Andrew Gargett · Eleni Gregoromichelaki · Ruth Kempson · Matthew Purver · Yo Sato

Grammar Resources for Modelling Dialogue Dynamically

August 6, 2009

Abstract This paper argues that by analysing language as a mechanism for growth of information [32,9], not only does a unitary basis for ellipsis become possible, otherwise thought to be irredeemably heterogeneous, but also a whole range of sub-types of ellipsis, otherwise thought to be unique to dialogue, emerge as natural consequences of use of language in context. Dialogue fragment types modelled include *reformulations*, *clarification requests*, *extensions*, and *acknowledgements*. Buttressing this analysis, we show how incremental use of fragments serves to progressively narrow down the otherwise mushrooming interpretational alternatives in language use, and hence is central to fluent conversational interaction. We conclude that, by its ability to reflect dialogue dynamics as a core phenomenon of language use, a grammar with inbuilt parsing dynamics opens up the potential for analysing language as a mechanism for communicative interaction.

A. Gargett, E. Gregoromichelaki, R. Kempson
Department of Philosophy, King's College London
Strand, London WC2R 2LS, UK
Tel.: +44-20-77848 2231
Fax.: +44-20-7848 2270
E-mail: {andrew.gargett,eleni.gregor,ruth.kempson}@kcl.ac.uk

M. Purver
Department of Computer Science, Queen Mary University of London
Mile End Road, London E1 4NS, UK
E-mail: mpurver@dcs.qmul.ac.uk

Y. Sato
Science and Technology Research Institute, University of Hertfordshire
College Lane, Hatfield AL10 9AB, UK
E-mail: y.sato@herts.ac.uk

1 Introduction

In confronting the challenge of providing formal models of dialogue, with its plethora of fragments and rich variation in modes of context-dependent construal, it might seem that linguists face two types of methodological choice. Either (a) conversation demonstrates genre-specific characteristics, for which a grammar specific to such activity must be provided ([57, 49, 45, 16, 2, 20] cf. [58]); or (b) the cross-speaker flexibility and variation characteristic of dialogue has to be seen as due to the specifics of the parsing/production systems which are based upon, but nevertheless distinct from, the mode-neutral grammar characterising the individual's competence in that language (see e.g. [11]). Both alternatives raise issues for an empirical account of what language users' knowledge of a given language amounts to. One core phenomenon where such issues are vividly displayed is *ellipsis*, a phenomenon observed both in monologue and dialogue genres. Ellipsis simultaneously constitutes one of the most fundamental characteristics of dialogue while, nevertheless, being subject to restrictions diagnostic of grammar-internal mechanisms. Thus ellipsis would seem to straddle the remits of dialogue-modelling and grammar-design in a way that does not allow for a unified explanation given conventional grammar formalisms. Preserving the separation between discourse phenomena and core grammar constraints, this has been taken to indicate a division between discourse ellipsis and grammar-internal ellipsis [64], confirming what is widely taken to constitute the heterogeneity of ellipsis phenomena [42]. But such a move notably fails to reflect the fact that the single most defining feature of all types of ellipsis is systemic context-dependence. Even when this feature of ellipsis is recognised, the presumed split between discourse-based and grammar-constrained ellipsis has been taken to vindicate the claim that a dialogue-specific grammar is necessary to deal with dialogue elliptical phenomena.

In this paper, to the contrary, we will show that ellipsis phenomena in general are characterisable by grammar-internal mechanisms without any stipulation specific to their use in dialogue. This result is achieved by developing a grammar formalism that directly reflects the dynamics of language processing as the basis for explaining structural properties of language. More specifically, the grammar tracks the incremental and progressive growth of interpretation in real time. *Procedures*, equivalently *actions*, determining growth of represented information are central to the formal specification of the grammar itself, not a characteristic of just the parser or the generator. Accordingly, the methodology of preserving a gulf between the design of the grammar-formalism and performance considerations is no longer sustained.

The approach within which we set out our account of fragment construal, is that of *Dynamic Syntax* (*DS*) [32, 9]. In this framework, the usually static notion of "syntax" as a set of principles assigning structures over strings of words is replaced by "syntax" as the progressive construction of semantic representations set in context, reflecting the dynamics of human parsing. Language production is presumed to be parasitic on the very same information-growth processes, following the dictates of the grammar formalism.

Such a framework is well suited to modelling ellipsis in all its forms: the diverse grammar-internal constraints on ellipsis postulated by standard frameworks emerge in DS from the definition of the dynamics of the time-linear accumulation of information. This consists of three dimensions:

(a) as the parser operates incrementally left-to-right, the set of words processed progressively increases;
(b) the semantic representations, which in DS take the form of gradually updated tree structures, progressively grow in that they become more specified both in terms of structure and content annotations;
(c) the set of actions used in constructing tree representations accumulates.

In [51] and [10] it is argued that *context* should then be accordingly defined as incorporating a record of (a) words processed,[1] (b) structures established, and (c) actions that led to this structure. This model of context is what allows DS to deal with ellipsis in a unified way as context-dependency (see [51,10] for details). Storage of actions allows a straightforward account of cases of ellipsis that have been argued to provide crucial evidence for idiosyncratic syntactic/semantic restrictions. In particular, both semantic and syntactic restrictions on fragments are dealt with in the same way, as constraints on tree growth. Sets of actions/procedures, what lexical items are presumed to consist in, are sequentially executed and subsequently stored in the context record. They can then be available for recovery and re-use, and some types of ellipsis crucially rely on this feature of context for their resolution. So, on this account, the concept of *procedure* emerges as central to what context amounts to as well.

In this paper we adopt this analysis as background for demonstrating how a range of elliptical phenomena which might, at first sight, seem specific to conversational dialogue – acknowledgements, clarifications, reformulations, utterance-exchanges involving interruption and intra-sentential switch of speaker-hearer roles – can be modelled without any stipulation specific to such functions. The immediate advantage of such an account is that it preserves, indeed directly reflects, the intuition that ellipsis occurs when the context fully determines construal as far as the message conveyed is concerned except for the elliptical fragment presented. In addition, we shall conclude that grammars that are defined as dynamical systems directly reflecting mechanisms for growth of interpretation in real time also allow for dynamic adjustment of the information conveyed according to the feedback received by the other interlocutor. Accordingly, grammars can be seen as a mechanism for communicative interaction: the human capacity for language is thus grounded directly in the interactional activity which it serves.

---

[1] There are two main reasons for the storage of words: firstly, there is the fact that people can remember them, at least in the short term; secondly, word/action pairs need to be available for the modelling of lexical/syntactic *alignment* (see [46, 51]). A reviewer points out that recall of words (qua phonological units) decays faster than content (but cf. [28,33]). To integrate this assumption, to the extent that it holds, in the DS model we have to introduce mechanisms that model decay of information in context (see e.g. [38]), an extension which is in any case required for modelling decay of accessibility of competing antecedents for anaphora/ellipsis.

2 Background

In natural language use, people often talk to each other with apparently fragmentary utterances, a phenomenon broadly known as *ellipsis*:

(1) A: Have you seen Mary?
     B: (a)   Mary?
         (b)   No, I haven't.
         (c)   No, but I have Bill.
         (d)   Yeah, Tom too.

Despite a robust folk intuition that ellipsis can occur whenever the context makes obvious how the apparently fragmentary expression of thought is to be "completed", ellipsis has very generally been seen as a heterogeneous set of phenomena not subject to a uniform explanation. The reason for this are the various complications that arise in defining the way the interpretation of such fragments is achieved in each case.

Amongst the rich array of ellipsis effects, there are cases of elliptical fragments where the the linguistic surface form of the antecedent[2] provides the resolution. These cases have been subject to a great deal of study over recent years - *VP-ellipsis* as in (1b), *sluicing*, *stripping* (as in (1d)), *gapping*, *pseudo-gapping* (as in (1c)), etc. Each sub-type is said to display different structural constraints on their construal, justifying their distinct analyses. What brings these together is that they, in some sense, constitute a complete sentence, given their interpretation as suitably completed by the provided antecedent, so that ellipsis can be seen as requiring syntactic/semantic rules defined over adjoined sentences (though see [64]).

But, even for these cases, the effects ellipsis resolution induces remain very puzzling as they appear, from that point of view, unresolvably heterogeneous. For example, there are cases where relative to a single antecedent source ambiguity may arise, with a single string admitting either a *strict* or a *sloppy* reading:

(2) John checked over his mistakes, and so did Bill/Bill too.
     'Bill checked Bill's mistakes' (sloppy)
     'Bill checked John's mistakes' (strict)

This phenomenon would seem to be broadly a phenomenon of construal, requiring a semantic basis for explanation under which the CONTENT of the antecedent has somehow to be manipulated by making available alternative forms of abstracts to match what is required at the ellipsis site (hence the classic account in terms of various abstraction operations applicable to the content of the first conjunct to yield a novel predicate applicable to the subject of the second, *higher order unification* [14]). However, the semantic account is not adequate on its own: elliptical phenomena also appear to be

---

[2] The *antecedent* of an elliptical utterance is whatever in the previous discourse/context provides the necessary "completion" for its understanding; thus in e.g. (1b) the string *seen Mary* is the antecedent because it provides whatever material is needed for the resolution of the *ellipsis site* provided by the auxiliary *haven't*.

sensitive to the very same constraints that affect surface syntactic operations. For example, as a general observation across languages and phenomena, relative clause constructions interfere with "movement" operations (they are *islands*). In the syntactic literature, this constraint is called the *Complex NP constraint*. Now exactly the same type of interference seems to be responsible for the ungrammaticality of the VP-ellipsis construction in (4) below:

(3) John interviewed every student who Bill already had.
(4) *John interviewed every student who Bill ignored the teacher who already had.

