

Eleni Gregoromichelaki, Ronnie Cann and Ruth Kempson  
**Language as tools for interaction: Grammar  
and the dynamics of ellipsis resolution<sup>1</sup>**

**Abstract:** In this paper we explore the consequences of defining natural-language syntax as procedures for context-dependent interpretation as argued within Dynamic Syntax. Dynamic Syntax is a formalism where both representations of content and context are defined dynamically and structurally, with time-linear growth across sequences of partial trees as the core notion, presumed to be manipulated by both speaker and hearer alike. The challenge of modelling the notorious heterogeneity of ellipsis effects in conversational dialogue is taken as a case study for establishing the appropriate concept of context. In the light of the success of these assumptions in capturing the diversity of ellipsis within an overall explanation, we argue that foundational assumptions like the competence/performance distinction as regards the nature of linguistic knowledge need to be reconsidered. We argue that the grammar needs to be defined in action-based terms with incrementality and context-dependence as crucial explanatory notions of syntactic/semantic phenomena. As a consequence of taking this perspective, we argue that success in communication may rest in the ongoing interaction between participants, as underpinned by low-level mechanisms of language use like the grammar, rather than in recognising some intended propositional content.

**Keywords:** dialogue ellipsis; Dynamic Syntax; intentions

---

**Dr Eleni Gregoromichelaki:** Philosophy Department, King's College London.

E-mail: eleni.gregor@kcl.ac.uk

**Professor Ronnie Cann:** Linguistics and English Language, School of Philosophy, Psychology and Language Sciences, University of Edinburgh. E-mail: R.Cann@ed.ac.uk

**Professor Ruth Kempson:** Philosophy Department, King's College London.

E-mail: ruth.kempson@kcl.ac.uk

---

<sup>1</sup> This paper reports work done under the ESRC project RES-062-23-0962, *The Dynamics of Conversational Dialogue*, and Leverhulme Trust Major Research Fellowship F00158BF, who we thank for financial support. For comments we are grateful to Arash Eshghi, Patrick Healey, Chris Howes, Wilfried Meyer-Viol, Gregory Mills, Matt Purver and Graham White.

# 1 Introduction: Context and the nature of natural-language grammars

Readdressing the question of what assumptions to make about the grammar has become urgent, since, despite all the work done by semanticists and pragmatists in emphasising just how important context is to language processing, we still have no clear understanding of what an appropriate notion of context consists in. *Ellipsis* provides striking illustration. Arguably it might be expected to constitute a window on context, given the informal observation that ellipsis is where things can be left out because context enables the missing parts to be filled in. Yet, no such uniform explanation is in sight: to the contrary, ellipsis is taken to be “fractally heterogeneous” (Ginzburg and Cooper 2004; Merchant forthcoming). In our view, this is because current sentence-based grammars bifurcate ellipsis, and all context-dependent phenomena, into those which are sentence-bound, and hence explicable grammar-internally, and those which are not, despite the apparent similarities between both types. Data deemed to be outside the remit of standard grammatical frameworks, along with the ambiguities generated by the multiple forms of resolutions allowed by the grammar, are then presumed to fall under some performance-based account. This paper shows that such a view does not survive the challenge posed by the data of ellipsis in conversational dialogue where neither standard Gricean performance models nor competence-based grammatical analyses appear adequate to provide the requisite uniform accounts. We take these in order below, exploring first the general problems which current Gricean theoretical assumptions are increasingly facing as accounts of processing in dialogue. We then turn to the data of ellipsis in conversational dialogue as the catalyst for exploring a set of alternatives.

## 1.1 Challenges to Gricean programmes for pragmatics

The theoretical hurdle to providing a unitary explanation for some requisite concept of context is the now familiar discrepancy between representations delivered by the grammar, i.e. syntax/semantics mappings (“sentence meanings” or encoded content), and “speaker meanings” (conveyed content). This led to the Gricean account of meaning<sub>NN</sub>, (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, among

others).<sup>2</sup> From that perspective, it has been seen as necessary that, beyond some modular linguistic knowledge, itself inadequate to deliver semantically-evaluable objects, communication should essentially involve notions of rationality and cooperation. In certain versions, this is interpreted as the requirement that communication must be guided by reasoning about mental states. Under some definitions, speakers must, in order to communicate, have (at least) fourth order thoughts and hearers must recover the speaker's meaning through reasoning about these thoughts. At least then for some forms of ellipsis ('pragmatic ellipsis', Stainton 2006), the most prevalent in conversational data, this threatens to involve a model that relies heavily in mind-reading in order to resolve the gap between the fragment uttered and the speech act performed (see e.g. Poesio and Rieser 2010).

