plan recognition and common ground/mutual knowledge computation (see, e.g. Levinson 1995), computational models of dialogue, even when based on generally Clarkian theories of common ground, have now largely been developed without explicit high-order meta-representations of other parties' beliefs or intentions except where dealing with complex dialogue domains (e.g. non-cooperative negotiation, Traum et al. 2008). With algorithmically defined concepts such as *dialogue gameboard*, *QUD*, (Ginzburg 2012; Larsson 2002) and default rules incorporating rhetorical relations (Lascarides and Asher 2009; Asher and Lascarides 2008), the necessity for rational reconstruction of inferential intention recognition is largely sidestepped (though see Lascarides and Asher 2009; Asher and Lascarides 2008 for discussion). Even models that avow to implement Gricean notions (see e.g. Stone 2005, 2004) have significantly weakened the Gricean reconstruction of the notion of "communicative intention" and meaning$_{NN}$, positing instead representations whose content does not directly reflect the logical structure (e.g. reflexive or iterative intentions) required by a genuine Gricean account.

The philosophical underpinnings of dialogue models that rely on Gricean notions are sought in accounts that explicate intentions as mental states, independent of and prior to intentional action. However, the tradition following late Wittgensteinian ideas sees 'intention' as part of a discursive practice (Anscombe 1957) rather than a term referring to an actual mental state. Accordingly, language is to be understood as action, rather than the means of allowing expression of inner, unobservable cognitive entities. Such approaches criticise standard dialogue models, e.g. H. Clark's theory, based on the claim that that these approaches retain a communication-as-transfer-between-minds view of language treating intentions and goals as pre-existing private inner states that become externalised in language (see, e.g. Hutto 2004). In contrast, philosophers like Brandom (1994) eschew the individualistic character of accounts of meaning espoused by the Gricean perspective, analysing meaning/intentionality as arising out of linguistic social practices, with meaning, beliefs and intentions all accounted for in terms of the *linguistic* game of giving and asking for reasons. This view has been adopted in the domain of computational semantics and dialogue modelling by Kibble (2006a, b) among others (e.g. Matheson et al. 2000; Walton and Krabbe 1995; Singh 1999). The guiding principle behind such social, non-intentionalist explanations of communication and dialogue understanding is to replace mentalist notions such as 'belief' with public, observable practical and propositional 'commitments', in order to resolve the problems arising for dialogue models associated with the intersubjectivity of beliefs and intentions, i.e. the fact that such private mental states are not directly observable and available to the interlocutors. A further motivation arises from the fact that it has been shown that beliefs, goals and intentions underdetermine what "rational" agents will do in conversation: social obligations or conversational rules may in fact either displace beliefs or intentions as the motivation for agents' behaviour or enter as an additional explanatory factor (e.g. the (social) obligation to answer a question might displace/modify the

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 195/215

"intention" not to answer it, see, e.g. Traum and Allen (1994)). Brandom's account presents an inferentialist view of communication which seeks to replace mentalist notions with public, observable practical and propositional commitments. Under this view, commitment does not imply 'belief' in the usual sense. A speaker may publicly commit to something which she does not believe. And 'intention' can be cashed out as the undertaking of a practical commitment or a reliable disposition to respond differentially to the acknowledging of certain commitments.[8]

From our point of view, the advantage of such non-individualistic, externalist accounts (see also Millikan 1984, 2005; Burge 1986) is that, in not giving supremacy to an exclusively individualist conception of psychological processes, they break apart the presumed exhaustive dichotomy between behaviourist and mentalist accounts of meaning and behaviour (see e.g. Preston 1994) or code versus inferential models of communication (see e.g. Krauss and Fussell 1996). Instead, ascribing contents to behaviours is achieved by supra-individual social or environmental structures, e.g. conventions, "functions", embodied practices, routinisations, that act as the context that guides agents' behaviour. The mode of explanation for such behaviours then does not enforce a representational component, accessible to individual agents, that analyses such behaviours in folk-psychological mentalistic terms, to be invoked as an explanatory factor in the production and interpretation of social action or behaviour. Individual agents instead can be modelled as operating through low-level mechanistic processes (see e.g. Böckler et al. 2010) without necessary rationalisation of their actions in terms of mental state ascriptions (see e.g. Barr 2004 for the establishment of conventions and Pickering and Garrod 2004 for coordination). This view is consonant with recent results in neuroscience indicating that notions like 'intentions', 'agency', 'voluntary action' etc. can be taken as post hoc "confabulations" rather than causally efficacious (work by Benjamin Libet, John Bargh and Read Montague, for a survey see Wegner 2002): according to these results, when a thought that occurs to an individual just prior to an action is seen as consistent with that action, and no salient alternative "causes" of the action are accessible, the individual will experience conscious will and ascribe agency to themselves.

Accordingly, when examining human interaction, and more specifically dialogue, notions like intentions and beliefs may enter into common sense psychological explanations that the participants themselves can invoke and manipulate, especially when the interaction does not run smoothly. As such, they do operate as resources that interlocutors can utilise explicitly to account for their own and others' behaviour. In this sense, such notions constitute part of the metalanguage participants employ to make sense of their actions in conscious, often externalised reflections (see e.g. Heritage 1984; Mills and Gregoromichelaki 2010; Healey

[8] An intermediate position is presented by Lascarides and Asher (2009); Asher and Lascarides (2008) who also appeal to a notion of public commitment associated with dialogue moves but which they link to a parallel cognitive modelling component based on inference about private mental states (see also Traum and Allen 1994; Poesio and Traum 1997).

2008). Cognitive models that elevate such resources to causal factors in terms of plans, goals etc. either risk not doing justice to the sub-personal, low-level mechanisms that implement the epiphenomenal effects they describe, or they frame their provided explanations as competence/computational level descriptions (see e.g. Stone 2005, 2004). The stance such models take may be seen as innocuous preliminary idealisation, but this is acceptable only in the absence of either emerging internal inconsistency or alternative explanations that subsume the phenomena under more general assumptions. For example, there are well-known empirical/conceptual problems with the reduction of agent coordination in terms of Bratman's joint intentions (Searle 1990; Gold and Sugden 2007)[9]; and there are also psychological/practical puzzles in cognitive/computational implementations in that the plan recognition problem is known to be intractable in domain-independent planning (Chapman 1987).[10] But, in addition, empirical linguistic phenomena seem to escape adequate modelling in that the assumption that speakers formulate and attempt to transmit determinate meanings in conversation seems implausible when conversational data is examined. We turn to a range of such phenomena next.

2.3 *Emergent Intentions*

The fundamental role of intention recognition and the primary significance of speaker meaning in dialogue has been disputed in interactional accounts of communication where intentions, instead of assuming causal/explanatory force can be characterised as "emergent" in that the participants can be taken to jointly construct the content of the interaction (Gibbs 2001; Haugh 2008; Mills and Gregoromichelaki 2010; Mills 2011). This aspect of joint action has been explicated via the assumption of the "non-summativity of dyadic cognition" (Arundale and Good 2002; Arundale 2008; Haugh 2012; Haugh and Jaszczolt 2012) or in terms of "interactive emergence" (Clark 1997; Gibbs 2001). This view gains experimental backing through the observation of the differential performance of participants versus over-hearers in conversation (Clark and Schaefer 1987; Schober and Clark 1989) and the gradual emergence of intentional explanations in task-oriented dialogue (Mills and Gregoromichelaki 2010). Standard dialogue systems, by contrast, are serial, modular and operate on complete utterances underpinned by a speaker plan and its recognition. Typically, such models include a parser responsible for syntactic and semantic analysis, an interpretation manager, a

[9] In addition, such accounts of coordination are not general enough in that they are discontinuous with explanations of collective actions, in e.g. crowd coordination, individuals walking past each other on the sidewalk, etc.

[10] In addition, it has been argued that use of such folk-psychological constructs are culture/occasion-specific (Du Bois 1987; Duranti 1988), hence should not be seen as underpinning general cognitive abilities.