What is displayed in (3) is the phenomenon of *antecedent contained deletion* (or *antecedent-contained ellipsis*), so called because the ellipsis site appears to be contained within the antecedent (*interviewed every student who ...*) from which the interpretation of the ellipsis site itself is to be constructed. Such circularity apart (for which different solutions are promulgated depending on the framework), the observation is that such structures preclude any dependency of the relative pronoun, *who*, on a verb appearing across a relative-clause boundary. Accordingly, (4) is ungrammatical because there is no possibility of resolving the ellipsis site indicated by *had* by means of the containing matrix verb in the way that (3) appears to allow. This pattern is reminiscent of restrictions on so-called *long-distance dependencies* which are taken to be diagnostic of a syntactic phenomenon, non-reducible to semantic terms:

(5) The man who Sue is worried that her sister is planning to marry lives in Austria.
(6) *The man who Sue is concerned about her sister who is planning to marry lives in Austria.

Though (5) is wellformed with *who* construed as the object of *marry*, no such dependency for the first relative pronoun in (6) is possible, because there is a further relative clause boundary (an island) between the relative pronoun *who* and the same verb *marry* that *who* has to be associated with (the observation goes back to [53]). Because such restrictions are not expressible in semantic terms – for example, the lambda calculus (the logic taken to underpin semantic combinatorics) would impose no such restriction – they have been taken as evidence for a concept of syntax independent of both semantics and phonology, and a diagnostic of what constitutes a syntactic process. Some types of ellipsis at least are thus argued to be within the remit of natural-language syntax, involving low-level deletion of phonological material (*PF Deletion*: see [17, 41, 42]).[3]

2.1 Dialogue Ellipsis

What all these grammar-internal characterisations of ellipsis miss, given their remit of characterising only sentence-internal properties, is the broad array

---

[3] Further cases indicate ellipsis construal sensitivity to morphological idiosyncracies (see [43, 44, 21, 51, 10] for discussion).

of elliptical effects in dialogue. One of the most striking characteristics of conversational dialogue is the extent, and freedom, with which participants make use of utterance fragments. Indeed fragmentary expressions that occur in dialogue apparently allow interpretations that indicate many sorts of conversational interaction. For instance, interlocutors can extend each other's utterances, while at the same time displaying their acceptance/understanding of the other's presentation:

(7)   A: Bob left.
      B: the accounts guy, (yeah).

They can interrupt and finish each other's utterances:

(8) Conversation from A and B, to C:
      A: We're going to
      B: Bristol, where Jo lives.

They can even use each other's utterances as the basis for what they themselves have to say, without waiting for their interlocutor to finish:

(9)   A: Most of the ones that we brought seem to have erm
      B: survived
      B: survived. Which I'm glad. [BNC: KBP 2507-09, from [54]]

And, as these examples illustrate, such switch of roles between hearer and speaker can take place across any syntactic dependency whatsoever: across a determiner-noun dependency (11), across a preposition-NP dependency (8), or auxiliary-verb dependency (9) (see [52] for corpus evidence for this claim).

Such fragmentary utterances may be interpretable only relative to partial contents currently being presented by other interlocutors. Such utterances seem in some sense to be constructed jointly by participants, relying on feedback by clarification, disagreement, or correction (henceforth, A female, B male):

(10) A: Have you mended
      B: any of your chairs? Not yet.
(11) (A smelling smoke coming from the kitchen:)
      A: Have you burnt the
      B: buns. Very thoroughly.
      A: But did you
      B: burn myself? No. Luckily.
(12) A: And er they X-rayed me, and took a urine sample,
           took a blood sample.
      A: Er, the doctor
      B: Chorlton?
      A: Chorlton, mhm, he examined me, erm, he, he said now they
           were on about a slide [unclear] on my heart.
           [from the BNC: KPY 1005-1008]
(13) A: Bob left.
      B: Rob?
      A: (No,) (Bob,) the accounts guy.

(14)   A: Is this yours or
        B: Yours.

It might seem, and it has traditionally been assumed, that the grammar/syntax, which under orthodox assumptions has to do with sentence strings and propositions, should have nothing to say about such fragmented dialogue turns, as they constitute "performance" data. But firstly, as was pointed out early on [43], there are grammatical sequences of words other than sentences, as can be seen in all the examples above; and the grammar should be able to characterise those and distinguish them from plain "word-salad". Secondly, the grammar is responsible for characterising certain dependencies among lexical items in grammatical sentences. Consider then the licensing of the negative polarity item (NPI) in (10): Such NPIs are only licensed by appearing in sentences that contain some explicit "affective" element, namely negation, question etc. (see e.g. [34]). Now, in (10), the only element that can license the NPI is the interrogative morphology registered at A's turn. It would seem then that A and B's turns should somehow be joined together to form a single utterance, otherwise we would not be able to match the intuition that the discourse is perfectly well-formed. Now one might be tempted to conclude that, indeed, this vindicates a grammar which characterises sentence strings as it would seem that this licensing occurs only when we assume that a single string of words is spread over two turns. But this would be too hasty: Notice what happens with the second A and B exchange in (11). The licensing of the reflexive anaphor *myself* is only possible because its antecedent, namely B, is part of the CONTENT of the turn started by A. But if we now try to join the two strings together (#*but did you burn myself*)[4] the result is not what the exchange was meant to convey. The same problem would arise with (14). From that point of view, it seems then that this phenomenon of turn-sharing, loosely characterised as *split utterances*, has to do with the sharing of contents rather than strings.

Given this conclusion, one might presume that what is needed is a grammar for dialogue, with the specific remit of determining how content is jointly constructed over turns by interlocutors (see e.g. [57,45]). Since content computation is known to rely crucially on context, such a grammar would have to regiment the pragmatic inferencing that results in the derivation of what interlocutors actually do with their utterances, including specification of the utterance's *speech-act* function in the dialogue (see e.g. [22]). Let's see whether this would be a desirable move according to the data presented here. While the NP fragments in (11)–(14) might be characterised as distinct utterance types, serving severally rather different functions of clarification, acknowledgement, correction, they also illustrate how both interlocutors may progressively contribute to the joint enterprise of establishing some shared communicative content. In such examples, each speaker contributes parts to a single collaborative utterance (as seen by the preservation of such grammatical constraints as those discussed above) and thus these also fall under the phenomenon loosely characterised as *split utterances*. Here the dialogue

---

[4] The point is that even though this string might be grammatical it does not carry the intended meaning.

participants jointly construct an utterance even though they may well NOT have identical messages in mind guiding what they are going to say. Speakers can divert or preempt the other interlocutor's planned utterance and this is a facet of dialogue that is easily handled by the participants. Moreover, each speaker's contribution is not necessarily intended to have the same function as the previous speaker's incomplete utterance which it continues. Even (7), an *acknowledgement*, can be seen as part of a split utterance, for it is similar in form to an afterthought *extension* added to A's sentential utterance. As (12) shows, joint construction of content can proceed incrementally: B provides a *reformulation* as a *clarification request*, resolved by A within the construction of a single proposition. In (13), the fragment A provides functions as a *correction* of some aspect of B's understanding, with A and B having to negotiate as to whose information is more reliable in order to secure coordination. Nevertheless the correction in (12) also constitutes an *extension*, so that a single conjoined propositional content is derived DURING which coordination is achieved. (14) represents an intermediate case, in which the respondent, realising what the question is, answers THROUGH his *completion* of the question; the fragment thus serves simultaneously both as question and answer (see e.g. [6]).

Such diversity of fragment uses might, then, seem evidence of conversation-specific rules disambiguating the function of such fragments as part of a grammar. Taking such a line, [16] presents a thorough taxonomy, as well as detailed formal and computational modelling of *non-sentential utterances*, referring to contributions like e.g. (7) as *repeated acknowledgements* involving *reformulation*. [16] models such constructions via type-specific "accommodation rules" which make a constituent of the antecedent utterance "topical". The semantic effect of the acknowledgement is then derived by applying an appropriately defined utterance type for such fragments to the newly constructed context. A distinct form of contextual accommodation is employed to model so-called *helpful rejection* as in (13) and so on. Under this view, the distinct dialogue acts that can be accomplished by a fragment consisting of a single NP, e.g. *the accounts guy*, have to be characterised by postulating distinct appropriate grammatical and contextual rules that constitute part of the dialogue management process. However, if, as seems to be the case,[5] what actions speakers and hearers can perform and recognise in dialogue is limitless and nondeterministic, shaped through feedback,[6] an a priori taxonomy like this is not a desirable move, for it requires that a single linguistic form such as an NP-fragment will have to be characterised as syntactically multiply ambiguous, relative to a set of predetermined dialogue acts. And moreover, such a taxonomy still faces the problem of specifying the exact function(s) each form performs, as this depends on a full specification of its context of occurrence, which is also necessary for resolving the postulated ambiguity. Additionally, what such an account fails to bring out is how these phenomena are structurally replicated in monologue in *apposition* structures, possibly extraposed (for a similar observation see also [48]):

---

[5] See e.g. [37,36,63,58].
[6] See e.g. [12].

(15) Bob, the accountant, is coming to stay.
(16) Bob left, the accountant.

Furthermore, even these construction-specific analyses of fragments in dialogue are taken to require a sentential form of analysis: the fragment is assigned a type which in combination with a suitable abstract with respect to what is provided in context will yield a propositional form of meaning. But, as (13) indicates, such fragments can be used at early points in a dialogue exchange when there may be no complete propositional content available in the context relative to which the fragment can be provided a suitable abstract. Nonetheless the participants in the dialogue can exchange a clarificatory request and reply unproblematically so that the communicative exchange can proceed.[7]

The challenge posed by split exchanges for orthodox grammar, parser and production mechanisms is considerable, and they have not been much addressed (though see [48]). The problem is as follows: speakers shift into hearing as though they had been hearing all along: parsers shift into production as though they had been speaking all along. But, until recently, parsing and generation systems have been defined relative to a grammar whose remit is syntactic and semantic analysis of complete sentence-strings. And, even now, though parsing and generation systems are increasingly reflecting incrementality [2,62,66], such incrementality must generally come from the processing model, with the grammar defined statically and independently. Yet, to deal with split utterances, parsing/generation systems have to be defined with a flexibility allowing either one to take up from where there has been a switch, despite the fact that both the string preceding or following the switch may fall outside the set of strings licensed as well-formed by the grammar. Many incremental systems adopt ad hoc grammar relaxation rules to deal with the incremental nature of dialogue. A grammar defined in terms of incremental growth of interpretation and, as a corollary, context is not faced with this problem though. To the contrary, if, by assumption, the same mechanisms for incremental monotonic growth are shared by both systems, parsing and production, it is this type of framework that, uniquely, can model this unproblematic shift of role as a wholly natural, indeed a predictable consequence. This suggests that incrementality in both generation and parsing, and the potential to provide update from whatever intermediate point speaker-hearer switch takes place, is at the core of the explanation for the frequency of split utterances in conversation.