In contrast to mindreading accounts of communication, Millikan (1984: ch. 3) argues that the Gricean conception of meaning<sub>NN</sub> in ruling out causal effects on the audience does not capture ordinary linguistic understanding, which is not at basis inferentially mediated. The alternative that she proposes, 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).<sup>3</sup> Objects of ordinary perception, e.g. vision, are no less abstract than linguistic meanings, both requiring contextual enrichment in order to be comprehended. Yet, in the case of ordinary perception, this processing does not require any consideration of someone'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 along a continuum with natural meaning based on the *function* that linguistic devices have been selected to perform (their survival value). 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 but without representing hearers' mental states or having intentions about hearers' mental states (see also Csibra

---

<sup>2</sup> 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.

<sup>3</sup> The strict dichotomy between "meaning<sub>NN</sub>" and "showing" has also been disputed within Relevance Theory (see, e.g., Wharton 2003).

and Gergely 1998). Similarly, hearers understand speech through direct perception of what the speech is about without necessary reflection on speaker intentions.<sup>4</sup>

In addition, 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 the 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 medium of conversational dialogue, these results appear to show that at least communication with and by children cannot rely on higher-order intention recognition (see also Tollefsen 2005).

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. This is corroborated by an increasingly large body of research demonstrating that the Gricean notion of "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, among others).<sup>5</sup> Given this type of observation, checking in parsing or producing utterances that information is jointly held by the dialogue participants – the perceived *common ground* – cannot be a necessary condition on such activities. And there is psycholinguistic evidence that such neglect of common ground does not significantly

---

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 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).

impede successful communication and is not even detected by participants (Engelhardt et al. 2006, among others).

### 1.1.1 Re-thinking the Gricean view as a basis for dialogue modelling

As the Gricean individualistic view that speaker's intention is the sole determinant of meaning underestimates the role of the hearer, current dialogue models have turned to work in philosophy that has begun to explore notions of joint agency/joint action/joint intentions (see e.g. Bratman 1990, 1999; Searle 1990; Gilbert 2003, among others). Bratman's account of *joint intentions* has been taken as the basis of modelling participant coordination (see, e.g., Poesio and Rieser 2010). Here, the controversial notion of 'intention' as a psychological state is underpinned by 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*) (see, e.g., Clark 1996: 59). On such a view, a strong Gricean element reflecting reasoning about speakers' intentions continues to underlie the psycholinguistic and computational modelling of dialogue even though it is 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 in linguistic processing to interaction and coordination in dialogue does not mark a radically new perspective but instead presents joint action as derivative.

However, joint action seems to involve a number of lower-level cognitive phenomena that cannot be easily explicated in Gricean terms. We 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 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 & Garrod 2004) or emotional, sensory-motor practices that are, crucially, nonconceptual (Gallagher 2001: 81; Hutto 2004) and, therefore, not accountable through the high-level logical structures that Gricean explanations require.

### 1.1.2 Joint action and coordination in dialogue: empirical issues

Running alongside this work have been Conversational Analysis accounts of the sequential coherence of conversations which emphasise the importance of the turn-by-turn organisation of dialogue in allowing juxtaposition of displays of participant understandings and providing 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. In similar spirit, 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 (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. 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 performance (see e.g. Suchman 1987/2007; Agre and Chapman 1990).

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 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 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 de-

scriptions (see e.g. Stone 2005). 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);<sup>6</sup> 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; Levinson 1995). But, in addition, linguistic phenomena seem to escape adequate explanation. We turn to a range of such phenomena next where the operation of a low-level mechanism like the grammar seems to be crucially implicated in the explanation but where appropriate modelling of the phenomena cannot be accounted for grammar-internally given the Gricean assumption that speakers formulate and attempt to transmit determinate meanings in conversation.

## 1.2 Dialogue and interactive structure-building

Amongst the most striking empirical evidence of the problems caused by the bifurcation of sentence- and discourse-based analyses of ellipsis are the data of conversational dialogue. In dialogue, utterances are often collaboratively constructed, with what is said by individual contributors being highly elliptical because they rely on the context in which the conversation takes place for their interpretation:

(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]

---

<sup>6</sup> 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.





Dalrymple et al. (1991) ellipsis resolution algorithm are not applicable without further assumptions:

- (7) A: They X-rayed me, and took a urine sample, took a blood sample. Er, Mary  
 B: Your sister?  
 A: My sister, she insists. . . . .

There is the further problem that the participants may disagree, so that the follow on is not at all an attempt at collaborative completion:

- (8) (A and B arguing:)  
 A: In fact what this shows is  
 B: that you are completely wrong.

Moreover, participants' intentions may only emerge/develop during the exchange, so cannot be intrinsic to all communicative understanding, contra all standard assumptions about communication made by pragmatists (Grice 1975, Sperber and Wilson 1986/1995 etc.):

- (9) A: Oh. They don't mean us to be friends, you see. So if we want to be . . .  
 B: which we do  
 A: then we must keep it a secret.
- (10) (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?