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 197/215

dialogue manager and a generation module. The output of each module is the input for another with speaking and listening seen as autonomous processes. This goes against the observation that, in ordinary conversation, utterances are shaped genuinely incrementally and "opportunistically" according to feedback by the interlocutor (as already pointed out by Clark 1996) thus genuinely engendering co-constructions of utterances, structures and meanings (see e.g. Lerner 2004). In our view, the main reason for this inadequacy in dialogue modelling are methodological assumptions justified by the competence/performance distinction, separating the grammar from the parser/generator and the pragmatic modules, with the result that the grammatical models employed lack the capability to fully manipulate and integrate partial structures in an incremental manner (for recent incremental systems see Petukhova and Bunt 2011; Poesio and Rieser 2010).

2.4 Incrementality in Processing and Split Utterances

The incrementality of on-line processing is now uncontroversial. It has been established for some considerable time now that language comprehension operates incrementally; and, standardly, psycholinguistic models assume that partial interpretations are built more or less on a word-by-word basis (see e.g. Sturt and Crocker 1996). More recently, language production has also been argued to be incremental (Kempen and Hoenkamp 1987; Levelt 1989; Ferreira 1996; Bock and Levelt 2002). Guhe (2007) further argues for the incremental conceptualisation of observed events resulting in the generation of preverbal messages in an incremental manner guiding semantic and syntactic formulation. In all the interleaving of planning, conceptual structuring of the message, syntactic structure generation and articulation, psycholinguistic incremental models assume that information is processed as it becomes available, reflecting the introspective observation that the end of a sentence is not planned when one starts to utter its beginning (see e.g. Guhe et al. 2000). In accordance with this, in dialogue, evidence for radical incrementality is provided by the fact that participants incrementally "ground" each other's contribution through *back-channel* contributions like *yeah*, *mhm*, etc. (Allen et al. 2001). In addition, as shown in (1), interlocutors clarify, repair and extend each other's utterances, even in the middle of an emergent clause (*split utterances*):

1. Context: Friends of the Earth club meeting
   A: So what is that? Is that er… booklet or something?
   B: It's a book
   C: Book
   B: Just… talking about al you know alternative
   D: On erm… renewable yeah
   B: energy really I think
   A: Yeah [BNC:D97].

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 198/215

In fact, such completions and continuations have been viewed by Herb Clark, among others, as some of the best evidence for cooperative behaviour in dialogue (Clark 1996: 238).

But even though, indeed, such joint productions demonstrate the participants' skill to collaboratively participate in communicative exchanges, this ability to take on or hand over utterances raises the problem of the status of intention-recognition within human interaction when the aim is an explicit procedural model of how such exchanges are achieved. Firstly, on the Gricean assumption that pragmatic inference in dialogue operates on the basis of reasoning based on evidence of the interlocutor's intention, delivered by establishing the semantic propositional structure licensed by the grammar, the data in (1) cannot be easily explained, except as causing serious disruptions in normal processing, hence the view of dialogue as "degenerate" language use in formal analyses. Secondly, on the assumption that communication necessarily involves recognising the propositional content intended by the speaker, there would be an expected cost for the original hearer in having to infer or guess this content before the original sentence is complete, and for the original speaker in having to modify their original intention, replacing it with that of another in order to understand what the new speaker is offering and respond to it. But, wholly against this expectation, interlocutors very straightforwardly shift out of the parsing role and into the role of producer and vice versa as though they had been in their newly adopted role all along. Indeed, it is the case that such interruptions do sometimes occur when the respondent appears to have guessed what they think was intended by the original speaker, what have been called *collaborative completions*:

2. Conversation from A and B, to C:
   A: We're going to…
   B: Bristol, where Jo lives.
3. A: Are you left or
   B: Right-handed.

However, this is not the only possibility: as (4)–(5) show, such completions by no means need to be what the original speaker actually had in mind:

4. Morse: in any case the question was
   Suspect: a VERY good question inspector [Morse, BBC radio 7].
5. Daughter: Oh here dad, a good way to get those corners out
   Dad: is to stick yer finger inside
   Daughter: well, that's one way (from Lerner 1991).

In fact, such continuations can be completely the opposite of what the original speaker might have intended as in what we will call *hostile continuations* or *devious suggestions* which are nevertheless collaboratively constructed from a grammatical point of view:

6. (A and B arguing:)
   A: In fact what this shows is
   B: that you are an idiot.
7. (A mother, B son)
   A: This afternoon first you'll do your homework, then wash the dishes and then
   B: you'll give me £10?

Furthermore, as all of (1)–(7) show, speaker changes may occur at any point in an exchange (Purver et al. 2009), even very early, as illustrated by (8), with the clarification *Chorlton?* becoming absorbed into the final in-effect collaboratively derived content:

8. A: They X-rayed me, and took a urine sample, took a blood sample. Er, the doctor
   B: Chorlton?
   A: Chorlton, mhmm, he examined me, erm, he, he said now they were on about a slide <unclear> on my heart [BNC: KPY 1005–1008].

This phenomenon has consequences for accounts of both utterance understanding and utterance production. On the one hand, incremental comprehension cannot be based primarily on guessing speaker intentions: for instance, it is not obvious why in (4)–(7), the addressee has to have guessed the original speaker's (propositional) intention/plan before they offer their continuation.[11] On the other hand, speaker intentions need not be fully-formed before production: the assumption of fully-formed propositional intentions guiding production will predict that all the cases above where the continuation is not as expected would have to involve some kind of revision or backtracking on the part of the original speaker. But this is not a necessary assumption: as long as the speaker is licensed to operate with partial structures, they can start an utterance without a fully formed intention/plan as to how it will develop (as the psycholinguistic models in any case suggest) relying on feedback from the hearer to shape their utterance (Goodwin 1979).

While core pragmatic research has largely left on one side the phenomenon of collaborative construction of utterances, the emergence of propositional contents in dialogue has been documented over many years in Conversation Analysis (CA) (see e.g. Lerner 2004). The importance of feedback in co-constructing meaning in communication has been already documented at the propositional level (the level of speech acts, 'adjacency pairs') within CA (see e.g. Schegloff 2007). However, it

[11] These are cases not addressed by DeVault et al. (2009), who otherwise offer a method for getting full interpretation as early as possible. Lascarides and Asher (2009); Asher and Lascarides (2008) also define a model of dialogue that partly sidesteps many of the issues raised in intention recognition. But, in adopting the essentially suprasentential remit of SDRT, their model does not address the step-by-step incrementality needed to model split-utterance phenomena.

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 200/215

seems here that the same processes can operate sub-propositionally, but this can be demonstrated only relatively to models that allow the incremental, sub-sentential integration of cross-speaker productions. We turn to two such models next.

3 Grammar and Dialogue

It seems to be a standard assumption that linguistic knowledge has to be modelled as providing constraints on linguistic processing (see e.g. Bosch 2008, a.o.). In this sense linguistic knowledge is (often) characterised in abstract static terms whereas linguistic processing is argued to be characterised by three indispensable features, namely: immediacy (i.e. context-dependence), incrementality, multi-modality (see Marslen-Wilson and Tyler 1980; Altmann and Steedman 1988). However, against this view, work on linguistic phenomena, e.g. ellipsis, that cross-cut monologue and dialogue, sentence and discourse, has shown that a unified story requires all these three processor properties to be included in the theory of linguistic knowledge/grammar (see, e.g. Gargett et al. 2009; Kempson et al. 2009a, b). Otherwise, separating linguistic knowledge (grammar) from processing results in a view of dialogue as "degenerate" language use. Notably, this separation has led even dialogue-oriented psycholinguists, e.g. Clark (1996), to distinguish language_S (language structure) versus language_U (language-in-use).