Accordingly, we now turn to exploring the use of grammar-internal resources to capture such phenomena. In this type of grammar, as we shall see, it will be actions/procedures for interpretation which replace the static notions of (representations of) denotational content as the central notion. The relevant conceptual point here is that conversational dialogues emerge incrementally over the course of the interaction through the distinct contributions of the participants, each employing no more than resources internal to the grammar.

---

[7] Incidentally such interruptive conversational exchanges pose problems for another type of "grammar of conversation", namely, models which describe *turn-taking* organisation and *adjacency pair* structure, see e.g. [55,56].
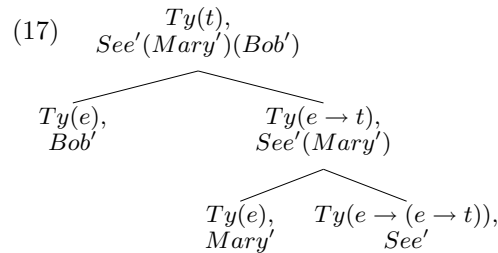
3 Dynamic Syntax: A Sketch

3.1 Introduction

Dynamic Syntax (*DS*) has three main characteristics underlying the modelling of how contextual information can be incorporated AS IT ARISES with linguistic information during interaction in dialogue. First, DS provides a fully incremental parsing model, based on a strictly monotonic process of interpretation update. Second, this process of update is goal-directed as it is dynamically driven by *requirements* for update at each initial or intermediate stage. As will be shown shortly, these two characteristics are closely related, and together allow for the third DS distinguishing feature, the fact that the same mechanisms are exploited to model both parsing (= understanding) and generation (= production).

3.2 Content representations

Taking up formulation of content first and its tree-theoretic representation, in the DS the output of a parse is modelled as formulae annotating trees. Such trees are formalised using *LOFT* [4], a modal logic designed to allow the processor to refer to partial, complete or required tree structure. LOFT makes available modal operators $\langle\uparrow\rangle, \langle\downarrow\rangle$ defining the concepts of *mother* and *daughter* nodes and their iterated underspecified counterparts, $\langle\uparrow_*\rangle, \langle\downarrow_*\rangle$,[8] defining the notions *be dominated by* and *dominate*, thus allowing for partially specified trees in terms of structure. Annotations (*decorations*) on those nodes determine *formula*, *type* ('*Ty*') and *tree-node* ('*Tn*') position, and a pointer, $\diamondsuit$, indicates the node under development. Complete individual trees are taken to correspond to predicate-argument structures, with nodes in such trees decorated with sub-terms of the propositional formula appearing at the root. Formulae on mother nodes are composed from information decorating their daughters using the combinatorics of the lambda calculus (however we omit the lambda-binding operators in the graphics below to simplify illustration):

(17)
$$Ty(t),\ See'(Mary')(Bob')$$
$$Ty(e),\ Bob' \qquad Ty(e \rightarrow t),\ See'(Mary')$$
$$Ty(e),\ Mary' \qquad Ty(e \rightarrow (e \rightarrow t)),\ See'$$

　　The representation language for content in DS is the *epsilon calculus*, a conservative extension of predicate logic. Thus quantificational NPs contribute terms of low type *e* (rather than some higher generalised-quantifier

---

[8] The symbol appearing in these annotations is the Kleene star used to indicate zero or more iterations of the dominance relation.

type as in [3]). Their processing results in the construction of the natural-deduction counterparts of quantifiers, *arbitrary names*, encoded as terms in the epsilon calculus. For example, the natural language NP *a man* contributes the content $(\epsilon, x, Man'(x))$ which stands for the witness of the formula $\exists x Man'(x)$ according to the following equivalence:

(18) $\quad \dfrac{\exists x \phi(x)}{\phi(\epsilon, x, \phi(x))} \qquad \dfrac{\text{predicate logic formula}}{\text{epsilon calculus equivalent}}$

The advantage of such terms is that they can be extended to include their context of occurrence inside their restrictor. So, for example, the semantic evaluation rules for a proposition derived from the string *A man cried* will produce not simply the term $\epsilon, x, (Man'x)$ but instead $\epsilon, x, (Man'x \wedge Cry'x)$ which denotes a *witness* of the set of men who cried:

(19) $Ty(t), Cry'(\epsilon, x, Man'x) \qquad \mapsto$
    $Ty(t), Cry'(\alpha) \qquad$ where $\alpha = (\epsilon, x, (Man'x \wedge Cry'x))$

This allows a straightforward account of cross-sentential anaphoric dependence on quantifier antecedents[9] as the term abbreviated as α above can now serve as the referent associated with the pronoun *he* in a possible continuation such as *He was upset*.[10] Similarly, names and definites can be analysed as *iota terms* (epsilon terms of widest scope). So a name like *Bill* or a definite description like *the man* will contribute terms $(\iota, x, Bill'(x))$ and $(\iota, x, Man'(x))$ to the tree representation (in the graphics below we omit the epsilon terms for simplicity of illustration unless they have a bearing on the analysis).[11]

3.3 Dynamics

DS employs an underspecification-plus-enrichment model of update in context. Underspecification is employed at all levels of tree relations (mother, daughter etc.), as well as formulae and type values, each with an associated *requirement* driving the goal-directed process of update. Tree updates are executed with lexical and general computational actions, which can be understood as transition functions from one tree representation to another. *Computational actions* govern general tree-constructional processes, introducing/updating structure and compiling interpretation for all non-terminal nodes in the tree once individual leaf nodes are successfully decorated (with no outstanding requirements). This may include the construction of only weakly specified tree relations, characterised only as dominated by some node from which they are constructed (*unfixed nodes*), with subsequent update

---

[9] *E-type* anaphora: [15] and many others since

[10] Relative scope is not expressed by the hierarchical structure of the tree but involves incremental collection of scope-dependency constraints (either lexical or structurally determined) with the output formulae and the set of scope dependencies being subject to an evaluation algorithm determining their combined effect on interpretation.

[11] The account of names and definites is simplified here for exegesis, but see [9].

through *unification* with some other node (unlike [35], here partial trees are part of the model).

In DS, individual lexical items are conceived of as packages of *lexical actions* for building structure and introducing content, expressed in exactly the same terms as the more general processes inducing nodes and decorations. Thus *partial trees* grow incrementally, driven by procedures associated with words as encountered, with the pointer, $\diamondsuit$, tracking the parse progress, and thus taking care of word-order – see Figure 1.

Initial Axiom: $?Ty(t), \diamondsuit$ $\boxed{0}$

Introduction/prediction: $\hookrightarrow$

$?Ty(t)$

$?Ty(e), \diamondsuit$ $?Ty(e \to t)$ $\boxed{1}$

Parsing *"Bob"*: $\hookrightarrow$

$?Ty(t)$

$Ty(e), \boldsymbol{Bob'}$ $?Ty(e \to t),$ $\boxed{2}$
$\diamondsuit$

Parsing *"sees"*: $\hookrightarrow$

$?Ty(t)$

$Ty(e),$ $?Ty(e \to t)$ $\boxed{3}$
Bob'

$?Ty(e),$ $Ty(e \to (e \to t)),$
$\diamondsuit$ $\boldsymbol{See'}$

Parsing *"Mary"*: $\hookrightarrow$

$Ty(t), \diamondsuit$
$See'(Mary')(Bob')$

$Ty(e),$ $Ty(e \to t),$ $\boxed{4/T_g}$
Bob' $See'(Mary')$

$Ty(e),$ $Ty(e \to (e \to t)),$
$\boldsymbol{Mary'}$ See'

Fig. 1 Monotonic tree growth in DS

An expected tree starts out as a node with the decoration $?Ty(t)$, an entirely underspecified minimal tree requiring a proposition, but this will be enriched or specified progressively. A node in general may, for example, be specified so as to determine that its only legitimate updates are logical expressions of individual type ($?Ty(e)$), or the requirement may also take a modal form, e.g. $?\langle\uparrow\rangle Ty(e \to t)$, a restriction that the mother of this node

be decorated with a formula of predicate type. *Requirements* are essential to the DS dynamics: all requirements must be satisfied if the construction process is to lead to a successful ("grammatical") outcome.

3.4 Parsing vs. Generation

A crucial feature of the DS dialogue model is that it is *bi-directional*: parsing and generation use the same action definitions and build the same representations, following essentially the same procedure of left-to-right updates through actions. This is due to DS's monotonic incrementality and goal-directedness. In fact, we propose that these two features can be characterised as closely related through *predictivity*, a term often used in literature on human online processing (e.g. [19]). Incrementality is often characterised informally as information growth at each word input, but it can be more precisely defined as the input word being incorporated into a predicted structure, in a manner close to the notion of *connectedness* in some psycholinguistic literature [67,13]. Information growth, that is, is always ensured at each input, as a word (lexical action) either fulfills a requirement or creates a requirement as well as contributing the word meaning. Monotonicity is also ensured, as what is 'required' or predicted initially in a DS parse is the underspecified type t node, and a series of more specific tree-goals are created from this as the parse proceeds. It is this predictive feature that renders DS bi-directional. In parsing, the hearer builds a succession of partial parse trees, of course without record of what the eventual proposition is going to be but with that partial tree including predictions about what can follow, in the form of as yet unsatisfied type requirements (see Figure 1 steps 0-4). Generating the same sentence proceeds in exactly the same fashion, provided that a GOAL TREE T_g (tree 4 in Figure 1) is available for the speaker representing what they wish to say. Each possible step in generation, each uttered word, is governed by whatever step is licensed by the parsing formalism. This is further constrained by the required *subsumption* relation of the thus-far constructed "parse" (partial) tree to the GOAL TREE T_g: the parse tree must be extendible to T_g according to the DS principles in order for the currently considered generation step to be licensed. By updating their growing "parse" tree relative to the GOAL TREE (via a combination of incremental parsing and lexical search), speakers produce the associated natural language string.