In short, the phenomenon of context-dependence as posed by the data of conversational dialogue is highly problematic for all orthodox assumptions. In this paper we explore the extent to which the intrinsically dynamic perspective of *Dynamic Syntax* gets a better handle on these data, and in particular these *split-utterances*.

## 2 Dynamic Syntax: Addressing the context challenge

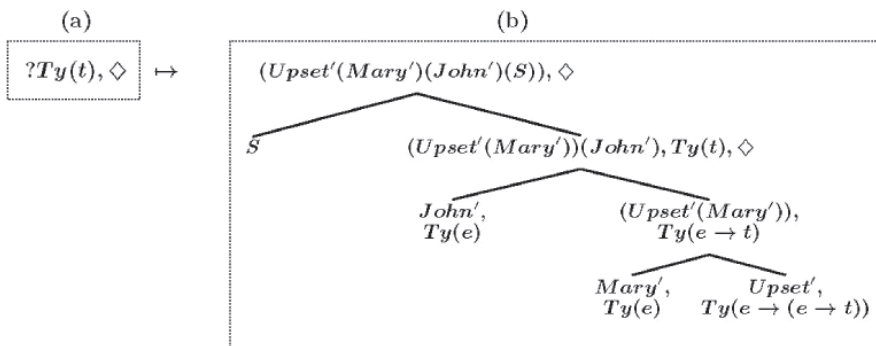
Dynamic Syntax (henceforth DS) is an action-based model of processing of which the core notion is incremental information growth following the time-linear flow of parsing/generation. Utterance contents, represented as binary tree-structures of predicate-argument form, are built up relative to context, and, similarly to

performance accounts, central to articulating this concept of growth are concepts of underspecification and update. There is a range of types of underspecification: of content, of structure, and of structural relations. Indeed, this process of building up structure is what is taken to constitute natural-language syntax: syntactic mechanisms are procedures that define how parts of representations of content can be incrementally introduced and updated. The bonus of such explicit adoption of a perspective in which the grammar reflects key properties of the dynamics of how language processing takes place is, as we shall see, that we have a natural basis for an integrated account of ellipsis.

## 2.1 Tree structures and tree-growth processes

The general process of parsing is taken to involve building as output a tree whose nodes reflect the content of some utterance – in the simple case of a sentence uttered in isolation, a complete propositional formula. The input to this task, in such a simple case, is a tree that does nothing more than state at the root node the “goal” of the interpretation process to be achieved, namely, to establish some propositional formula ( $?Ty(t)$  in (11a) below, goals are represented with ? in front of annotations). For example, in the parse of the string *John upset Mary*, the output tree in (11) to the right of the  $\mapsto$  constitutes some final end result: it is a tree in which the propositional formula itself annotates the root node, and its various subterms appear on the dominated nodes rather like a proof tree in which all the nodes are labelled with a formula and a type:

(11) John upset Mary.



These DS trees are invariably binary, and, by convention, the argument always appears on the left branch, and the functor on the right branch (a pointer,  $\diamond$ , identifies the node under development). Each node in a complete tree is anno-



example, each selected strategy for update as shown in (12), has to be checked for subsumption with respect to the goal tree  $T_g$  in step [4] representing the content to be conveyed. So parsers and producers are modelled as performing incremental synchronised actions in that they use complementary but mutually available strategies for building up representations of content in context, either to establish interpretation for a sequence of words, or to find words which match the content to be conveyed. In this process, there may be partial specifications of trees (as in the non-final steps of (12), partial specifications of formulae (as for anaphoric expressions, which project but a place-holding metavariable, e.g. **U**, **V**, etc. to be updated from elsewhere), and even partial specifications of tree relations (long-distance dependencies, for example, involving initial projection of some “unfixed” node needing to be resolved later in the incremental derivation). Such incomplete specifications, across all parameters, are associated with goals, i.e. requirements for update ( $?X$ ) that drive the tree growth process (see Cann et al. 2005 for details).

## 2.2 Towards a dynamic concept of context

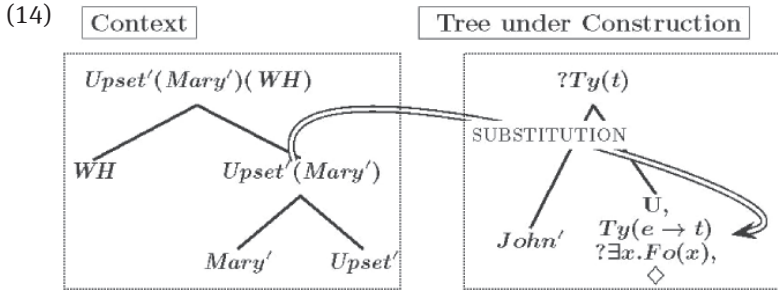
Turning now to what ellipsis data indicate, we will find that we have to see ‘semantics’ as structured representations of content, ‘syntax’ as the set of actions for constructing these representations, and ‘context’ as a store of content, i.e. structures, plus the actions involved, hence the semantic representations and the syntactic process combined. For with these concepts in place, we anticipate that hearers/speakers can retrieve actions and content stored in context and re-use them both to build up interpretation irrespective of which role they had previously been adopting; and a principled general account of ellipsis follows immediately. More specifically, the types of information that ellipsis can pick up from context are of three basic sorts: content derived by some antecedent utterance (or extralinguistic information), the structure made available by some antecedent utterance, and the processes made use of by some previous utterance. We take each in turn.

