9
Syntax as the dynamics of language understanding

Ruth Kempson

9.1 Grammar and language processing: building structure in context

This chapter sets out the case for setting linguistic theory and the grammars linguists write in a perspective reflecting real-time processing, as offered by the Dynamic Syntax framework. The assumptions advocated by this framework represent a radical departure from the methodology adopted by all who have worked on the formal modelling of natural language since the Chomskyan revolution of the 1960s. According to the familiar methodology a language is assumed to be in some sense isolatable, taken to be the set of well-formed sentences of the language taken in abstraction from their use. A grammar for a language is then a set of rules or principles which associate with each such sentence a triple: the phonological properties of the words made available by some underlying phonological system; the projection of fixed structure definable over those words (syntax); the meaning of the string defined over that syntactic structure (semantics). All attributes of language use in real time are treated as external to the system and explained by ‘performance’ theories to be grounded in such ‘competence’ grammars. What is not often pointed out is that this methodology omits any explanation of the two properties arguably diagnostic of natural language: (1) natural languages are systems underpinning an activity that takes place in real time (this is precluded in principle); (2) their token structures and interpretation display systemic dependence on the context within which such activity takes place (at best, this gets only a partial characterisation). Furthermore, the methodology commits the linguist to an encapsulated system, isolated from other cognitive systems.

On the new view which this chapter will introduce, the domain-specific concept of syntax is abandoned. So too is the assumption that syntactic tree structures are defined as holding over a sequence of words. To the contrary, words, and the general procedures encoded in language, are mechanisms for building up conceptual representations of content in an incremental left-to-right fashion, and it is only conceptual representations which have structure attributed to them. The consequence of this perspective shift is a much leaner perspective, in which the competence system is much closer though not identical to the performance system it underpins. The multiple levels of a grammar formalism with distinct phonology, morphology, morphosyntax, multilevelled syntax are replaced by two levels:
phonology and representations of content, with mechanisms defined in terms of some underspecified input and updates from that input that in combination yield mappings from phonological input onto representations of content. All are defined in terms of their contribution to the progressive building up of content, carrying over to syntax the assumptions of underspecification and context-relative update that are familiar from semantics and pragmatics (Kamp and Reyle 1993; Sperber and Wilson 1995; and many others). In consequence, the natural language system as a whole becomes dynamic. The system is a specification of general and lexically stored procedures which users can implement to build up such representations, both as hearers or as speakers.

9.1.1 Background perspectives

This shift of perspective came from two independent sources. It came initially in response to the increasing unease displayed by semanticists addressing the context-dependence of language interpretation, for which frameworks were needed which could directly reflect the incremental unfolding of representations of content. Then there was the realisation that incorporating this aspect of performance into the grammar transforms well-known syntactic puzzles, dissolving the difficulties they pose. In this chapter, I shall take ellipsis in conversational dialogue as our case study, to illustrate both issues.

Ever since model-theoretic semantics was articulated by Montague in the late 1960s and early 1970s, it was apparent that any semantic explication of language has to reflect the context-relativity of its construal; and over the years, it became increasingly clear that this context-relativity of language is endemic. Encoded meaning of an expression can always be enriched to yield some context-specific interpretation; so the articulation of that input meaning must reflect its underspecification with respect to the assignable output interpretation. This is not merely a one-shot context relativism definable from some characterisation of denotational truth-conditional content constituting sentence-meaning as the output of the grammar. Every step of the way involves context dependency, in anaphora, tense and aspect construal, quantifiers, adjuncts etc. Moreover, there is apparent necessity to invoke concepts of semantic representation over and above any such syntactic structures that may be defined over sequences of those expressions, so the directness of mapping from sentence-string onto model-theoretically definable content can no longer be presumed. The lead move in defining a structured concept of content within formal semantics was Hans Kamp’s Discourse Representation Theory (DRT: Kamp and Reyle 1993), whose aim was to provide a unitary characterisation of pronominal anaphora. Kamp and colleagues provided a single base from which pronoun construal was defined for such varied environments as (1)–(3) which show that the anaphoric linkage can either be inter-sentential and able to be independent, or intra-sentential and dependent on another quantifier:

(1) A colleague is arriving late from New York. He will be tired by the end of the day.

(2) If a colleague is arriving late from New York, he will be tired by the end of the day.

(3) Every time a colleague arrives late from New York, he will be tired by the end of the day.