4 Context, anaphora and ellipsis in DS

4.1 Anaphora

Content and structural underspecification both play important roles in facilitating successful linguistic interaction. Linguistic items like pronouns are paradigm cases of such underspecification in terms of their content. This type of content underspecification is represented in DS as involving a placeholding metavariable, noted as e.g. **U**, **V** etc., plus an associated requirement for replacement by an appropriate term value: $?\exists \mathbf{x}.Fo(\mathbf{x})$. This value has to
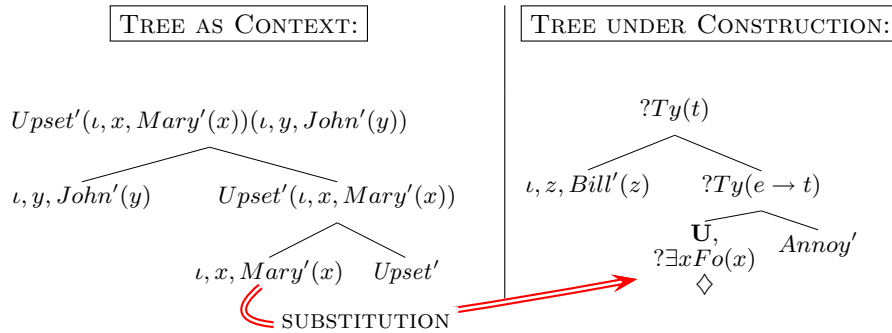
be supplied by the context of the discourse. *Context* in DS involves storage of entire *parse states* which includes

(a) a record of the words processed to date,
(b) the tree structures built up and
(c) the actions utilised to build these structures.

So consider the parsing of B's utterance below in the context of A's utterance:

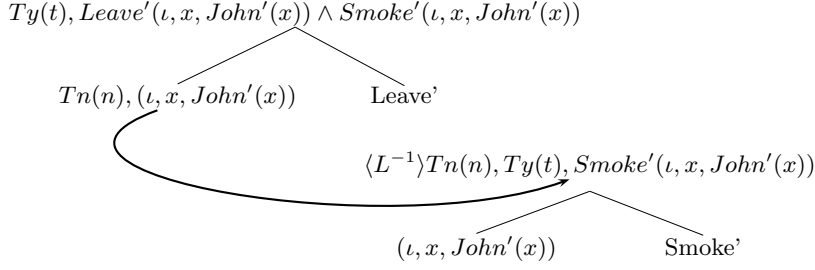(20) A: John upsets Mary.
        B: Bill annoys her.

(21)

$$\boxed{\text{Tree as Context:}}$$

$Upset'(\iota, x, Mary'(x))(\iota, y, John'(y))$

$\iota, y, John'(y)$ $Upset'(\iota, x, Mary'(x))$

$\iota, x, Mary'(x)$ $Upset'$

SUBSTITUTION

$$\boxed{\text{Tree under Construction:}}$$

$?Ty(t)$

$\iota, z, Bill'(z)$ $?Ty(e \rightarrow t)$

U,
$?\exists x Fo(x)$ $Annoy'$
$\diamond$

The process of *Substitution* licenses copying of a term from the context tree to replace a metavariable awaiting such replacement.

4.2 Relative clauses

Context also accumulates incrementally while processing the utterance itself. One instance of this phenomenon is the case of *relative clauses* in English, which require more complex structures than the simple binary predicate-argument structures we have seen so far. In DS, these are modelled via a general tree adjunction operation defined to license the construction of a tree sharing some term with another newly constructed one, yielding so-called LINKed *trees* (Kempson et al. 2001). The resulting combined information from the adjoined trees appears as a conjunction of formulae at the tree FROM which the link is made. In such constructions, the relative pronoun provides a copy of the head noun which is to appear inside the LINKed tree. The content derived on this tree is then incorporated in the main structure as an extension of the local content formulae decorations. In other words, each (potentially partial) tree is used in turn as context for the processing of the other with flow of information between them as more and more linguistic input is processed. For instance, processing of *non-restrictive relatives* gives rise to the following type of structure as output:[12]

---

[12] $\langle L^{-1} \rangle Tn(n)$ is an address annotation indicating that $Tn(n)$ is the node from where the LINK relation originates.

(22) John, who smokes, left:

$$Ty(t), Leave'(\iota, x, John'(x)) \land Smoke'(\iota, x, John'(x))$$

$Tn(n), (\iota, x, John'(x))$ Leave'

$$\langle L^{-1}\rangle Tn(n), Ty(t), Smoke'(\iota, x, John'(x))$$

$(\iota, x, John'(x))$ Smoke'

4.3 Apposition

This mechanism allowing the construction of trees in pairs is also applicable in the modelling of *apposition* devices as in (15), (23):

(23) A friend, a musician, smokes.

Here the constraint on LINKed structures as sharing a term is met through the construction of a compound term made up of a restrictor derived from the paired formulae derived from the NPs in apposition. The derivation of such formulae involves building a transition from a node of type e to a LINKed tree, constrained to host a term of the same type. For (23), the content resulting from the combined processing of *a friend* and *a musician* is as follows:

(24) Parsing *A friend, a musician*

$$?Ty(t)$$

$Ty(e), (\epsilon, x, Friend'(x)),$
$\diamond$
     $?Ty(e \to t)$

$$Ty(e)$$
$$(\epsilon, x, Musician'(x))$$
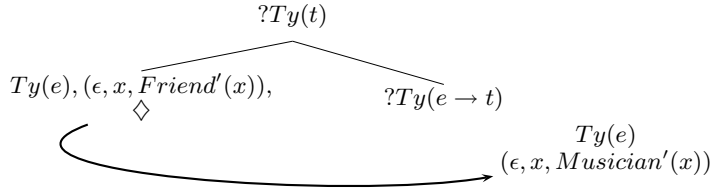
Evaluation of the LINKed nodes, both of type e, yields the composite term:
$(\epsilon, x, (Friend'(x) \land Musician'(x)))$:

(25) Parsing *A friend, a musician*

$$?Ty(t)$$

$(\epsilon, x, (Friend'(x) \land Musician'(x))), \diamond$ $?Ty(e \to t)$

$$Ty(e)$$
$$(\epsilon, x, Musician'(x))$$
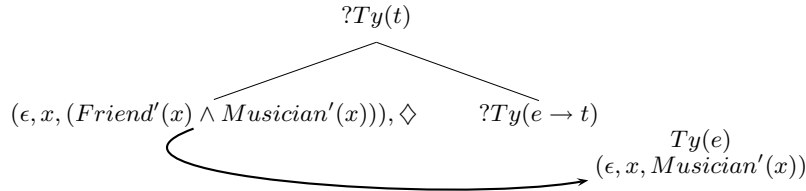
The final formula that is derived from the parse of *A friend, a musician, smokes* is: $Smoke'(\epsilon, x, (Friend'(x) \land Musician'(x)))$ (for details see [9]).

4.4 Ellipsis

In DS, these various mechanisms are brought together to provide a uniform basis for ellipsis construal ([51,9,10]). Matching the folk intuition, in DS, ellipsis is taken as providing a window on the concept of context required in linguistic explanation. Context itself as we have seen includes a record of parse states, primarily the most immediate one. Parse states are triples of word-sequence, (partial) tree structure which is the output of the processing of the words, and the set of actions that led to the build up of the structure. With this new concept of context (departing from regular denotational assumptions as in [14,65] etc.), it is argued that the full range of semantic and syntactic effects displayed in ellipsis can be expressed while preserving a unitary account of the construal process.

For example, in (2), the strict interpretation can be established by presuming that the content of the predicate established for the first conjunct is copied at the ellipsis site, with the effect that the predicate is identically applied in both conjuncts, one form of *parallelism*:

(2) John checked over his mistakes, and so did Bill/Bill too.
    'Bill checked John's mistakes' (strict)
    'Bill checked Bill's mistakes' (sloppy)

The sloppy interpretation also can be established by identity with some construct taken from context. But, in this case, it is the sequence of structure-building actions that is identically applied in both conjuncts resulting in another type of parallelism effect. The relevant sequence of actions required for the processing of the first conjunct of (2) are:

(a) the actions introducing a predicate term *Check-over'* by processing the verb *checked over*, and
(b) the actions associated with *his mistakes* which introduce as the object argument of the predicate *Check-over'* a term whose restrictor is a relation between this object and some individual identified as the subject of *Check-over'*.

For the resolution of the ellipsis, it is this sequence of actions, (a) and (b), that was used to construct the first conjunct, that is repeated now a second time. But in the new local environment, with a new, distinct subject (*Bill*) this same sequence can result in a distinct reading, the sloppy interpretation. So even if the derivation of this interpretation exactly parallels the mode of construal of the first conjunct it is nevertheless denotationally distinct from that of the first conjunct because of the new local context in which it applies. Formally this involves having the ellipsis site introduce a temporary value, a metavariable **V** of predicate type, $Ty(e \rightarrow t)$, constraining the choice of sequence of actions. In (2), there is only one such available sequence of actions in the context, but in other cases, there may be more than one, leading to ambiguity:

(26) Sue was checking her results because she was worried her teacher was checking them and Molly was too.