### 2.2.1 Context-dependent processing (a): Recovery of content

The type of ellipsis construal familiar from the linguistics literature is that ellipsis can select terms from (linguistic) context:

- (13) Q: Who upset Mary?  
 Ans: John did. (strict readings)

We display this as follows:



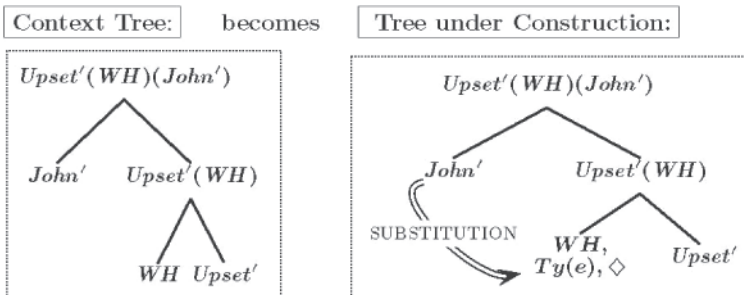
The formulation of this is exactly analogous to pronominal anaphora, though for some predicate type. The encoded expression projects a metavariable, **U**, of type  $(e \rightarrow t)$ , which then acts as trigger to license recovery of some appropriate content from the immediate context. It is notable that this commits us to the correct observation that ellipsis can be interpreted relative to the non-linguistic, indeed non-verbal, context if the predicate is sufficiently salient (contra Hankamer and Sag 1976 and others):

- (15) (Parent to teenage son with surf-board standing in shallows:  
 I *wouldn't* if I were you. The flag's flying, so it'll be dangerous

**2.2.2 Context-dependent processing (b): Re-use of structure**

In addition, the very structure made available by the immediate context can also be used, in some sense directly, with the respondent presuming on its availability as the point of departure for their own utterance. This is characteristic of question-answer pairs in particular:

- (16) Q: Who did John upset?  
 Ans: Himself.



As the above display shows, the structure is in some sense shared between speaker and hearer in these exchanges, and the question is what that amounts to. In this case, this structure contains a specialised metavariable, **WH**, which the *wh*-expression is defined to provide as a place-holding device for a subsequent answer (see Kempson et al. 2001 ch. 5). In the reply this metavariable is replaced through the processing of a reflexive pronoun, which duly has to be identified as the term inhabiting the subject node. Hence the interpretation of the fragment as providing an answer to the question, by update of the very structure which it itself provides. As we shall very shortly see, it is this hand-over of structure which is characteristic of the split-utterance phenomenon.

### 2.2.3 Context-dependent processing (c): Re-using actions from context

It is not merely the output structure or content which interlocutors can make use of but also the actions which each of them may severally use in building up such representations, for it is such re-use which enables patterns of structure-building to be replicated without strict identity of the resulting content:

- (17) A: Who hit himself?  
 B: John did. Bill too.

In (17), as in (13), B's (initial) answer involves the predicate-anaphoric device, the verb *do*, which constitutes a trigger to establish some appropriate predicate from what context provides. But in this case, it is actions recovered from context which yield that value (rather than some given predicate-formula as in the strict interpretation). What is recovered are the actions lexically encoded in the verb *hit* immediately followed by the actions encoded in the reflexive pronoun. As applied in the new environment provided by the ellipsis site – *John did* – this sequence of actions will again involve the projection of the two-place predicate *hit*' and a place-holding metavariable at the object node, but the subsequent local identification of that metavariable as subject will, in this new environment, ensure that the object node decoration will be *John*'. The result is the 'sloppy interpretation' 'John hit himself'. A similar pattern of reiterated actions can be applied also in interpreting the following add-on of *Bill too* – all without any repeat of the word *himself* to trigger such distinct local bindings. It is simply the selection of an action sequence, as indicated by the type-specification of the trigger, which are taken over from context and reiterated.

This re-use of actions is what underpins the broad array of ‘sloppy readings’ that linguists have identified:

(18) John upset his mother. Harry too.

(19) The man [who arrested *John*] failed to read *him his* rights. The man [who arrested Tom] did too.