In contrast, here we would like to argue for a reconciliation between the "language-as-action" and "language-as-product" traditions, at the same time shifting the boundaries between grammar and pragmatics. The reason for this is that the two approaches should be seen, in our view, as constituting not a dichotomy but a continuum. However, in order to substantiate such a view, linguistic knowledge has to be reconceptualised as encompassing the update dynamics of communication which crucially involves:

- representations integrating multiple sources of information
- word-by-word incrementality within the grammar system
- NL grammars as mechanisms for communicative interaction relative to context.

This is because what we see as inherent features of the grammar architecture, utilised to solve traditional grammatical puzzles (see e.g. Kempson et al. 2001; Cann et al. 2005; Kempson et al. 2011b), also underlie many features of language use in dialogue. Firstly, the function of items like inserts, repairs, hesitation markers etc. interact with the grammar at a sub-sentential level (Clark and Fox Tree 2002). Hence the grammar must be equipped to deal with those in a timely and integrated manner. In addition, the turn-taking system (see, e.g., Sacks et al. 1974) seems to rely on the grammar, based on the predictibility of (potential) turn endings; in this respect, recent experimental evidence have shown that this predictability is grounded on syntactic recognition rather than prosodic cues etc. (De Ruiter et al. 2006); and further evidence shows that people seem to exploit such predictions to manage the timing of their contributions (Henetz and Clark 2011). More

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 201/215

importantly for our concerns here, incremental planning in production allows the grammar to account for how the interlocutors interact sub-sententially in dialogue to derive joint meanings, actions and syntactic constructions taking in multi-modal aspects of communication and feedback, a fact claimed to be a basic characteristic of interaction (Goodwin 1979, 1981).

3.1 Modelling the Incrementality of Split Utterances

The challenge of modelling the full word-by-word incrementality required in dialogue has recently been taken up by two models which employ distinct approaches: a neo-Gricean model by Poesio and Rieser (2010) (*P&R* henceforth) and Dynamic Syntax (Kempson et al. 2001).

P&R set out a dialogue model for German, defining a thorough, fine-grained account of dialogue interactivity. Their primary aim is to model *collaborative completions*, as in (2) and (3) in cooperative task-oriented dialogues where take-over by the hearer relies on the remainder of the utterance taken to be understood or inferrable from mutual knowledge/common ground.[12] Their account is an ambitious one in that it aims at modelling the generation and realisation of joint intentions which accounts for the production and comprehension of co-operative completions. The P&R model hinges on two main points: the assumption of recognition of interlocutors' intentions according to shared joint plans (Bratman 1992), and the use of incremental grammatical processing based on LTAG. With respect to the latter, this account relies on the assumption of a string-based level of syntactic analysis, for it is this which provides the top-down, predictive element allowing the incremental integration of such continuations. However, exactly this assumption would seem to impede a more general analysis, since there are cases where split utterances cannot be seen as an extension by the second contributor of the proffered string of words/sentence:

9. Eleni: Is this *yours* or
   Yo: *Yours* [natural data].
10. with smoke coming from the kitchen:
    A: I'm afraid I burnt the kitchen ceiling
    B: But have *you*
    A: burned *myself*? Fortunately not.

In (9), the string of words (sentence) that the completion yields is not at all what either participant takes themselves to have constructed, collaboratively or otherwise. And in (10) also, even though the grammar is responsible for the dependency that licenses the reflexive anaphor *myself*, the explanation for B's continuation in the third turn of (10) cannot be string-based as then *myself* would not be locally

[12] Thus, notably, the P&R data involve data collected after task training.

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 202/215

bound (its antecedent is *you*). Moreover, in LTAG, P&R's syntactic framework, parsing relies in the presence of a head that provides the skeleton of the structure. Yet, as (1)–(10) indicate, utterance take-over can take place without a head having occurred prior to the split (see also Purver et al. 2009, Howes et al. 2011), and even across split syntactic dependencies (in (10) an antecedent-anaphor relation and in (11) between a Negative Polarity Item and its triggering environment, the question):

11. A: *Have* you mended
    B: *any* of your chairs? Not yet.

Given that such dependencies are defined grammar-internally, the grammar has to be able to license such split-participant realisations. But string-based grammars cannot account straightforwardly for many types of split utterances except by treating each part as elliptical sentences requiring reconstruction of the missing content with case-specific adjustments to guarantee grammaticality/interpretability (as is needed in (9)–(10)).

Furthermore, if the attempt is to reconstruct speaker's intentions as the basis for the interpretation recovered, as P&R explicitly advocate, there is the additional problem that such fragments can play multiple roles at the same time (e.g. the fragments in (3) and (9) can be simultaneously taken as question/clarification/completion/acknowledgment/answer; see also Sbisà, this volume). Notice also that co-construction at the sub-propositional level can be employed for the performance of speech acts by establishing (syntactic) conditional relevances,[13] i.e. exploiting grammatical mechanisms as a means to induce the coordination of social actions. For example, such completions might be explicitly invited by the speaker thus forming a question–answer pair:

12. A: And you're leaving at
    B: 3.00 o'clock.
13. A: And they ignored the conspirators who were …
    B: Geoff Hoon and Patricia Hewitt [radio 4, Today programme, 06/01/10]
14. Jim: The Holy Spirit is one who <pause> gives us? Unknown: Strength
    Jim: Strength. Yes, indeed. <pause> The Holy Spirit is one who gives us? <pause>
    Unknown: Comfort. [BNC HDD: 277–282]
15. George: Cos they <unclear> they used to come in here for water and bunkers you see
    Anon 1: Water and?
    George: Bunkers, coal, they all coal furnace you see,… [BNC, H5H: 59–61]

Within the P&R model, such multifunctionality would not be capturable except as a case of ambiguity or by positing hidden constituent reconstruction that has to be

[13] For the concept of *conditional relevance* in conversation see, e.g., Schegloff (1996).

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 203/215

subject to some non-monotonic build-and-revise strategy that is able to apply even within the processing of an individual utterance. But, in fact, in some contexts, invited completions have been argued to exploit the vagueness/covertness of the speech act involved to avoid overt/intrusive elicitation of information (Ferrara 1992):

16. (Lana = client; Ralph = therapist)
    Ralph: Your sponsor before…
    Lana: was a woman

Hence, the resolution of such fragments cannot be taken to rely on the determination of a specific speaker-intended speech-act (see also Sbisà, this volume).

It has to be said that the P&R account is not intended to cover such data, as the setting for their analysis is one in which participants are assigned a collaborative task with a specific joint goal, so that joint intentionality is fixed in advance and hence anticipatory computation of interlocutors' intentions can be fully determined; but such fixed joint intentionality is decidedly non-normal in dialogue (see e.g. Mills and Gregoromichelaki 2010) and leaves any uncertainty or non-determinism in participants' intentions an open challenge. Nonetheless, by employing an incremental model of grammar, the P&R account marks a significant advance in the analysis of such phenomena. Relative to any other grammatical framework, dialogue exchanges involving incremental split utterances of any type are even harder to model, given the near-universal commitment to a static performance-independent methodology. Thus, first of all, in almost all standard grammar frameworks, it is usually the sentence/proposition that is the unit of syntactic/semantic analysis. Inevitably, fragments are then assigned sentential analyses with semantics provided through ellipsis resolution involving abstraction operations as in Dalrymple et al. (1991) (see e.g. Purver 2006; Ginzburg and Cooper 2004; Fernandez 2006). The abstraction is defined over a propositional content provided by the previous context to yield appropriate functors to apply to the fragment. Of course, multiple options of appropriate "antecedents" for elliptical fragments are usually available (one for each possible abstract) resulting in multiple ambiguities which are then relegated to some performance mechanism for resolution. Such mechanisms are defined to appeal to independent pragmatic assumptions having to do with recognizing the speaker's intention in order to select a single appropriate interpretation. But the intention recognition required for disambiguation is unavailable in sub-sentential split utterances as in (1), (3), (9)–(15) in all but the most task-specific domains. This is because, in principle, attribution of recognition of the speaker's intention to convey some specific propositional content is unavailable until the appropriate propositional formula is established. This is particularly clear where an antecedent is required too early in the emergent proposition so that no appropriate abstract definable from context is available as in (8) above.