These variable but parallel effects in ellipsis construal are predicted in a model where the context stores and makes available not only semantic representations but also procedures for constructing these representations. In being predicted to be available, *parallelism* effects can thus be explained, while nevertheless capturing the diversity of interpretations apparently developed from a single antecedent source: what is reconstructed at the ellipsis site is reiteration of the very same actions used in building up interpretation for the first conjunct, applied now to the partial tree induced by proceessing the fragment.

In a similar vein, syntactic constraints are also analysed in tree-growth terms. In particular, the island restriction debarring dependency across relative-clause boundaries (the *Complex NP constraint*) is reconstructed as a tree-update constraint affecting all long-distance dependencies. Since the structural information projected by relative pronouns is required to be initially underspecified, inevitably, its update will be constrained by island restrictions too:

(5) The man [who Sue is worried that her sister is planning to marry] lives in Austria.
(6) *The man [who Sue is concerned about her sister [who is planning to marry]] lives in Austria.

With this perspective on structural constraints, the parallel restriction on antecedent-contained ellipsis emerges unproblematically, for the linguistic formulation of the fragment itself provides all that is necessary to predict that ellipsis construal will exhibit sensitivity to such a constraint:

(3) John interviewed every student [who Bill already had].
(4) *John interviewed every student [who Bill ignored the teacher [who already had]].

The fragment in these cases is minimally made up of a determiner (*every*), a nominal (*student*) and a relative pronoun (*who*) initiating a relative clause which contains the ellipsis site (*had*). This sequence has to be processed according to the usual principles governing the processing of relative clauses. First, the noun *student* induces a restrictor for the quantificational term (*every student who...*) which includes a variable (x). Second, a LINKed tree is constructed with the requirement to include this variable as one of the arguments at some node. Initiating the building of this tree a temporarily structurally underspecified node, an unfixed node, is introduced in order to process the relative pronoun which, in surface structure, appears dislocated from the position in which it should be interpreted (as the object of the verb inside the relative). Processing the relative pronoun then annotates this unfixed node with a second copy of the variable x. It is then the underspecified domination relation associated with the unfixed node, $(\langle \uparrow_* \rangle Tn(n))$, which independently imposes the constraint that its position must be resolved within the domain of a single tree, not across another LINKed one as, by definition, there are no dominance relations holding across LINKed structures. :[13]

---

[13] τ-*terms*, (τ, x, Px), are the terms contributed by universal quantifiers like *every*.
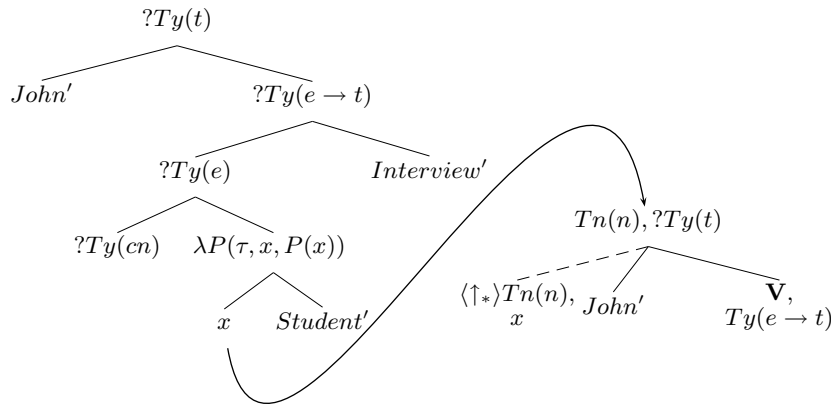
Now, in coming to resolve the metavariable $\mathbf{V}$ which the ellipsis site *had* inside the relative clause has contributed, a sequence of actions from the context has to be retrieved that will result in a subtree of $Ty(e \to t)$. But the choice of which sequence to select is constrained: the selected sequence that will resolve the ellipsis has to conform to the independent restriction on unfixed nodes imposed on the the partial tree already constructed from the fragment. Hence the variable x can only appear in the local tree and not at another LINKed one. This explains the island sensitivity yielding ungrammaticality in (6) where this constraint cannot be satisfied.

Notice the significance of this result: In other frameworks, island constraints would be articulated within the component of syntax, independent of any interpretation considerations, hence not expected to interact with ellipsis construal. In DS, however, with syntax defined in terms of growth of representations of content, such restrictions, modelled as constraints on tree growth, are directly predicted to also constrain ellipsis. This is because ellipsis is modelled as a process of tree growth and is therefore required to conform to any restrictions applying independently to such processes.

5 Fragments in Dynamic Syntax

5.1 Fragments and grammar

As part of this shift into a procedural perspective, [10] defines a concept of *wellformedness* with respect to context, opening the way for arbitrary fragments to be seen as wellformed as long as they occur in a particular environment. Under this definition, fragment construals and the context which they extend can both be partial and dependent on the presence of each other for wellformedness. This provides a basis from which phenomena like (7)–(13) can be analysed using the same mechanisms for structure-building as made available in the core grammar. As we have already discussed, the range of interpretations these fragments receive in actual dialogue seems not to involve well-defined boundaries (see also [58]). We suggest, therefore, that, in being a mechanism for progressive, cumulative growth of information in context, the grammar itself provides the primary means for processing and

integrating such fragments. Computation of the precise contribution of such exchanges to the communicative interaction may need, in addition, pragmatic inferencing (as argued in [58,48]). But this is not a necessary feature of the processing of such fragmental exchanges. These, as modelled in DS, reveal devices specifically externalising the process of hypothesised interpretation as context update and as such inviting confirmation/disconfirmation by the interlocutor (see [30]).

5.2 Processing of fragments

With this in mind, we now turn to the DS account of fragment processing in dialogue. Given that both parsing and generation make use of the very same mechanisms of tree growth, split/joint utterance data are directly predicted. Given the incremental nature of the grammar itself, switch from hearer to speaker is predicted to be possible at any arbitrary point in the dialogue without such fragments having to be interpreted as propositional in type (as is standard elsewhere, e.g. [50]). The parser turned generator simply continues from the partial parse tree that has been established, relative to their own, possibly novel, goal as to how that emergent tree should be completed; and the generator merely loses the initiative, but has already a corresponding partial tree from which to process their respondent's attempt at completing it.

   In DS this basic insight provides the ground for taking completions, clarifications etc. as building on what has previously been said, even though that might be partial, in order to complete/request clarification of some aspect of that previous utterance.[14] Recall examples (7) and (12), repeated below:

(7)   A: Bob left.
       B: the accounts guy, (yeah).
(12)   A: And er they X-rayed me, and took a urine sample,
             took a blood sample.
       A: Er, the doctor
       B: Chorlton?
       A: Chorlton, mhm, he examined me, erm, he, he said now they
             were on about a slide [unclear] on my heart.

Fragments which reformulate an interlocutor A's utterance can occur in two ways: either as (a) non-interruptive confirmations/extensions of A's utterance after the whole of her utterance has been integrated, see (7), or as (b) interruptions of her, A's, utterance, see (12). However, in DS, both are modelled in the same way: as incremental additions. We take these in turn below.[15]

---

[14] Given the DS concept of LINKed trees projecting propositional content, we anticipate that this mechanism will be extendable to fragment construal involving inference (see e.g. [58,59]).

[15] DS makes available variable strategies for the processing of such fragments, we simplify the presentation here but for some alternatives see [29,18].

5.3 Non-interruptive fragments

In (7), B's response *the accounts guy* constitutes both a reformulation of A's utterance, and an extension of A's referring expression, in effect providing the appositive expression 'Bob, the accounts guy'. B has presumably processed A's original utterance and achieved some identification of the individual associated with the name *Bob*: that is to say, B has constructed a full content representation for this utterance. In DS terms, B's context after processing A's utterance contains the following tree:[16]

(27) B's CONTEXT:

$$Ty(t), Leave'(\iota, x, Bob'(x)), \diamondsuit$$

$$(\iota, x, Bob'(x)) \qquad Leave'$$

B can now re-use this contextual representation as the point of departure for generating the expression *the accounts guy*, using the apposition mechanism defined in section 4.3. In this case his own GOAL TREE, the message he wishes to convey, will be annotated at the relevant node with a composite term made up both from the term recovered from parsing A's utterance and his new addition:[17]

(28) B's GOAL TREE for uttering *the accounts guy*:

$$Ty(t), Leave'(\iota, x, Bob'(x) \wedge Leave'(x))$$

$$Tn(n), (\iota, x, Bob'(x) \wedge Acc.guy'(x)) \qquad Leave'$$

$$\langle L^{-1} \rangle Tn(n),$$
$$(\iota, x, Acc.guy'(x))$$

In order for B to produce the fragment *the accounts guy* relative to this GOAL TREE, according to DS, he has to incrementally test-parse the fragment first and check that the subsumption relation is preserved between each parsing transition and his GOAL TREE. Simplifying for illustration purposes, the (test-)parsing steps include: (a) moving the pointer to the relevant node in the context tree (27), (b) initiating a LINKed tree from this node and moving the pointer there and (c) processing the fragment as an apposition to yield its content as an iota term which will be subsequently unified with the content of the already present iota term in the context:

---

[16] For simplicity of illustration we do not show the internal tree-structure of epsilon/iota terms on the graphics.

[17] Items like *yeah* have a metacommunicative function in dialogue (*backchannels*) and are not therefore included as part of the main DS propositional content.