In all these cases, it is not the output content which is replicated, or needs manipulation in order to create some novel content: it is simply reiteration of actions stored in the evolving context of the emergent dialogue.

*Context*, thus, is defined in DS as involving both the sequence of partial trees and the formulae that annotate them, whether linguistically derived or not; and, crucially, the action sequences underpinning the transitions between them. With this rich concept of context, ellipsis can indeed be seen as the recovery from context of information of these various types (for detailed analyses see Kempson et al. 2011).

## 2.3 Context-dependent processing: Co-construction in dialogue

The account of context as defined has, as a bonus, the prediction of seamless switching between speaker and hearer roles that is diagnostic of conversational dialogue. Unlike other frameworks, for which such split utterances pose very considerable problems, under DS assumptions, this phenomenon is predicted to be wholly straightforward, indeed their existence is a consequence of the DS account of production (Purver et al. 2006). According to DS, the very same mechanisms are used in production as in parsing, hence, tight coordination between the parties is expected. With each party modelled as building up structure incrementally relative to context, at any point, they can switch roles making use of that individually constructed representation whether as parser or producer. So, even within a role transition and the first and second person pronouns switching interpretation, the mechanism for processing them remains identical (see Purver et al. 2010):

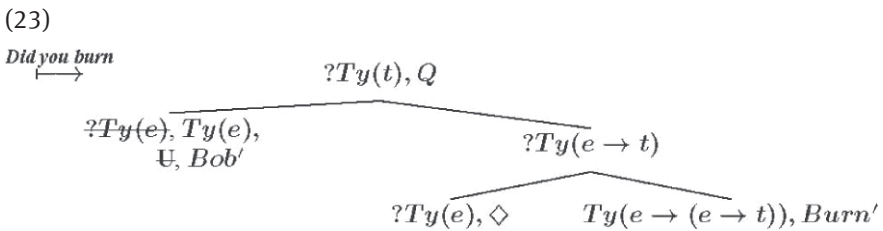
(20) A: Did **you** give me back  
B: **your** penknife? It’s on the table.

(21) A: I heard a shout. Did **you**  
B: Burn **myself**? No, luckily.

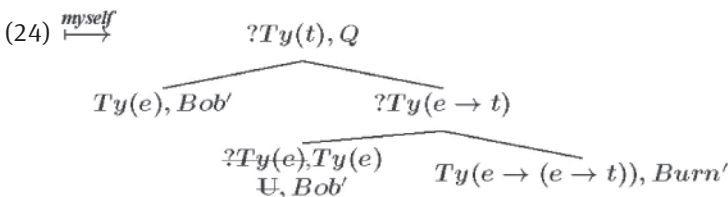
As these display, the context used by a participant as a hearer is exactly that of the context they use in their shifted role as a speaker and vice versa. By way of illustration, we take a simplified variant of (21):

- (22) Ann: Did you burn  
Bob: myself?

Here, of course, the reconstruction of the string as *\*Did you burn myself?* is unacceptable (at least with a reflexive reading of *myself*), illustrating the problem for purely syntactic accounts of split utterances. But under DS assumptions, with representations only of structured content, not of putative structure over strings of words, the switch of person is entirely straightforward. Consider the partial tree induced by parsing Ann's utterance *Did you burn* which involves a substitution of the metavariable (**U**) projected by *you* by the constant standing for the addressee/parser (*Bob'*):



At this point, Bob can complete the utterance with the reflexive as what such an expression does, by lexical definition, is copy a formula from a local co-argument node onto the current node, just in case that formula satisfies the conditions set by the person, number and, here, participant role of the uttered reflexive. So, in this case, the restriction is that the metavariable stands for a local co-argument that is *currently* the speaker:



Hence the absence of a 'syntactic' level of representation distinct from that of semantic representations allows the direct successful integration of such frag-



ments through the grammatical mechanisms themselves, rather than necessitating their analysis as sentential ellipsis (for detailed analyses see Kempson et al. 2011; Purver et al. 2010).

We see then that the incorporation of the dynamics of structure-building in the grammar itself allows the required licensing of a considerably larger dataset. The split utterances, so signally ignored in accounts of ellipsis that purport to be a subpart of sentence-based grammars become core data, relative to which competing grammars can be compared (see also Poesio and Rieser 2010; Ginzburg 2012).

### 3 Grammars for linguistic processing and communication

With this sketch of ellipsis and the attendant concept of context, we now turn finally to what the concept of *grammar* that the DS system provides suggests about the nature of communication itself. In our view, the data presented here, as well as further dialogue phenomena such as self-corrections, restarts, backchannels, cannot be accounted for in a principled way unless foundational assumptions about the nature of grammar, and linguistic knowledge in general, are reconsidered. All such phenomena require syntactic/semantic licensing but at a subsentential level and display radical context dependence. In order to provide a principled explanation then the grammar is defined in terms of mechanisms for word-by-word incremental parsing and production in context. ‘Syntax’ constitutes constraints on tree growth, with lexical specifications inducing actions for tree growth. There is considerable significance in this shift of perspective. No concept of sentence meaning is directly reconstructed. The system is, rather, a set of mechanisms for incrementally building content in context. An immediate consequence is that the system is unencapsulated. Crucially, actions of tree-building induced by grammatical and lexical processes may be interspersed with pragmatic actions of substitution and enrichment as driven by requirements associated with underspecified input.