In response to the challenge that such data provide, we turn to *Dynamic Syntax* (DS: Kempson et al. 2001; Cann et al. 2005) where the correlation between parsing

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 204/215

and generation, as they take place in dialogue, can provide a basis for modelling recovery of interpretation in communicative exchanges without reliance on recognition of specific intentional contents.

3.2 *Dynamic Syntax*

DS is an action-based formalism. It models "syntax" in procedural terms as the goal-directed, incremental, stepwise transition from strings of words to meaning representations which dynamically integrate both linguistic and extra-linguistic or inferred information. These are the only representations constructed during the interpretation of utterances, hence no distinct syntactic level of representation is assumed. As in DRT and related frameworks (see also Jaszczolt 2005), semantic, truth-conditional evaluation applies solely to these contextually enriched representations, hence no semantic content is ever assigned to strings of words (sentences).

3.2.1 Radically Contextualist Representations

The examination of linguistic data seems to indicate evidence of structure underlying the linear presentation of strings. Similar types of evidence can also be found in dialogue. First of all, it has been shown both by corpus research (Fox and Jasperson 1995) and experimental results (Eshghi et al. 2010) that repair processes in dialogue target primarily 'constituents' whereas other factors like pauses, time units etc. play a secondary role. For example, Fox and Jasperson, who examine self-repairs, claim that "in turn beginnings, if repair is initiated after an auxiliary or main verb, the verb and its subject are always recycled together; the verb is never recycled by itself." (1995:110). Moreover, the use of fragments ("elliptical" utterances) during interaction, follows syntactic constraints indicating their appropriate integration in some structured representation. This is more evident in languages with rich morphology and case systems. For example, although it has been established that speakers can use fragments like the following in (16) to perform speech acts that do not presuppose the recovery of a full sentence ('non-sentential speech acts': Stainton 2005), languages like German and Greek require that the fragment bears appropriate case specifications, otherwise it is perceived as ungrammatical:

17. Context: A and B enter a room and see a woman lying on the floor:
    A to B: Schnell, den Arzt/*der Arzt (German)
    "Quick, the doctor_ACC/*the doctor_NOM"

One might take these as evidence for a separate (possibly autonomous) level of syntactic analysis. Indeed, based on similar observations, standard grammatical

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 205/215

models postulate an independent level of structure over strings (see e.g. Ginzburg and Cooper 2004; Ginzburg 2012) whereas categorial grammars that deny the existence of any level of independent structuring with syntactic relevance have difficulty in explaining such data. Both these types of account are not sustainable as there is also evidence that explanations for such phenomena cannot be string-based. As shown below in (17) and earlier in (9)–(10), splicing together the two partial strings gives incorrect interpretations since elements like indexicals have to switch form in order to be interpretable as intended or for grammaticality:

18. G: when you say it happens for a reason, it's like, it happened to get ***you*** off
    D: off ***my*** ass [Carsales 3 cited in Ono and Thompson (1995)]

In contrast, even though DS, like categorial grammar, takes the view that syntactic constraints and dependencies do not justify a separate level of representation for structures over stings, nevertheless, it handles such data successfully via the definition of constraints on the updates of the semantic representations induced by the processing mechanism. So the reduction in representational levels, instead of impeding the definition of syntactic licensing, allows in fact the handling of a wider range of data via the same incremental licensing mechanisms. So, instead of data such as those in (9)–(10) and (17) being problematic, use of the licensing mechanisms across interlocutors illustrates the advantages of a DS-style incremental, dynamic account over static models (for detailed analyses see Kempson et al. 2009a, b, 2011a; Purver et al. 2010, 2011; Gregoromichelaki et al. 2009, 2011; Gargett et al 2008). Given that linguistic processing has to be incrementally interleaved with processes of inference and perceptual inputs, this is essential for dialogue as not only is comprehension heavily reliant on context and multimodal input but also dialogue management issues are handled by interaction of linguistic and non-linguistic resources. For example, Goodwin (1979) suggests that in face-to-face interaction completion, extension and allocation of turns are managed through a combination of gaze and syntactic information.

3.2.2 Incrementality

Because of this procedural architecture, two features usually associated with parsers, *incrementality* and *predictivity*, are intrinsic to the DS grammar and are argued to constitute the explanatory basis for many idiosyncrasies of NLs standardly taken to pose syntactic/morphosyntactic/semantic puzzles. As can be seen in (1) above, dialogue utterances are fragmentary and subsentential. This implies that dialogue phenomena like self-repair, interruptions, corrections etc. require modelling of the incremental understanding/production and if the grammar needs to license such constructions it needs to deal with partial/non-fully-sentential constructs. Modular approaches to the grammar/pragmatics interface deny that this is an appropriate strategy. Instead they propose that the grammar delivers under-specified propositional representations as input to pragmatic processes that achieve

full interpretations and discourse integration (see e.g. Schlangen 2003, following an SDRT model). However, an essential feature of language use in dialogue is the observation that on-going interaction and feedback shapes utterances and their contents (Goodwin 1981), hence it is essential that the grammar does not have to licence whole propositional units for semantic and pragmatic evaluation to take place. And this is the strategy DS adopts as it operates with partial constructs that are fully licensed and integrated in the semantic representation immediately. This has the advantage that online syntactic processing can be taken to be implicated in the licensing of fragmentary utterances spread across interlocutors without having to consider such fragments as elliptical sentences or non well-formed in any respect. And this is essential for a realistic account of dialogue as corpus research has shown that speaker/hearer exchange of roles can occur across all syntactic dependencies (Purver et al. 2009):

19. Gardener: I shall need the mattock.
    Home-owner: **The**…
    Gardener: **mattock**. For breaking up clods of earth [BNC].
20. A: or we could just haul: a:ll the skis in [[the:]] dorms
    B: [[we could]]
    [[haul all the skis into the dorm]]
    C: [[hh uh hhuhhuh]] (1.0)
    B: **which** (0.3)
    A: **might** work
    B: might be the best [BNC].
21. Jack: I just returned
    Kathy: **from**…
    Jack: Finland [from Lerner 2004]
22. Teacher: Where was this book lub- published?
    Teacher: Macmillan publishing company **in**? (.)
    Class: New York ((mostly in unison))
    Teacher: Okay, [from Lerner 2004].
23. Therapist: What kind of work do you do?
    Mother: on food service
    Therapist: **At_**
    Mother: uh post office cafeteria downtown main point office on Redwood
    Therapist: °Okay° [Jones and Beach 1995].
24. S: You know some nights I just- (0.2) if I get bad flashes I c- I can't mo:ve.
    C: No: =
    S: So some nights he's got the baby and me:huh(.)
    C: hhhh Uh by flashes you mean flashbacks
    S: Yea:h.
    C: **To**::,
    S: To- To the bi:rth

C: To the birth itse:lf. mm.(0.2)
S: And thee uhm (.) the- the labor an' thee the week in the hospital afterwa:rd[s.]
C: [Y]e:s. Ye:s. [from Lerner 2004]

But if the grammar is conceived as operating independently of the dialogue processes that manage turn handling and derivation of content across participants there is no way to account for the licensing, the formal properties and eventual interpretations of such fragmentary utterances (see also Morgan 1973). Instead, DS grammar constraints operate incrementally, on a word-by-word basis, thus allowing participants to progressively integrate contents and modify each other's contributions.