(29) B's test parse (1): moving the pointer

$$Ty(t), Leave'(\iota, x, Bob'(x))$$

$$(\iota, x, Bob'(x)), \Diamond \quad Leave'$$
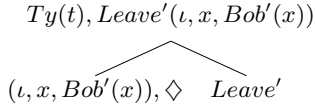
(30) B's test parse (2): LINK adjunction [18]

$$Ty(t), Leave'(\iota, x, Bob'(x))$$

$$Tn(n), (\iota, x, Bob'(x)) \qquad Leave'$$

$$\langle L^{-1} \rangle Tn(n), \Diamond$$
$$?Ty(e), ?\langle \downarrow_* \rangle x$$
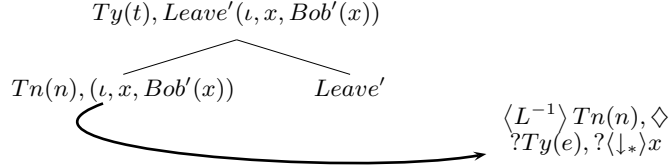
(31) B's test parse (3): processing *the accounts guy*

$$Ty(t), Leave'(\iota, x, Bob'(x))$$

$$Tn(n), (\iota, x, Bob'(x)) \qquad Leave'$$

$$\langle L^{-1} \rangle Tn(n), \Diamond$$
$$Ty(e), \cancel{?Ty(e)},$$
$$(\iota, x, Acc.guy'(x)) \cancel{?\langle \downarrow_* \rangle x}$$
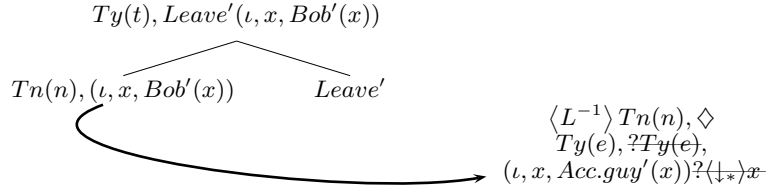
(32) B's test parse (4): LINK evaluation

$$Ty(t), Leave'(\iota, x, Bob'(x))$$

$$Tn(n), \Diamond,$$
$$(\iota, x, Bob'(x) \wedge Acc.guy'(x)) \qquad Leave'$$

$$\langle L^{-1} \rangle Tn(n),$$
$$Ty(e),$$
$$(\iota, x, Acc.guy'(x))$$
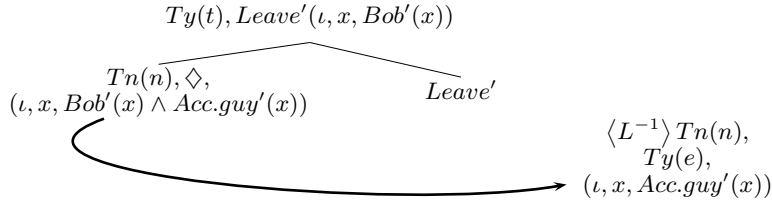
If all these steps generate structure that subsumes the GOAL TREE, (28), license to produce the words *the accounts guy* will ensue[19] – as is the case with the indicated derivation. Since in this case B's content corresponds to that intended by A, and the resulting LINKed trees are therefore consistent with each other, production of the reformulation will also have the indirect effects

---

[18] $?\langle \downarrow_* \rangle x$ is the requirement indicating that a copy of the variable at the LINK-initiating node must appear inside the LINKed tree.
[19] In fact, given the incrementality of DS, each single word is uttered individually upon the subsumption check but we suppress these steps here for simplicity.

of acknowledgement, confirmation of B's understanding of A's utterance, agreement etc.

We now turn to fragments that directly exploit the incremental nature of the parser/grammar, namely, fragments whose production interrupts the original speaker at some sub-sentential point.

5.4 Interruptive Clarification

In the acknowledgement case (7), the tree relative to which the LINKed structure is built is complete; but the very same mechanism can be used when the interlocutor needs clarification, and the tree being built is still partial. In (12), repeated below, B has built only a partial tree at the point of interruption:
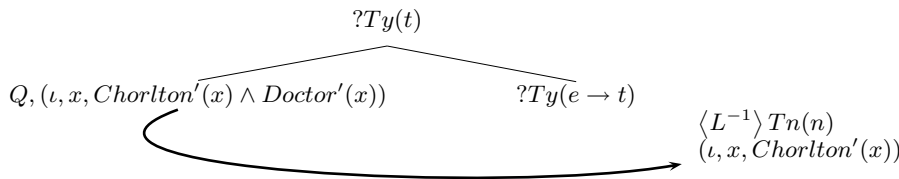
(12)   A: And er they X-rayed me, and took a urine sample,
            took a blood sample.
       A: Er, the doctor
       B: Chorlton?
       A: Chorlton, mhm, he examined me, erm, he, he said now they
            were on about a slide [unclear] on my heart.

(33)   B's CONTEXT after parsing *the doctor*:

$$?Ty(t)$$

$$\Diamond, (\iota, x, Doctor'(x)) \qquad\qquad ?Ty(e \to t)$$
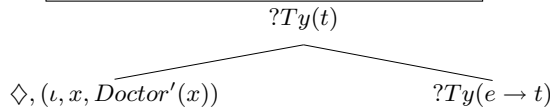
In order to request clarification of the intended referent of *the doctor*, B again takes as his GOAL TREE a tree decorated with an expansion of the term constructed from parsing A's utterance, as a means of requesting more specific information to aid in the task of identifying who is being talked about. The fact that this time B's GOAL TREE is partial (he has not completed the parse of a full proposition before asking for clarification) causes no problem for the analysis:[20]
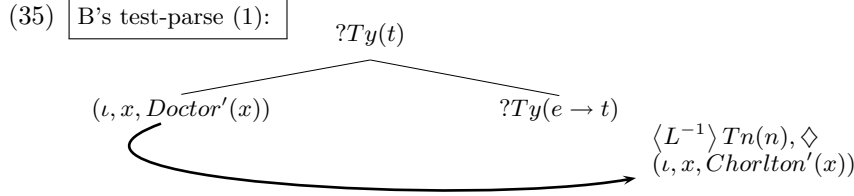
(34)   B's GOAL TREE:

$$?Ty(t)$$

$$Q, (\iota, x, Chorlton'(x) \wedge Doctor'(x)) \qquad\qquad ?Ty(e \to t)$$

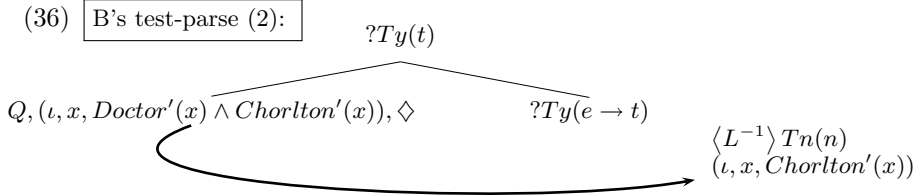$$\left\langle L^{-1} \right\rangle Tn(n)$$
$$(\iota, x, Chorlton'(x))$$

Using the very same mechanism as in (7) of building a LINKed tree constrained to induce shared terms, B can generate (and A can parse) the name

---

[20] We ignore here any discusion of question-hood, apart from the annotation Q on the relevant nodes, since our emphasis is on common mechanisms. See [29] for preliminary discussion.

Chorlton, no matter that neither has completed the parse tree for A's original (unfinished) utterance. This name, contributing a term with the restrictor that the individual picked out must be named 'Chorlton', is used to decorate the LINKed node:

(35) B's test-parse (1):

$$?Ty(t)$$

$$(\iota, x, Doctor'(x)) \qquad\qquad ?Ty(e \to t)$$

$$\langle L^{-1} \rangle Tn(n), \Diamond$$
$$(\iota, x, Chorlton'(x))$$

The outcome of this process, when the LINKed structure is evaluated, is a composite term $(\iota, x, Doctor'(x) \wedge Chorlton'(x))$ at the node at which the LINKed tree was attached, extending the initial iota term. As this subsumes the GOAL TREE of (34), the name *Chorlton* is licensed to be uttered:

(36) B's test-parse (2):

$$?Ty(t)$$

$$Q, (\iota, x, Doctor'(x) \wedge Chorlton'(x)), \Diamond \qquad\qquad ?Ty(e \to t)$$

$$\langle L^{-1} \rangle Tn(n)$$
$$(\iota, x, Chorlton'(x))$$

This extension of the term is confirmed by A, this time trivially replicating the composite term derived from B's utterance (see [29] for discussion). This process, therefore, is identical to that employed in processing B's utterance in (7), though to rather different effect, a clarification request and reply, at this intermediate stage in the interpretation process. The eventual effect of the process of inducing LINKed structures to be decorated by coreferential type e terms may thus vary across monologue and different dialogue applications, but the mechanism is the same.
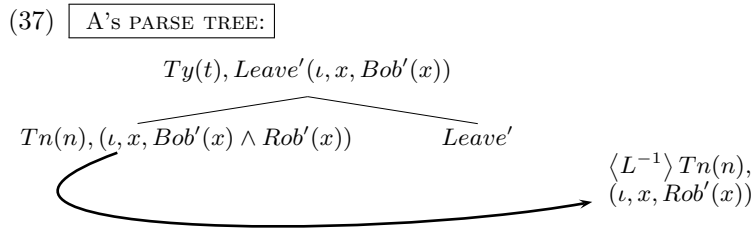
5.5 Correction

It might be argued nonetheless that the phenomenon of correction as a repair of the interlocutor's understanding is intrinsically a dialogue phenomenon having to do with metarepresentation of the other's information state. However, from our point of view, this is not a necessary conclusion. In some cases, the recognition that some information presented by the interlocutor (or indeed assumed temporarily by the agent themselves) is inconsistent with other assumptions in the context requires an inferential step providing access to the parser's more general knowledge base. However, it does not require metarepresentation of the other's information state in that this inconsistency is

locally resolvable through the information that the interlocutor themselves have in their immediate context. Suppose that B mishears and requests confirmation of what he has perceived A as saying, but that he is mistaken; and A in turn rejects B's utterance and provides more information:

(4) A: Bob left.
     B: Rob?
     A: (No,) (Bob,) the accounts guy.

Of course, A can process B's clarification request in exactly the same way as set out in section 5.4 above, as an extension of her own context via LINKed tree construction. This leads to a representation as follows:

(37) $\boxed{\text{A's PARSE TREE:}}$

$$Ty(t), Leave'(\iota, x, Bob'(x))$$

$$Tn(n), (\iota, x, Bob'(x) \wedge Rob'(x)) \qquad Leave'$$

$$\langle L^{-1} \rangle Tn(n),$$
$$(\iota, x, Rob'(x))$$

In order for A to establish that the information here leads to inconsistency (the set denoted by $Bob'(x) \wedge Rob'(x)$ is empty, the individual so described not having two such names), she has to be able to retrieve information independently assumed by her (again this too needs a specified interface with an inference model that is not provided here). But assuming that this is available, the tree can be recognised as specifying information that is inconsistent, which would lead to rejection. Rejection is therefore analysed here as simple disagreement: B's utterance has been understood, but simply judged as incorrect.