There is a more radical consequence than this, as the structure-building activity intrinsic to language use does not have to be modelled as reliant on mind-reading. Speakers may, that is, start off without fixed intentions, contribute without completing any fixed propositional content, and rely on others to complete the initiated structure, and so on. The potential of modelling these assumptions seems compatible both with psycholinguistic observations (Levelt 1989) and empirical analyses of dialogue (Goodwin 1979).

## 4 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 2008; 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 vs. over-hearers in conversation (Schober and Clark 1989) and the gradual emergence of intentional explanations in task-oriented dialogue (Mills and Gregoromichelaki 2010). Standard dialogue models, by contrast, are serial, modular and operate on complete utterances underpinned by a speaker plan and its recognition. This goes against the observation that, in ordinary conversation, utterances are shaped incrementally and ‘opportunistically’ according to feedback by the interlocutor 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 a recent incremental system see e.g. Poesio and Rieser 2010). In contrast, the view presented here can genuinely model co-construction in dialogue even in cases where, as argued in Gregoromichelaki et al (forthcoming), the grammar itself is the sole mechanism underpinning the implicit performance of the speech act involved. In such cases, namely cases like (4) and (6), it is the set up of grammatical dependencies and their fulfilment that drives the interaction rather than (meta)representation of speaker intended speech acts.

## 5 Conclusion

The notion of emergent intentionality that can be modelled through an unencapsulated, incremental grammar that incorporates the dynamics of interaction is consonant with non-individualistic accounts of meaning (see e.g. Millikan 1984; Burge 1986). The advantage of such accounts 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 vs. 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 rationalisations rather than causally efficacious (for a survey see Wegner 2002).

Accordingly, we view coordination in dialogue as grounded in the control of (sub-personal, low-level) mechanisms like the “grammar”, 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.

## References

- Agre, Philip E. & David Chapman. 1990. What are plans for? *Robotics and Autonomous Systems* 6(1/2): 17–34.
- Allott, Nicholas E. 2008. *Pragmatics and rationality*. London: UCL PhD dissertation.
- Arundale, Robert B. 2008. Against (Gricean) intentions at the heart of human interaction. *Intercultural Pragmatics*, 5(2): 229–258.
- Arundale, Robert B. & David Good. 2002. Boundaries and sequences in studying conversation. In Anita Fetzer & Christiane Meierkord (eds.), *Rethinking sequentiality: Linguistics meets conversational interaction*, 121–150. Amsterdam: John Benjamins.
- Bach, Kent. 1997. The semantics-pragmatics distinction: What it is and why it matters. *Linguistische Berichte* 8: 33–50.
- Bach, Kent & Robert M. Harnish. 1982. *Linguistic communication and speech acts*. Cambridge, MA: MIT Press.