3.2.3 Predictivity

As we said earlier, the turn-taking system (see Sacks et al. 1974) relies heavily on the grammar via the notion of predictibility of (potential) turn endings. Fluent speaker/hearer role switch relies on participants' being able to monitor the ongoing turn and project constituent completions so that they can time their exits and entries appropriately. Experimental results have shown that this ability is primarily grounded on syntactic recognition (rather than prosodic clues etc. see, e.g. De Ruiter et al. 2006). The ability of recipients to project the upcoming turn completion so that they can plan their own contribution seems to favour predictive models of processing (e.g. Sturt and Lombardo 2005) over head-driven or bottom-up parsers. DS incorporates exactly such a notion of predictivity/goal-directedness inside the grammar formalism itself in that processing (and hence licensing) is driven by the generation and fulfilment of goals and subgoals. This architectural feature of DS is fully compatible with observations in interactional accounts of conversation where it is noted that 'anticipatory planning' takes place (Arundale and Good 2002). In addition, given the format of the semantic representations employed by DS (linked trees annotated with conceptual content in functor-argument format), a second stage of composition of what has been built incrementally also occurs at constituent boundaries thus giving the opportunity for 'retroactive assessment' of the derived content (as noted again by Arundale and Good 2002).

Because DS is *bidirectional*, i.e. a model of both parsing and production mechanisms that operate concurrently in a synchronized manner, its goal-directedness/predictivity applies symmetrically *both* in parsing and generation (for predictivity in production see also Demberg-Winterfors 2010). And the consequences in this domain are welcome. Given that the grammar licenses the generator to operate with partial sub-propositional objects, speakers can be modelled as starting to articulate utterances before having planned a complete proposition. Split utterances follow as an immediate consequence of these assumptions: given the general predictivity/goal-directedness of the DS architecture, the parser/

generator is always predicting top-down structural goals to be achieved in the next steps. But such goals are also what drives the search of the lexicon ('lexical access') in generation, so a hearer who achieves a successful lexical retrieval before processing the anticipated lexical input provided by the original speaker can spontaneously become the generator and take over. As seen in all cases (1)–(15) above, the original hearer is, indeed, using such anticipation to take over and offer a completion that, even though licensed, i.e. a grammatical continuation of the initial fragment, might not necessarily be identical to the one the original speaker would have accessed had they been allowed to continue their utterance as in (7)–(9). And since the original speaker is licensed to operate with partial structures, without having a fully-formed intention/plan as to how it will develop (as the psycholinguistic models in any case suggest), they can integrate immediately such offerings without having to be modelled as necessarily revising their original intended message[14] (for detailed analyses see Kempson et al. 2009a, b; Purver et al. 2010, 2011; Gregoromichelaki et al. 2009, 2011; Gargett et al 2008).

Thus DS reflects directly and explicitly, from within the grammar itself, how the possibility arises for joint-construction of utterances, meanings and structures in dialogue and how this is achieved. And these explanations are fundamentally based on the same mechanisms underlying language structure: since the grammar licenses partial, incrementally constructed objects, speakers can start an utterance without a fully formed intention/plan as to how it will develop relying on feedback from the hearer to shape its structure and its construal. Moreover, the syntactic constraints themselves can be exploited ad hoc as a source of "conditional relevances" (Schegloff 2007) by setting up sequences (joint speech acts or 'adjacency pairs') sub-sententially (see (20)–(22) above). Thus, syntactic devices and their goal-directed, projectible nature can be manipulated by interlocutors to manage conversational organisation and perform speech acts without fully-formed propositional contents.

Given these results, in our view, the dichotomy between $language_S$ (language structure) and $language_U$ (language use) postulated in standard linguistic models does not withstand the test of application in dialogue, the primary site of language use. Instead, the grammar has to be seen as underpinning communication with, as DS suggests, the syntactic architecture viewed in dynamic terms as the crystallisation of action patterns derived from language use and wider cognitive/social considerations.

4 Conclusion

With grammar mechanisms defined as inducing incremental context-dependent growth of information and employed symmetrically in both parsing and generation, the availability of derivations for genuine dialogue phenomena, like split

[14] But, of course, this is not excluded either.

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 209/215

utterances, from within the grammar, shows how core dialogue activities can take place without any other-party meta-representation at all.[15] On this view, as we emphasised earlier, communication is not at base the intention-recognising activity presumed by Gricean and post-Gricean accounts. Rather, speakers can be modelled as able to air propositional and other structures with no more than the vaguest of planning and commitments as to what they are going to say, expecting feedback to fully ground the significance of their utterance, to fully specify their intentions (see e.g. Wittgenstein 1953: 337). Hearers, similarly, do not have to reconstruct the intentions of their interlocutor as a filter on how to interpret the provided signal; instead, they are expected to provide evidence of how *they* perceive the utterance in order to arrive at a joint interpretation. This view of dialogue, though not uncontentious, is one that has been extensively argued for, under distinct assumptions, in the CA literature. According to the proposed DS model of this insight, the core mechanism is incremental, context-dependent processing, implemented by a grammar architecture that reconstructs "syntax" as a goal-directed activity, able to seamlessly integrate with the joint activities people engage in.

This then enables a new perspective on the relation between linguistic ability and the use of language, constituting a position intermediate between the philosophical stances of Millikan and Brandom, and one which is close to that of Recanati (2004). Linguistic ability is grounded in the control of (sub-personal, low-level) mechanisms (see e.g. Böckler et al. 2010) which enable the progressive construction of structured representations to pair with the overt signals of the language. The content of these representations is ascribed, negotiated and accounted for in context, via the interaction among interlocutors and their environment. From this perspective, constructing representations of the other participants' mental states, rational deliberation and planning, though a possible means of securing communication, is seen as by no means necessary.

Acknowledgments We would like to thank Matt Purver, Greg Mills, Ronnie Cann, Pat Haley and Wilfried Meyer-Viol for their various contributions in this work. We gratefully acknowledge helpful discussions with Alex Davies, Arash Eshghi, Chris Howes, Graham White, Hannes Rieser, Yo Sato and Robin Cooper. This work has been supported by The Dynamics of Conversational Dialogue (DynDial) ESRC-RES-062-23-0962.

[15] Though, of course, use of reasoning over mental states is not precluded either; such richer contexts and consequent derived implications are modelled via the construction of appropriately linked representations, whose mechanisms for construction are independently available in DS, see Purver et al (2011).

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 210/215

References

Agre, P. E., & Chapman, D. (1990). What are plans for? *Robotics and Autonomous Systems, 6*(1–2), 17–34.

Allen, J., Ferguson, G., & Stent, A. (2001). *An architecture for more realistic conversational systems. In Proceedings of the 2001 International Conference on Intelligent User Interfaces (IUI)*, January 2001.

Allott, N. E. (2005). Paul Grice, reasoning and pragmatics. *UCL Working Papers in Linguistics, 17*, 217–243.

Allott, N. E. (2008). *Pragmatics and rationality*. (PhD thesis, University of London).

Altmann, G., & Steedman, M. (1988). Interaction with context during human sentence processing. *Cognition, 30*(3), 191–238.

Anscombe, G. E. M. (1957). *Intention*. Oxford: Basil Blackwell.

Arundale, R. B. (2008). Against (Gricean) intentions at the heart of human interaction. *Intercultural Pragmatics, 5*(2), 229–258.

Arundale, Robert B., & Good D. (2002). Boundaries and sequences in studying conversation In A. Fetzer & C. Meierkord (Eds.), *Rethinking sequentiality. Linguistics meets conversational interaction* (pp. 121–150). Amsterdam: John Benjamins.

Asher, N., & Lascarides, A. (2008). Commitments, beliefs and intentions in dialogue. *Proceedings of LonDial*, 35–42.

Austin, J. L. (1975). *How to do things with words*. Oxford: Oxford University Press.

Bach, K. (1997). The semantics-pragmatics distinction: What it is and why it matters. *Linguistische Berichte, 8*(1997), 33–50.

Bach K., & Harnish R. M. (1982). *Linguistic communication and speech acts*. MIT Press.

Barr, DJ. (1998). Trouble in mind: Paralinguistic indices of effort and uncertainty in communication. In S. Santi, I. Guaitella, C. Cave & G. Konopczynski (Eds.), *Oralité et gestualité. Interactions et comportements multimodaux dans la communication* (pp. 597–600). Paris: L'Harmattan.