To generate her subsequent correction, A need only establish as the current most recent representation in context the tree A originally built by generating *Bob left* (the most recent *consistent* tree available). This can be monotonically achieved by recovering and copying this tree[21] to serve as the current most immediate context. Generating the corrective *Bob, the accounts guy* then proceeds exactly as in the previous sections. Note that this analysis does not assume that B's (mistaken) utterance content is at any stage (non-monotonically) removed from the context, even for A: corrected representations must be maintained in the context as they can provide antecedents for subsequent anaphoric expressions, as in:[22]

(38) A: Bob left.
     B: Rob?
     A: No. HE'S in Beijing these days. Bob, the accounts guy.

---

[21] As an anonymous reviewer pointed out this tree will not include the LINKed node which led to the inconsistency so this context tree can now be extended consistently to extend the description of the particular individual intended.

[22] It is notably harder to recover ellipsis construal appropriately across an intervening utterance, but it is by no means impossible (see [24]).

In conclusion, these fragments and their construal have demonstrated that, despite serving distinct functions in dialogue – as acknowledgements, corrections, extensions, etc – the mechanisms that underpin these distinct functions are nonetheless general strategies for tree growth that are independently available.

6 Summary and Evaluation

Even though DS is a grammar formalism and so not, in principle, providing a full theory of either utterance understanding or dialogue interactivity, in closing, it is of some interest to reflect whether, nevertheless, there are clear advantages in adopting this model for language processing.

6.1 Context, grammar and dialogue

In DS, context directly interacts with the parsing process at any point and both speaker and hearer can presume on it. Hence, for either speaker or hearer, any reiteration of what is provided by the context may be unnecessary, given that some fragment and the state of the parser at this point provides sufficient information to achieve a complete interpretation. Modelling both fragment resolution and the transition between speakers as the transition between parse states means being able to capture the dialogue dynamics more directly via key aspects of the grammar. Firstly, a distinct advantage for this stance is a continuum discerned from what are standardly seen as grammar-internal phenomena (e.g. VP-ellipsis) to what are usually taken as distinct dialogue phenomena (fragments such as clarifications, extensions, reformulations and corrections) which, in some accounts, are even considered as alien data of "performance". Under the modelling offered by DS, there simply is no essential difference: mechanisms for interpretation apply equally intra-sententially, inter-sententially, and across participants.

Secondly, in all cases, the advantage which use of ellipsis provides is a "least effort" means of re-employing previous content/structure/actions which constitute the context: hence its prevalent use in conversational dialogue is expected. It is not merely that fewer words are used in such minimal utterances, hence preferred on a trivial cost basis. More importantly, in fact, in such elliptical re-iterations, a whole sequence of, in principle, independent production/parsing choices are already uniquely resolved by the way in which such decisions were taken in the processing of the antecedent string. Such a determinism means that, with information culled from context, representations do not have to be constructed afresh via costly processes of lexical retrieval, choice of alternative parsing strategies, etc.

6.2 Grammar and the disambiguation challenge

The account thus also sheds new light on the grammar-parser contribution to disambiguation and the complex linking between contextual and linguis-

tic information (see also [31]). Given the fine-grainedness of the DS model regarding the question of how interlocutors link current utterances with previous (discourse) contextual information, the familiar challenge of facing the problem of disambiguation opens up. The occurrence of recurrent, often overlapping fragments, as in (12) might seem to raise the issue of how to manage the step-wise multiplying interpretive and structural options in the processing of such ellipses. Two features of DS are crucial in responding to this challenge: (a) *incrementality* and (b) parsing/generation *bi-directionality*.

The DS account enables one to express how interlocutors are able to exploit the inherent incrementality afforded by the grammar to manage the rapid increase in available processing options. By employing fragments incrementally in the build up of construal, hearers are able to immediately respond to a previous utterance at any point in the construction process, hence sub-sententially as well as sententially. In consequence, interlocutors can constrain each other's interpretation choices in an ongoing way, by clarification, acknowledgement etc., during the construction of even a single propositional formula. The modelling of this aspect of human interaction, which is fundamental to its efficiency, is not open to more conventional sentence-based frameworks where the locus of context dependency of linguistic processing is external to the core grammatical resources. Further, incremental processing of fragments allows reduction of uncertainty as regards the structural antecedent of the elliptical element. This is because the fragment is processed while the parse state indicates exactly the intended antecedent (through the position of the DS pointer which is located at the relevant node). In this respect, no computation of "saliency" or "accessibility" of potential antecedents is necessary in such an incremental account, hence another source of complexity in parsing is minimised.

The parsing/generation bi-directionality of DS, in addition, makes it straightforward to model switching between speaker and hearer. And the ease of this switching at sub-sentential points is crucial for efficient participant coordination. Indeed, in split exchanges, the parse tree transparently reveals exactly where need of clarification or miscommunication may be arising, as it will be at that node from which a sub-routine extending it takes place. According to the DS model of generation, repeating or extending a constituent of an interlocutor's utterance is licensed only if the current GOAL TREE matches or extends a parse tree updated with the relevant subpart of that utterance. Indeed, this update is what the interlocutor is seeking to clarify, correct or acknowledge. So an interlocutor can reuse the already constructed (partial) parse tree in their context, thereby starting at this point, rather than having to construct an entire propositional tree or subtree (e.g. of type e). With a continuous cycle of contribution-response-contribution, the effect of clarifications and the like, despite appearing to indicate misunderstanding or "repair", is in fact to narrow the focus to a specific point of query, enabling interlocutors to make quite fine-grained adjustments to their own understandings (see also [23, 20]).

So not only does the reflection of the time-linear parsing dynamics in the grammar formalism narrows down the competence-performance gap by definition, it also serves to directly address the complexity issues normally

taken to be a performance consideration associated with parser/generator design.

6.3 Cognitive modelling

A further advantage of modelling dialogue directly in terms of grammatical resources is that attractive properties of the latter are potentially available for the former. This connection has been indicated above for a range of dialogue phenomena regarding the advantages of incrementality in reducing complexity and advancing efficiency. Of further interest is the way in which the dynamics of the grammar formalism might also be potentially directly available to the dialogue system, reflected here in the way that dialogue is being modelled as essentially the chaining of transitions between parse states (of successive speakers). Being able to directly draw on grammatical resources in this way means avoiding a model of interlocutor coordination via external mechanisms superimposed on a mode-neutral grammar formalism. From the DS viewpoint, the key mechanisms involved in agent coordination are essentially system-internal to the grammar and thus underpin the dialogue model directly (see also [30]).

As regards the tight coupling of parsing and generation proposed by DS, parity of representations and mechanisms across coordinated behaviours is not an isolated idea in the cognitive modelling literature. Parity between speaking and hearing has prompted cognitive neurodynamic research, for instance, to model turn-taking as a kind of coupled oscillation of the cognitive processes across speakers and hearers (e.g. [69,5]). In a similar vein, here we are aiming to characterise part of the dynamics underlying such coordination in dialogue, without appealing to higher-level processes such as mental state modelling (contra [12,58,48], see also [27]). In modelling the transition between speakers as the transition between parse states defined grammar-internally, fundamental aspects of dialogue are modelled in a largely mechanistic manner (e.g. [47]), a move which is echoed in emerging results in cognitive science more generally.[23] This opens up the possibillity of characterising language as intrinsically a set of mechanisms for communicative interaction, where it is the application of these mechanisms in ongoing conversational exchange that yields the effect of coordinated activity without necessary high-level (meta)-representation by either party of their interolocutors' beliefs.

---

[23] There is a range of results linking action and perception within a common framework (e.g. [25,26]), on how various cognitive mechanisms for "sharing" representations may facilitate joint action [60,61], on research into common representations underlying both speaking and hearing (e.g. [39,40]), and on imitation as "behavior parsing" [7,8]. Research into such parity between action and perception is now well-established (see e.g. [68]), and this is beginning to be reflected in work on the role of such mechanisms in communication (e.g. [1]).