- Barr, Dale J. 2004. Establishing conventional communication systems: Is common knowledge necessary? *Cognitive Science* 28(6): 937–962.
- Böckler, Anne, Günther Knoblich & Natalie Sebanz. 2010. Socializing Cognition. In Britt Glatzeder, Vinod Goel & Albrecht Müller. *Towards a theory of thinking: Building blocks for a conceptual framework*, 233–250. Berlin: Springer.
- Bratman, Michael E. 1990. What is intention? In Philip R. Cohen, Jerry Morgan & Martha E. Pollack (eds.), *Intentions in communication*, 15–32. Cambridge, MA: MIT Press.
- Bratman, Michael E. 1999. *Faces of intention: Selected essays on intention and agency*. New York: Cambridge University Press.
- Burge, Tyler. 1986. Individualism and psychology. *The Philosophical Review* 95(1): 3–45.
- Cann, Ronnie, Ruth Kempson & Lutz Marten. 2005. *The dynamics of language*. Oxford: Elsevier.
- Chapman, David. 1987. Planning for conjunctive goals. *Artificial Intelligence* 32(3): 333–377.
- Clark, Andy. 1997. *Being there: Putting brain, body, and world together*. Cambridge, MA: MIT Press.
- Clark, Herbert H. 1996. *Using language*. New York: Cambridge University Press.
- Cohen, Philip R., Jerry Morgan & Martha E. Pollack (eds.), 1990. *Intentions in communication*. Cambridge, MA: MIT Press.
- Csibra, Gergely & György Gergely. 1998. The teleological origins of mentalistic action explanations: A developmental hypothesis. *Developmental Science* 1(2): 255–259.
- Dalrymple, Mary, Stuart M. Shieber & Fernando C.N. Pereira. 1991. Ellipsis and higher-order unification. *Linguistics and Philosophy* 14(4): 399–452.
- Engelhardt, Paul E., Karl G.D. Bailey & Fernanda Ferreira. 2006. Do speakers and listeners observe the Gricean Maxim of Quantity? *Journal of Memory and Language* 54(4): 554–573.
- Gallagher, Shaun. 2001. The practice of mind: Theory, simulation or primary interaction? *Journal of Consciousness Studies* 8(5/7): 83–108.
- Garrod, Simon & Anthony Anderson. 1987. Saying what you mean in dialogue: A study in conceptual and semantic coordination. *Cognition* 27(2): 181–218.
- Gibbs, Raymond W., Jr. 2001. Intentions as emergent products of social interactions. In Bertram F. Malle, Louis J. Moses & Dare A. Baldwin (eds.), *Intentions and Intentionality*, 105–122. Cambridge, MA: MIT Press.
- Gilbert, Margaret. 2003. The structure of the social atom: Joint commitment as the foundation of human social behavior. In Schmitt, Frederick F. (ed.), *Socializing metaphysics: The nature of social reality*, 39–64. Lanham, MD: Rowman & Littlefield.
- Ginzburg, Jonathan. 2012. *The interactive stance: Meaning for conversation*. Oxford: Oxford University Press.
- Ginzburg, Jonathan & Robin Cooper. 2004. Clarification, ellipsis, and the nature of contextual updates in dialogue. *Linguistics and Philosophy* 27(3): 297–365.
- Glüer, Kathrin & Peter Pagin. 2003. Meaning theory and autistic speakers. *Mind & Language* 18(1): 23–51.
- Gold, Natalie & Robert Sugden. 2007. Collective intentions and team agency. *Journal of Philosophy* 104(3): 109–37.
- Goodwin, Charles. 1979. The interactive construction of a sentence in natural conversation. In George Psathas (ed.), *Everyday language: Studies in ethnomethodology*, 97–121. New York: Irvington.
- Gregoromichelaki, Eleni, Ronnie Cann & Ruth Kempson. Forthcoming. On coordination in dialogue: Subsentential speech and its implications. In Lawrence Goldstein (ed.), *On Brevity*. Oxford: Oxford University Press.

- Grice, H. Paul. 1975. *Logic and conversation*. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and Semantics 3: Speech Acts*, 41–58. New York: Academic Press.
- Hanna, Joy E., Michael K. Tanenhaus & John C. Trueswell. 2003. The effects of common ground and perspective on domains of referential interpretation. *Journal of Memory and Language*, 49(1): 43–61.
- Hankamer, Jorge & Ivan Sag. 1976. Deep and surface anaphora. *Linguistic Inquiry* 7(3): 391–426.
- Haugh, Michael. 2008. Intention in pragmatics. *Intercultural Pragmatics*, 5(2). 99–110.
- Haugh, Michael, Kasia M. Jaszczolt. 2012. Speaker intentions and intentionality. In Keith Allan & Kasia M. Jaszczolt (eds.), *The Cambridge handbook of pragmatics*, 87–112. Cambridge: Cambridge University Press.
- Healey, Patrick. 2008. Interactive misalignment: The role of repair in the development of group sub-languages. In Robin Cooper & Ruth Kempson (eds.), *Language in flux*, 13–39. London: College Publications.
- Heritage, John C. 1984. *Garfinkel and ethnomethodology*. Cambridge: Polity.
- Horton, William S. & Richard J. Gerrig. 2005. Conversational common ground and memory processes in language production. *Discourse Processes* 40(1): 1–35.
- Hutto, Daniel D. 2004. The limits of spectatorial folk psychology. *Mind & Language* 19(5): 548–573.
- Kempson, Ruth, Wilfried Meyer-Viol & Dov Gabbay. 2001. *Dynamic syntax*. Oxford: Blackwell.
- Kempson, Ruth., Eleni Gregoromichelaki, Wilfried Meyer-Viol, Matthew Purver, Graham White & Ronnie Cann. 2011. Natural-language syntax as procedures for interpretation: The dynamics of ellipsis construal.” In Alain Lecomte, Samuel Tronçon (eds.), *Ludics, dialogue and interaction* (Lecture Notes in Computer Science, vol. 6505), 114–133. Berlin: Springer.
- Keysar, Boaz. 2007. Communication and miscommunication: The role of egocentric processes. *Intercultural Pragmatics* 4(1): 71–84.
- Krauss, Robert M., Susan R. Fussell. 1996. Social and psychological models of interpersonal communication. In E. Tory Higgins & Arie W. Kruglanski (eds.), *Social psychology: Handbook of basic principles*, 655–701. New York: Guilford Press.
- Lerner, Gene H. 2004. Collaborative turn sequences. In G. Lerner (ed.), *Conversation analysis: Studies from the first generation*, 225–256. Amsterdam: John Benjamins.
- Levelt, Willem J.M. 1989. *Speaking: From intention to articulation*. Cambridge, MA: MIT Press.
- Levinson, Stephen C. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- Levinson, Stephen C. 1995. “Interactional biases in human thinking.” In E. Goody (ed.), *Social intelligence and interaction*, 221–260. Cambridge: Cambridge University Press.
- McDowell, John Henry. 1980. Meaning, communication, and knowledge. In Zak van Straaten (ed.), *Philosophical subjects: Essays presented to P.F. Strawson*, 117–139. Clarendon Press. Oxford.
- Merchant, Jason. Forthcoming. Ellipsis. In Artemis Alexiadou, Tibor Kiss, and Miriam Butt (eds.), *Handbook of Contemporary Syntax*, 2nd edn. Walter de Gruyter. Berlin.
- Millikan, Ruth Garrett. 1984. *Language, thought and other biological categories*. MIT Press. Cambridge MA.
- Mills, Gregory J. 2011. The emergence of procedural conventions in dialogue. *Cognitive Science Society Conference (COGSCI)*. 33. 471–476.
- Mills, Gregory J. & Eleni Gregoromichelaki. 2010. Establishing coherence in dialogue: Sequentiality, intentions and negotiation. In Pavel Iupkowski & Matthew Purver (eds.), *Aspects of semantics and pragmatics of dialogue (SemDial)*, 17–24. Poznań: Polish Society for Cognitive Science.