Barr, D. J. (2004). Establishing conventional communication systems: Is common knowledge necessary? *Cognitive Science, 28*(6), 937–962.

Becchio, C., & Bertone, C. (2004). Wittgenstein running: Neural mechanisms of collective intentionality and we-mode. *Consciousness and Cognition, 13*, 123–133.

Birch, S. A. J., & Bloom, P. (2007). The curse of knowledge in reasoning about false beliefs. *Psychological Science, 18*(5), 382.

Bock, K., & Levelt, W. (2002). Language production: Grammatical encoding. In G. T. M. Altmann (Ed.), *Psycholinguistics* (pp. 405–452). London: Routledge (Critical concepts in psychology, 5).

Böckler, A., Knoblich, G., & Sebanz, N. (2010). Socializing Cognition. *Towards a Theory of Thinking*, 233–250.

Bosch, P. (2008). Processing definite determiners. In: P. Bosch, D. Gabelaia & J Lang (Eds.) *Proceedings of the 7th International Tbilisi Symposium on Language, Logic and Computation*.

Brandom, R. B. (1994). *Making it explicit: Reasoning, representing, and discursive commitment*. Harvard University Press.

Bratman, M. E. (1990). What is intention? In P. R. Cohen, J. Morgan, & E. M. Pollack (Eds.) *Intentions in communication*. Cambridge: MIT Press.

Bratman, M. E. (1992). Shared cooperative activity. *Philosophical Review, 101*, 327–341.

Bratman, M. E. (1993). Shared intention. *Ethics, 104*, 97–113.

Bratman, M. E. (1999). *Faces of intention: Selected essays on intention and agency*. Cambridge: Cambridge University Press.

Breheny, R. (2006). Communication and folk psychology. *Mind & Language, 21*(1), 74–107.

Brennan, S. E., & Schober, M. F. (2001). How listeners compensate for disfluencies in spontaneous speech. *Journal of Memory and Language, 44*(2), 274–296.

Burge, T. (1986). Individualism and psychology. *The Philosophical Review, 95*(1), 3–45.
Cann, R., Kempson, R., & Marten, L. (2005). *The dynamics of language*. Oxford: Elsevier.
Chapman, D. (1987). Planning for conjunctive goals. *Artificial Intelligence, 32*(3), 333–377.
Chomsky, N. (1965). *Aspects of the theory of syntax*. Cambridge: MIT Press.
Clark, H. H. (1996). *Using language*. Cambridge: Cambridge University Press.
Clark, A. (1997). *Being There.* Cambridge: MIT Press.
Clark, H. H., & Fox Tree, J. E. (2002). Using uh and um in spontaneous speech. *Cognition, 84*, 73–111.
Clark, A., & Lappin, S. (2011). *Linguistic nativism and the poverty of the stimulus*. London: Wiley-Blackwell.
Clark, H. H., & Marshall, C. R. (1981). Definite reference and mutual knowledge. In A. K. Joshi, B. L. Webber, & I. A. Sag (Eds.), *Elements of discourse understanding* (pp. 10–63). Cambridge: Cambridge University Press.
Clark, H. H., & Schaefer, E. F. (1987). Collaborating on contributions to conversations. *Language and Cognitive Processes, 2*(1), 19–41.
Cohen, P. R., Morgan, J., & Pollack, M. E. (1990). *Intentions in communication*. Cambridge: MIT Press.
Csibra, G. (2008). Goal attribution to inanimate agents by 6.5-month-old infants. *Cognition, 107*(2), 705–717.
Csibra, G., & Gergely, G. (1998). The teleological origins of mentalistic action explanations: A developmental hypothesis. *Developmental Science, 1*(2), 255–259.
Dalrymple, M., Shieber, S. M., & Pereira, F. C. N. (1991). Ellipsis and higher-order unification. *Linguistics and Philosophy, 14*(4), 399–452.
Demberg-Winterfors, V. (2010). *A broad-coverage model of prediction in human sentence processing*. (PhD thesis. University of Edinburgh).
DeVault, D., Sagae, K., & Traum, DR. (2009). *Can I finish? In Proceedings of the SIGDIAL 2009 Conference*, pp. 11–20, London.
Du Bois, J. W. (1987). Meaning without intention: lessons from divination. *Papers in Pragmatics*, *1*(2).
Duranti, A. (1988). Intentions, language, and social action in a Samoan context. *Journal of Pragmatics, 12*(1), 13–33.
Engelhardt, P. E., Bailey, K. G. D., & Ferreira, F. (2006). Do speakers and listeners observe the Gricean Maxim of quantity? *Journal of Memory and Language, 54*(4), 554–573.
Eshghi, A., Healey, P. G. T., Purver M., Howes, C., Gregoromichelaki, E., & Kempson, R. (2010). *Incremental turn processing in dialogue. In Proceedings of the 16th Annual Conference on Architectures and Mechanisms for Language Processing (AmLAP)*, New York, Sep 2010.
Fernandez, R. (2006). *Non-sentential utterances in dialogue: classification, resolution and use*. (PhD thesis, King's College London, University of London).
Ferrara, K. (1992). The interactive achievement of a sentence: Joint productions in therapeutic discourse. *Discourse Processes, 15*(2), 207–228.
Ferreira, V. S. (1996). Is it better to give than to donate: syntactic flexibility in language production. *Journal of Memory and Language, 35*, 724–755.
Fox, B. A., & Jasperson, R. (1995). A syntactic exploration of repair in English conversation. In P. W. Davis (Ed.), *Descriptive and theoretical modes in the alternative linguistics* (pp. 77–134). Amsterdam: John Benjamins.
Gallagher, S. (2001). The practice of mind: Theory, simulation or primary interaction? *Journal of Consciousness Studies, 8*, 83–108.
Gargett, A., Gregoromichelaki, E., Howes, C., & Sato, Y. (2008). *Dialogue-grammar correspondence in dynamic syntax. In Proceedings of the 12th SemDial (LonDial).*
Gargett, A., Gregoromichelaki, E., Kempson, R., Purver, M. & Sato, Y. (2009). Grammar resources for modelling dialogue dynamically. *Journal of Cognitive Neurodynamics*, *3*(4).
Garrod, S., & Anderson, A. (1987). Saying what you mean in dialogue: A study in conceptual and semantic coordination. *Cognition, 27*, 181–218.