References

1. Arbib, M.: From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. Behavioral and Brain Sciences **28**, 105–167 (2005)
2. Atterer, M., Schlangen, D.: RUBISC - a robust unification-based incremental semantic chunker. In: Proceedings of SRSL 2009, the 2nd Workshop on Semantic Representation of Spoken Language, pp. 66–73. Association for Computational Linguistics, Athens, Greece (2009). URL http://www.aclweb.org/anthology/W09-0509
3. Barwise, J., Cooper, R.: Generalized quantifiers and natural language. Linguistics and Philosophy **4**, 159–219 (1981)
4. Blackburn, P., Meyer-Viol, W.: Linguistics, logic and finite trees. Bulletin of the IGPL **2**, 3–31 (1994)
5. Bonaiuto, J., Thrisson, K.: Towards a neural model of realtime turntaking in face-to-face dialogue. In: I. Wachsmuth, M. Lenzen, G. Knoblich (eds.) Embodied Communication in Humans And Machines. Oxford University Press (2008)
6. Bunt, H.: Multifunctionality and multidimensional dialogue semantics. In: Proceedings of DiaHolmia, 13th SEMDIAL Workshop (2009)
7. Byrne, R.: Imitation without intentionality. using string parsing to copy the organization of behaviour. Animal Cognition **2**, 63–72 (1999)
8. Byrne, R.W.: Imitation as behaviour parsing. Philos Trans R Soc Lond B Biol Sci. **358(1431)**, 529536 (2003)
9. Cann, R., Kempson, R., Marten, L.: The Dynamics of Language. Elsevier, Oxford (2005)
10. Cann, R., Kempson, R., Purver, M.: Context and well-formedness: the dynamics of ellipsis. Research on Language and Computation **5**(3), 333–358 (2007)
11. Chomsky, N.: Aspects of the Theory of Syntax. MIT, Cambridge, MA (1965)
12. Clark, H.H.: Using Language. Cambridge University Press (1996)
13. Costa, F., Frasconi, P., Lombardo, V., Soda, G.: Towards incremental parsing of natural language using recursive neural networks. Applied Intelligence **19**(1-2), 9–25 (2003)
14. Dalrymple, M., Shieber, S.M., Pereira, F.C.N.: Ellipsis and higher-order unification. Linguistics and Philosophy **14**(4), 399–452 (1991)
15. Evans, G.: Pronouns. Linguistic Inquiry **11**(2), 337–362 (1980)
16. Fernández, R.: Non-sentential utterances in dialogue: Classification, resolution and use. Ph.D. thesis, King's College London, University of London (2006)
17. Fiengo, R., May, R.: Indices and Identity. MIT Press, Cambridge, Mass. (1994)
18. Gargett, A., Gregoromichelaki, E., Howes, C., Sato, Y.: Dialogue-grammar correspondence in dynamic syntax. In: Proceedings of the 12th SEMDIAL (LONDIAL) (2008)
19. Gibson, E.: Linguistic complexity: Locality of syntactic dependencies. Cognition **68**, 1–76 (1998)
20. Ginzburg, J.: The Interactive Stance: Meaning for Conversation. CSLI (forthcmg)
21. Ginzburg, J., Cooper, R.: Clarification, ellipsis, and the nature of contextual updates in dialogue. Linguistics and Philosophy **27**(3), 297–365 (2004)
22. Ginzburg, J., Sag, I., Purver, M.: Integrating conversational move types in the grammar of conversation. In: P. Kühnlein, H. Rieser, H. Zeevat (eds.) Perspectives on Dialogue in the New Millennium, *Pragmatics and Beyond New Series*, vol. 114, pp. 25–42. John Benjamins (2003)
23. Healey, P.: Interactive misalignment: The role of repair in the development of group sub-languages. In: R. Cooper, R. Kempson (eds.) Language in Flux. College Publications (2008)

24. Healey, P., Eshghi, A.: What is a conversation? distinguishing dialogue contexts. In prep, Queen Mary University of London
25. Hommel, B., Msseler, J., Aschersleben, G., Prinz, W.: The theory of event coding (tec): A framework for perception and action planning. Behavioural and Brain Sciences **24**, 849–937 (2001)
26. Hurley, S.: The shared circuits hypothesis: A unified functional architecture for control, imitation, and simulation. In: S. Hurley, N. Chater. (eds.) Perspectives on Imitation: From mirror neurons to memes. MIT Press (2005)
27. Kecskes, I., Mey, J. (eds.): Intention, Common Ground and the Egocentric Speaker-Hearer. Mouton de Gruyter (2008)
28. Keenan E., M.B.M.D.: Pragmatics in memory: a study of natural conversation. Journal of Verbal Learning and Behaviour **16**, 549560. (1977)
29. Kempson, R., Gargett, A., Gregoromichelaki, E.: Clarification requests: An incremental account. In: Proceedings of the 11th Workshop on the Semantics and Pragmatics of Dialogue (DECALOG) (2007)
30. Kempson, R., Gregoromichelaki, E., Purver, M., Mills, G., Gargett, A., Howes, C.: How mechanistic can accounts of interaction be? In: Proceedings of Diaholmia, the 13th Workshop on the Semantics and Pragmatics of Dialogue (2009)
31. Kempson, R., Gregoromichelaki, E., Sato, Y.: Incrementality, speaker-hearer switching and the disambiguation challenge. In: Proceedings of SRSL 2009, the 2nd Workshop on Semantic Representation of Spoken Language, pp. 74–81. Association for Computational Linguistics, Athens, Greece (2009). URL http://www.aclweb.org/anthology/W09-0510
32. Kempson, R., Meyer-Viol, W., Gabbay, D.: Dynamic Syntax. Blackwell, Oxford (2001)
33. Kintsch, W., Bates, E.: Recognition memory for statements from a classroom lecture. Journal of Experimental Psychology: Human Learning and Memory **3**(2), 150–159 (1977)
34. Ladusaw, W.A.: Polarity sensitivity as inherent scope relations. Ph.D. thesis, University of Massachusetts, Amherst, Amherst, MA (1979)
35. van Leusen, N., Muskens, R.: Construction by description in discourse representation. In: J. Peregrin (ed.) Meaning: The Dynamic Turn, chap. 12, pp. 33–65 (2003)
36. Levinson, S.: Activity types and language. In: P. Drew, J. Heritage (eds.) Talk at Work, pp. 66–100. Cambridge University Press (1992)
37. Levinson, S.C.: Pragmatics. Cambridge Textbooks in Linguistics. Cambridge University Press (1983)
38. Lewis, R.L., Vasishth, S.: An activation-based model of sentence processing as skilled memory retrieval. Cognitive Science **29**, 1–45 (2005). URL http://www.ling.uni-potsdam.de/ vasishth/Papers/cogsci05lewisvasishth.pdf
39. Liberman, A., Mattingly, I.: A specialization for speech perception. Science **243**(4890), 489–494 (1989)
40. Liberman, A., Whalen, D.: On the relation of speech to language. Trends in Cognitive Sciences **4**(5), 187–196 (2000)
41. Merchant, J.: The Syntax of Silence: Sluicing, Islands, and the Theory of Ellipsis. Oxford University Press, Oxford (2001)
42. Merchant, J.: Three kinds of ellipsis: Syntactic, semantic, pragmatic? (2007). Ms University of Chicago
43. Morgan, J.: Sentence fragments and the notion 'sentence'. In: Y.M. B. Jashru R. Lees (ed.) Issues in Linguistics, pp. 719–751. University of Illinois Press, Urbana (1973)
44. Morgan, J.: Sentence fragments revisited. In: R. Chametzky (ed.) Proceedings of the 25th Meeting of the Chicago Linguistics Society. CLS (1989)
45. Ono, T., Thompson, S.: What can conversation tell us about syntax? Amsterdam Studies in the Theory and History of LInguistic Sciences 4 pp. 213–272 (1995)
46. Pickering, M., Garrod, S.: Toward a mechanistic psychology of dialogue. Behavioral and Brain Sciences **27**, 169–226 (2004)

47. Pickering, M., Garrod, S.: Toward a mechanistic psychology of dialogue. Behavioral and Brain Sciences **27**, 169–226 (2004)
48. Poesio, M., Rieser, H.: Completions, coordination, and alignment in dialogue (to appear). Ms.
49. Polanyi, L., Scha, R.: A syntactic approach to discourse semantics. In: Proceedings of the 10th International Conference on Computational Linguistics (COLING10), 413 419. Stanford, CA. (1984)
50. Purver, M.: The theory and use of clarification requests in dialogue. Ph.D. thesis, University of London (2004)
51. Purver, M., Cann, R., Kempson, R.: Grammars as parsers: Meeting the dialogue challenge. Research on Language and Computation **4**(2-3), 289–326 (2006)
52. Purver, M., Howes, C., Gregoromichelaki, E., Healey, P.: Split utterances in dialogue: a corpus study. In: Proceedings of SigDial 2009 (2009)
53. Ross, J.: Constraints on variables in syntax. Ph.D. thesis, Massachusetts Institute of Technology (1967)
54. Rühlemann, C.: Conversation in context: a corpus-driven approach. Continuum (2007)
55. Sacks, H., Schegloff, E., Jefferson, G.: A simplest systematics for the organization of turn-taking for conversation. Language pp. 696–735 (1974)
56. Schegloff Emanuel A., .S.H.: Opening up closings. Semiotica **7**(4), 289–327. (1973)
57. Schegloff, E.: The relevance of repair to syntax-for-conversation. Discourse and syntax **12**, 261–88 (1979)
58. Schlangen, D.: A coherence-based approach to the interpretation of non-sentential utterances in dialogue. Ph.D. thesis, University of Edinburgh (2003)
59. Schlangen, D., Lascarides, A.: The interpretation of non-sentential utterances in dialogue. In: Proceedings of the 4th SIGdial Workshop on Discourse and Dialogue, pp. 62–71. Association for Computational Linguistics, Sapporo, Japan (2003)
60. Sebanz, N., Bekkering, H., Knoblich, G.: Joint action: Bodies and minds moving together. Trends in Cognitive Sciences **10**, 70–76 (2006)
61. Sebanz, N., Knoblich, G., Prinz, W., Wascher, E.: Twin peaks: An ERP study of action planning and control in co-acting individuals. Journal of Cognitive Neuroscience, **15**, 99–104 (2006)
62. Skantze, G., Schlangen, D.: Incremental dialogue processing in a micro-domain. In: Proceedings of the 12th Conference of the European Chapter of the ACL (EACL 2009), pp. 745–753. Association for Computational Linguistics, Athens, Greece (2009). URL http://www.aclweb.org/anthology/E09-1085
63. Sperber, D., Wilson, D.: Relevance: Communication and Cognition, second edn. Blackwell (1995)
64. Stainton, R.: Words and Thoughts: Subsentences, Ellipsis, and the Philosophy of Language. Oxford University Press (2006)
65. Stalnaker, R.: Context and Content. Oxford University Press (1999)
66. Stoness, S., Tetreault, J., Allen., J.: Incremental parsing with reference interaction. In: In ACL Workshop on Incremental Parsing, pp. 18–25 (2004)
67. Sturt, P., Crocker, M.: Monotonic syntactic processing: a cross-linguistic study of attachment and reanalysis. Language and Cognitive Processes **11**, 448–494 (1996)
68. various: Special issue on the brain mechanisms of imitation learning. Neural Networks **19** (2006)
69. Wilson, M., Wilson, T.: An oscillator model of the timing of turn-taking. Psychonomic Bulletin and Review pp. 957–968 (2005)