To obtain the appropriate basis for generality, a level of discourse representation structure (DRS) was defined via a mapping from whatever syntactic structure might be articulated
Syntax as the dynamics of understanding

over the elements in the string. The semantics for this system is incremental, but only sentence-by-sentence.

Even granting the problem that such anaphoric links appear not to be restricted by sentence boundaries, the question is what status the intermediate DRS should have in the grammar, being merely an intermediate construct which is given denotational substance only by the subsequent semantics. Inevitably, there were attempts to reduce all such constructs to model-theoretic alternatives in which the non-determinism of construal is expressed in the semantics (Groenendijk and Stokhof 1991). The problem of apparent need to posit some intermediate representation between the sentence form and its semantic evaluation is, however, not restricted to pronominal anaphora. It arises in temporal and modal construal, quantification, and strikingly in ellipsis, where truncated sequences can be interpreted as though these elliptical forms were, in some sense, a vehicle for a propositional mode of construal, some cases indeed explicitly marked as being underspecified for some predicate to complete such a proposition. Like anaphora, ellipsis dependencies may ignore sentence boundaries:

(4) If John is writing a book, then Bill won’t be.
(5) Every time John writes a book, Bill does.
(6) John is writing a book. An article too.

Ever since the various phenomena comprising ellipsis were identified, there have been arguments as to whether ellipsis falls within the remit of syntax. If so, then ellipsis will not provide evidence for any additional level of semantic representation. But this move led to the identification of a wealth of different ellipsis types and posited ambiguities (Fiengo and May 1984), so semanticists sought to develop accounts defining the variety via processes of abstraction over some preceding content re-applied at the ellipsis site to yield the various effects (Dalrymple et al. 1991). However, if ellipsis is deemed to be subject to semantic characterisation, it would seem that an additional level of semantic representation is required for full explication of ellipsis. This is because, amongst other arguments, there are fragments which obligatorily reflect morphosyntactic constraints of the language. Fragmentary Greek answers to a question for example have to reflect the case appropriate to a full reply, a pattern widespread in case-rich languages (7).

(7) I Maria den egrapse to grama? Oxi, ego/*emena

Didn’t Maria write the letter? No, I_Nom/*me_Acc

So ellipsis, where the basis for recovering interpretation is radically incomplete, provides evidence that there must be some concept of building up a representation of content with granularity reflecting syntactic/morphological detail.

Yet adding an additional level to the grammar does not address the problem of cross-sentential forms of ellipsis; and further problems with intra-sentential ellipsis abound. In conversational dialogue, well-formed sentence structures may emerge through a sequence of fragmentary contributions, each participant adding some fragment to a partial structure that emerges bit by bit across the exchange. As we shall see, this interaction depends on participants alternating listening and speaking, a phenomenon which grammars respecting the traditional sentence-based methodology are very poorly placed to explain. All such
sub-sentential interactions, however fluent and widespread, are beyond the remit of grammar, dismissed as mere performance dysfluencies. The result is that not only do such grammars fail to explain the systemic nature of context-dependence, a diagnostic property of natural language, but no morphosyntactic, syntactic and semantic dependency will be completely characterised.

There is a simple but radical solution. This is to shift to the assumption that syntax is the set of mechanisms inducing the dynamic ongoing projection of semantic representation: there is no separate level of syntactic representation over and above that of such emergent semantic representations. So it is not the level of semantic representation that is jettisoned: it is the old-style conception of syntax. Interpretation involves evolving growth of partial representations, secured through sequences of structure-building actions, to which any participant can contribute. The patterns of dialogue ellipsis follow directly, wholly without stipulation. This chapter seeks to give the reader a taste of this issue, introducing the data and then seeing how the dynamics of processing as modelled within DS directly secures the effects of interaction, this success illustrating how we can successfully model our linguistic competence as grounded in the dynamics of how people do what they do.

### 9.1.2 Split utterances in dialogue: the challenge of incrementality

In dialogue, switching of roles between speaking and hearing, across and within sentential structures, is widespread. There is much interactivity between participants in a dialogue, with people joining in on what the other person is saying, adding to it (8), often helping the other to find the appropriate add-on (9).

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

(9) A: I suppose I need a a
   B: mattock. For breaking up clods of earth.

Each such contribution can add unproblematically to whatever partial structure has been set out so far without any contributor having planned the overall structure at the outset. (10) turns out to be a conditional sentence only upon its completion, its consequent clause preceding its antecedent clause:

(10) Hugh: We’re going to London
   Alex: to see Granny
   Eliot: with the dogs?
   Hugh: if you promise to keep them under control.