Geurts, B. (2010). *Quantity implicatures*. Cambridge: Cambridge University Press.
Gibbs, Jr., & Raymond, W. (2001). Intentions as emergent products of social interactions. In B. F. Malle, L. J. Moses, & D. A. Baldwin (Eds.), *Intentions and intentionality* (pp. 105–122). Cambridge: MIT Press.
Gilbert, M. (1996). *Living together: Rationality, sociality, and obligation*. Lanham: Rowman & Littlefield.
Gilbert, M. (2003). The structure of the social atom: Joint commitment as the foundation of human social behavior. In F. F. Schmitt (Ed.), *Socializing metaphysics: The nature of social reality* (pp. 39–64). Lanham: Rowman & Littlefield.
Ginzburg, J. (2012). *The interactive stance: Meaning for Conversation*. Oxford: Oxford University Press.
Ginzburg, J., & Cooper, R. (2004). Clarification, ellipsis, and the nature of contextual updates in dialogue. *Linguistics and Philosophy, 27*(3), 297–365.
Ginzburg, J. (2012). *The interactive stance*. Oxford University Press.
Glüer, K., & Pagin, P. (2003). Meaning theory and autistic speakers. *Mind and Language, 18*(1), 23–51.
Gold, N., & Sugden, R. (2007). Collective intentions and team agency. *Journal of Philosophy, 104*(3), 109–137.
Goodwin, C. (1979). The interactive construction of a sentence in natural conversation. In G. Psathas (Ed.), *Everyday language: Studies in ethnomethodology* (pp. 97–121). New York: Irvington Publishers.
Goodwin, C. (1981). *Conversational organization: Interaction between speakers and hearers*. New York: Academic Press.
Gregoromichelaki, E., Sato, Y., Kempson, R., Gargett, A., Howes, C. (2009). *Dialogue modelling and the remit of core grammar. In: Proceedings of IWCS 2009*.
Gregoromichelaki, E., Kempson, R., Purver, M., Mills, G. J., Cann, R., Meyer-Viol, W., Healey, PGT. (2011). Incrementality and intention-recognition in utterance processing. *Dialogue and Discourse 2*(1), 199:233, special issue on Incremental Processing in Dialogue.
Grice, P. (1975). Logic and conversation. In P. Cole and J. L. Morgan (Eds.), *Syntax and semantics, vol. 3: Speech acts* (pp. 41–58). New York: Academic Press.
Grice, P. (1981). Presupposition and implicature. In P. Cole (Ed.), *Radical* pragmatics (pp. 183–198). New York: Academic Press. Reprinted in studies in the way of words. Cambridge, MA: Harvard University Press, 1989, pp. 269–82.
Grice, P. (2001). *Aspects of reason*. Oxford: Clarendon Press.
Guhe, M. (2007). *Incremental conceptualization for language production*. New Jeresy: Lawrence Erlbaum Associates.
Guhe, M., Habel, C., & Tappe, H. (2000). Incremental event conceptualization and natural language generation in monitoring environments. In *Proceedings of the 1st international conference on natural language generation*-(Vol. 14, pp. 85–92). Association for Computational Linguistics.
Hanna, J. E., Tanenhaus, M. K., & Trueswell, J. C. (2003). The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language, 49*(1), 43–61.
Haugh, M. (2008). Intention in pragmatics. *Intercultural Pragmatics, 5*(2), 99–110.
Haugh, M. (2012). Conversational interaction. In K. M. Jaszczolt & K. Allan (Eds.), *Cambridge handbook of pragmatics*. Cambridge: Cambridge University Press.
Haugh, M., & Jaszczolt, K. M. (2012). Speaker intentions and intentionality. In K. M. Jaszczolt & K. Allan (Eds.),*Cambridge Handbook of Pragmatics* (pp.87–112). Cambridge: Cambridge University Press.
Hawkins, J. A. (2004). *Efficiency and complexity in grammars.* Oxford: Oxford University Press.
Healey, P. G. T. (1997). *Expertise or expert-ese: The emergence of task-oriented sub-languages. In P. Langley & M. G. Shafto (Eds.), Proceedings of the 19th Annual Conference of the Cognitive Science Society* (pp. 301–306). Stanford: Stanford University, Aug 1997.

Healey, P. G. T. (2008). Interactive misalignment: The role of repair in the development of group sub-languages. In R. Cooper & R. Kempson (Eds.), *Language in Flux*. College Publications.

Henetz, T., & Clark, H. H. (2011). *Managing delays in speaking*. Paper presented at the IPrA meeting, Manchester, July 2011.

Heritage, J. (1984). *Garfinkel and ethnomethodology*. Cambridge: Polity Press.

Horton, W. S., & Gerrig, R. J. (2005). Conversational common ground and memory processes in language production. *Discourse Processes, 40*(1), 1–35.

Howes, C., Purver, M., Healey, P. G. T., Mills, G. J. (2011). On incrementality in dialogue: Evidence from compound contributions dialogue and discourse. (Vol 2, No 1). Special issue on Incremental Processing in Dialogue.

Hutto, D. D. (2004). The limits of spectatorial folk psychology. *Mind and Language, 19*(5), 548–573.

Jaszczolt, K. M. (2005). *Default semantics: Foundations of a compositional theory of acts of communication.* Oxford: Oxford University Press.

Jones, C. M., & Beach, W. A. (1995). Therapists' techniques for responding to unsolicited contributions by family members. In G. H. Morris & R. J. Chenail (Eds.), *The talk of the clinic: Explorations in the analysis of medical and therapeutic discourse* (pp. 49–70). Hillsdale: Lawrence-Erlbaum Associates.

Kempen, G., & Hoenkamp, E. (1987). An incremental procedural grammar for sentence formulation. *Cognitive Science, 11*(2), 201–258.

Kempson, R., Meyer-Viol, W., & Gabbay, D. (2001). *Dynamic syntax: The flow of langage understanding*. Oxford: Blackwell.

Kempson, R., Gregoromichelaki, E., & Sato, Y. (2009a). *Incrementality, speaker/hearer switching and the disambiguation challenge. In Proceedings of European Association of Computational Linguistics (EACL Athens 2009).*

Kempson, R., Gregoromichelaki, E., Purver, M., Mills, G. J., Gargett, A., Howes, C. (2009b). *How mechanistic can accounts of interaction be? In: Proceedings of Diaholmia, the 13th workshop on the semantics and pragmatics of dialogue.*

Kempson, R., Gregoromichelaki, E., Meyer-Viol, W., Purver, M., White, G., & Cann, R. (2011a). *Natural-language syntax as procedures for interpretation: the dynamics of ellipsis construal. In Proceedings of the PRELUDE Workshop on Games, Dialogues and Interaction*. Lecture Notes in Computer Science. Springer.

Kempson, R., Gregoromichelaki, E., & Howes, C. (2011b). *The dynamics of lexical interfaces*. Stanford: CSLI Publications.

Keysar, B. (2007). Communication and miscommunication: The role of egocentric processes. *Intercultural Pragmatics, 4*(1), 71–84.

Kibble, R. (2006a). Reasoning about propositional commitments in dialogue. *Research on Language and Computation, 4*(2), 179–202.

Kibble, R. (2006b). Speech acts, commitment and multi-agent communication. *Computational and Mathematical Organisation Theory*, *12*, 127–145 Special issue on normative multi-agent systems.

Krauss, R. M., & Fussell, S. R. (1996). Social and psychological models of interpersonal communication. In Kruglanski, A W. & Higgins, E. T. (Eds.), *Social psychology: Handbook of basic principles* (655–701). Guildford.

Larsson, S. (2002). *Issue-based dialogue management*. (PhD thesis, Göteborg University). Also published as Gothenburg Monographs in Linguistics 21.

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

Lerner, G. H. (1991). On the syntax of sentences-in-progress. *Language in Society, 20*, 441–458.

Lerner, G H. (2004). Collaborative turn sequences. In G. H. Lerner (Ed.) *Conversation analysis: Studies from the first generation* (pp. 225–256). Amsterdam: John Benjamins.

Levelt, W. J. M. (1989). *Speaking: From intention to articulation*. Cambridge: MIT Press.

Levinson, S. C. (1983). *Pragmatics. Cambridge textbooks in linguistics*. Cambridge: Cambridge University Press.

Levinson, S. C. (1995). Interactional biases in human thinking. In E. N. Goody (Ed.), *Social intelligence and interaction* (pp. 221–260). Cambridge: Cambridge University Press.

Matheson, C., Poesio, M., Traum, D. (2000). *Modelling grounding and discourse obligations using update rules. In: Proceedings of NAACL* (pp. 1–8).

Marr, D. (1982). *Vision: A computational investigation into the human representation and processing of visual information*. New York: Henry Holt and Co., Inc.

Marslen-Wilson, W. D., & Tyler, L. K. (1980). The temporal structure of spoken language understanding. *Cognition, 8*, 1–71.

McDowell, J. (1980). Meaning, communication, and knowledge. In Z. van Straaten (Ed.), *Philosophical subjects: Essays presented to PF Strawson*. Oxford: Clarendon Press.

Millikan, R. G. (1984). *Language, thought and other biological categories*. Cambridge: MIT Press.

Millikan, R. G. (2005). *Language: A biological model*. USA: Oxford University Press.

Mills, G. J. (2007). Semantic co-ordination in dialogue: the role of direct interaction. PhD thesis, Queen Mary University of London.

Mills, G. J. (2011). *The emergence of procedural conventions in dialogue. In: Proceedings of the 33rd Annual Conference of the Cognitive Science Society*. Boston.