- Perner, Josef. 1991. *Understanding the representational mind*. Cambridge, MA: MIT Press.
- Pickering, Martin J. & Simon Garrod. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences* 27(2): 169–226.
- Poesio, Massimo & Hannes Rieser. 2010. Completions, coordination, and alignment in dialogue. *Dialogue and Discourse* 1(1): 1–89.
- Preston, Beth. 1994. Behaviorism and mentalism: Is there a third alternative? *Synthese*, 100(2). 167–196.
- Purver, Matthew, Ronnie Cann & Ruth Kempson. 2006. Grammars as parsers: Meeting the dialogue challenge. *Research on Language and Computation* 4(2/3). 289–326.
- Purver, Matthew, Christine Howes, Eleni Gregoromichelaki & Patrick Healey. 2009. Split utterances in dialogue: A corpus study. *Special Interest Group on Discourse and Dialogue (SigDial)*. 10. 262–271.
- Purver, Matthew, Eleni Gregoromichelaki, Wilfried Meyer-Viol & Ronnie Cann. 2010. Splitting the “I”s and crossing the “You”s: Context, speech acts and grammar. In Paweł Łupkowski & Matthew Purver (eds.), *Aspects of semantics and pragmatics of dialogue (SemDial)*, 43–50. Poznań: Polish Society for Cognitive Science.
- Schegloff, Emanuel A. 2007. *Sequence organization in interaction: A primer in conversation analysis*. Cambridge: Cambridge University Press.
- Schober, Michael F. & Clark, Herbert H. 1989. Understanding by addressees and overhears. *Cognitive Psychology* 21(2). 211–232.
- Searle, John R. 1969. *Speech acts*. Cambridge: Cambridge University Press.
- Searle, John R. 1990. Collective intentions and actions. In Philip R. Cohen, Jerry Morgan & Martha E. Pollack (eds.), *Intentions in communication*, 401–415. Cambridge, MA: MIT Press.
- Sperber, Dan & Deirdre Wilson. 1986/1995. *Relevance: Communication and cognition*. Oxford: Blackwell.
- Stainton, Robert J. 2006. *Words and thoughts*. Oxford: Oxford University Press.
- Stone, Matthew. 2005. Communicative intentions and conversational processes in human-human and human-computer dialogue. In John C. Trueswell & Michael K. Tanenhaus (eds.), *Approaches to studying world-situated language use*, 39–70. Cambridge, MA: MIT Press.
- Suchman, Lucy A. 1987/2007. *Plans and situated actions: The problem of human-machine communication*. Cambridge: Cambridge University Press.
- Tollefsen, Deborah. 2005. Let’s pretend! Children and joint action. *Philosophy of the Social Sciences* 35(1). 75–97.
- Wegner, Daniel M. 2002. *The illusion of conscious will*. Cambridge, MA: MIT Press.
- Wellman, Henry M., David Cross, Julianne Watson. 2001. Meta-analysis of theory-of-mind development. *Child development* 72(3). 655–684.
- Wharton, Tim. 2003. *Pragmatics and the showing–saying distinction*. London: UCL PhD dissertation.