Mills, G. J. & Gregoromichelaki, E. (2008/in prep.) *Coordinating on joint projects*. Paper based on talk given at the Coordination of Agents Workshop, Nov 2008, KCL.

Mills, G. J. & Gregoromichelaki, E. (2010). *Establishing coherence in dialogue: sequentiality, intentions and negotiation. In Proceedings of the 14th SemDial*, PozDial.

Morgan, J. (1973). Sentence fragments and the notion 'sentence'. In B. Jashru, R. Lees, Y. Malkiel, A. Pietrangeli, & S. Saporta (Eds.), *Issues in linguistics* (pp. 719–751). Urbana: University of Illinois Press.

Newmeyer, F. J. (2010). Formalism and functionalism in linguistics. *Wiley Interdisciplinary Reviews*: *Cognitive Science*, *1*(3), 301–307, May/June 2010.

Ono, T., & Thompson, S. A. (1995). What can conversation tell us about syntax? In P. W. Davis (Ed.), *Descriptive and theoretical modes in the alternative linguistics* (pp. 213–271). Amsterdam: John Benjamins.

Perner, J. (1991). *Understanding the representational mind*. Cambridge: MIT Press.

Petukhova, V., & Bunt, H. (2011). *Incremental dialogue act understanding. In: IWCS '11 Proceedings of the 9th International Conference on Computational Semantics*.

Phillips, C. (1996). Order and structure. PhD dissertation, Massachusetts Institute of Technology.

Pickering, M. J., & Garrod, S. (2004). Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences, 27*, 169–226.

Poesio, M., & Rieser, H. (2010). Completions, coordination, and alignment in dialogue. *Dialogue and Discourse, 1*(1), 1–89.

Poesio, M. & Traum, D. R. (1997). Conversational actions and discourse situations. *Computational Intelligence*, *13*(3), 309–347.

Preston, B. (1994). Behaviorism and mentalism: Is there a third alternative? *Synthese, 100*(2), 167–196.

Purver, M. (2006). The theory and use of clarification requests in dialogue. (PhD thesis, University of London).

Purver, M., Howes, C., Gregoromichelaki, E., & Healey, P. G. T. (2009). *Split utterances in dialogue: a corpus study. In: Proceedings of SigDial London*, Sep 2009.

Purver, M., Gregoromichelaki, E., Meyer-Viol, W., & Cann, R. (2010). *Splitting the i's and crossing the you's: Context, speech acts and grammar. In Proceedings of SemDial 2010 (PozDial)*, Poznan, June 2010.

Purver, M., Eshghi, A., & Hough, J. (2011). *Incremental semantic construction in a dialogue system. In: Proceedings of 9th International Conference on Computational Semantics (IWCS)*, Oxford, Jan 2011.

Recanati, F. (2004). *Literal meaning*. Cambridge: Cambridge University Press.

Ruiter, De., Jan-Peter, H. M., & Enfield, N. J. (2006). Projecting the end of a speakers turn: A cognitive cornerstone of conversation. *Language, 82*(3), 515–535.

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

Saxton, M. (1997). The contrast theory of negative input. *Journal of Child Language, 24*(01), 139–161.

Schegloff, E. A. (1979). The relevance of repair to syntax-for-conversation. *Syntax and semantics, 12*, 261–286.

Schegloff, E. A. (1996). Turn organization: One intersection of grammar and interaction. In E. Ochs, E. A. Schegloff, & S. A. Thompson (Eds.), *Interaction and grammar* (pp. 52–133). Cambridge: Cambridge University Press.

Schegloff, E. A. (2007). *Sequence organization in interaction: A primer in conversation analysis*. Cambridge: Cambridge University Press.

Schiffer, S. R. (1972). *Meaning*. USA: Oxford University Press.

Schlangen, D. (2003). A coherence-based approach to the interpretation of non-sentential utterances in dialogue. (PhD thesis, University of Edinburgh).

Schober, M. F., & Clark, H. H. (1989). Understanding by addressees and overhears. *Cognitive Psychology, 21*, 211–232.

Searle, J. R. (1969). *Speech acts, an essay in the philosophy of language*. Cambridge: Cambridge University Press.

Searle, J. R. (1983). *Intentionality: An essay in the philosophy of mind*. Cambridge: Cambridge University Press.

Searle, J. R. (1990). Collective intentions and actions. In P. R. Cohen, J. Morgan, & M. E. Pollack (Eds.), *Intentions in communication* (pp. 401–415). Cambridge: MIT Press.

Searle, John. R. (1995). *The social construction of reality*. New York: Free Press.

Singh, M. P. (1999). *A social semantics for agent communication languages. In: Proceedings IJCAI'99 Workshop on Agent Communication Languages* (pp. 75–88). (Issues in agent communication lecture notes in computer science, 2000, Vol. 1916/2000, pp. 31–45, DOI: 10.1007/10722777_3).

Sperber, Dan., & Wilson, Deirdre. (1995). *Relevance: Communication and cognition*. Oxford: Blackwell. 2nd edn.

Stainton, R. J. (2005). In defense of non-sentential assertion. In Z. G. Szabo (Ed.), *Semantics versus Pragmatics* (pp. 383–457). Oxford: OUP.

Stone, M. (2004). Intention, interpretation and the computational structure of language. *Cognitive Science, 28*(5), 781–809.

Stone, M. (2005). Communicative intentions and conversational processes in human–human and human-computer dialogue. In J. C. Trueswell & M. K. Tanenhaus (Eds.), *Approaches to studying world-situated language use* (pp. 39–70). Cambridge: MIT Press.

Strawson, P. F. (1964). Intention and convention in speech acts. *The Philosophical Review, 73*(4), 439–460.

Sturt, P., & Crocker, M. W. (1996). Monotonic syntactic processing: A cross-linguistic study of attachment and reanalysis. *Language and Cognitive Processes, 11*, 448–494.

Sturt, P., & Lombardo, V. (2005). Processing coordinate structures: Incrementality and connectedness. *Cognitive Science, 29*, 291–305.

Suchman, L. A. (1987/2007). *Plans and situated actions: The problem of human-machine communication*. Cambridge: Cambridge University Press.

Tollefsen, D. (2005). Let's pretend! Children and joint action. *Philosophy of the Social Sciences, 35*, 75–97.

Tomasello, M. (2008). *Origins of human communication*. Cambridge: MIT Press.

Traum, D. R. & Allen, J. F. (1994). Towards a formal theory of repair in plan execution and plan recognition. *Proceedings of UK planning and scheduling special interest group.*

Traum, D. R., Marsella, S. C., Gratch, J., Lee, J., & Hartholt, A. (2008). *Multi-party, multi-issue, multi-strategy negotiation for multi-modal virtual agents. In 8th International Conference on Intelligent Virtual Agents.*

Tuomela, R. (1995). *The importance of us: A philosophical study of basic social notions*. Palo Alto: Stanford University Press.

Layout: T1 Standard SC | Book ID: 315597_1_En | Book ISBN: 978-3-319-01013-7
Chapter No.: 7 | Date: 1-10-2013 | Page: 216/215

Tuomela, R. (2005). We-intentions revisited. *Philosophical Studies, 125*(3), 327–369.
Tuomela, R. (2007). *The philosophy of sociality: The shared point of view*. Oxford: Oxford University Press.
Walton, D. N., & Krabbe, E. C. W. (1995). *Commitment in dialogue*. Albany: State University of New York Press.
Wegner, D. M. (2002). *The illusion of conscious will*. Cambridge: MIT Press.
Wellman, H. M., Cross, D., & Watson, J. (2001). Meta-analysis of theory-of-mind development: The truth about false belief. *Child Development, 72*(3), 655–684.
Wharton, T. (2003). *Pragmatics and the showing–saying distinction*. (PhD thesis, UCL).
Wittgenstein, L. (1953). *Philosophical investigations, trans GEM Anscombe*. Oxford: Blackwell.