Such speaker-switching may take place at what might be construed as sentence boundaries, so that both antecedent and add-on fragment might be said to be separate sentences, with the fragment trading on some proposition being recoverable from context. But this is not necessary: one can interrupt at any point in a clausal sequence, and by no means only collaboratively (11).

(11) A: What this shows is
    B: that you have completely missed the point.
Split dependencies can, moreover, be arbitrarily complex (12).

(12)  
A: Has every student handed in  
B: their homework?  
A: or even any assignments?

(12) contains a pronoun uttered by B interpreted as bound by a quantifying term uttered by A and a negative polarity item in A’s further response, that too requiring the environment of A’s initial utterance, even though A’s follow up is in response to what B said. The overall phenomenon is that in language use, participants in a discourse can take arbitrary structures, complete or partial as context, and use these as the point of departure from which to switch into speaking and so overtly contribute to the ongoing utterance. In an exactly converse pattern, a speaker can take where they have got to in producing some utterance as the background context relative to which they can understand what is then said to them. It is not a matter of going into reverse in either switch: it is simply a matter of keeping going from where one has got to.

The seamless fluency with which individuals take on or hand over utterance responsibility raises foundational issues for language modelling. The problem for syntacticians and semanticists is that these split points can separate sub-parts of a structure, apparently splitting dependencies of every sort that linguists have pinpointed, whether syntactic or semantic; yet the espoused sententialist methodology relative to which these dependencies are articulated preclude any straightforward modelling of the data. In some sense, these two parts have to go together to determine a whole syntactic structure; yet the result even so may not constitute a well-formed sentence string (13).

(13) (A emerging from a smoking kitchen)  
A: I’ve burnt the kitchen rather badly.  
B: have you burnt  
A: Myself? No.

In (13), the relevant sub-parts, put back together, yield Have you burnt myself, an ungrammatical string, yet they constitute a perfectly well-formed exchange.

What such data indicate is that these exchanges are not about putting word sequences together to form some single string. They involve speaker A taking over from speaker B to provide some extension of the content of what was initially put into construction by A. In (13), this involves A’s taking up the projection of the second part of the utterance in a way that reflects themselves now as the new speaker. So the projection of structure has not only to represent content, and local dependency on identification of some subject with the use of the reflexive pronoun, but context-dependent content as expressed by that local anaphor. This indicates the extent of the challenge these data pose; the dependencies have to be defined over semantically transparent structure and with whatever imposed context-relativity is determined by the words chosen, echoing the required level of granularity demonstrated by the Greek fragments as in (7).

This phenomenon has been almost completely ignored until very recently as a mystery. With their move to address the challenges posed by modelling dialogue, the new Type Theory with Records framework (TTR: Cooper 2012; Ginzburg 2012) allows a relatively rich concept of abstraction ranging over morphosyntactic as well as semantic content, both at sub-sentence and supra-sentence levels; and TTR has been the first framework to take up
the challenge of modelling dialogue data as systematically within the remit of grammar. Yet there are problems, as the underpinning sententialist methodology remains essentially undisturbed, with fragments taken to constitute syntactically non-standard sentence-types that involve abstractions over some previous sentential context in order to provide the appropriate functor to combine with the presented fragment. This may be a justifiable move where there is a sentence-based paraphrase, but there are many where there is not, as (13) illustrated. There are also problems where the interjected fragment occurs so early in an utterance that there is no appropriate sentential candidate for its interpretation (14).

(14) A: They X-rayed me, and took a urine sample, took a blood sample. Er, the doctor
B: Chorlton?
A: Chorlton, mhm, he examined me, erm, he, he said now they were on about a slight [shadow] on my heart. [BNC: KPY 1005–8]

What is required to characterise the fragment ‘Chorlton?’ in (14) is not any abstraction over the previous context: what is requested is clarification as to the entity described by the word doctor, yielding some token of an individual not sentence type. So it would seem that the attempt to force fragment construal into a propositional frame is imposed solely by dictates of the sententialist methodology in a way that does not correspond to the notion of context itself.

Split-utterance data are problematic also for pragmatists. The assumption shared by pragmatists is that utterance understanding is an act by the hearer of grasping the proposition which the speaker either has or could have in mind (Grice 1975; Sperber and Wilson 1995). This is a consequence of a sentence-based grammar methodology, as the challenge posed by such grammars is that it is in some performance theory, given the sentence outputs of the grammar, that explanation of issues of context-dependence must be articulated. Pragmatists broadly agree that successful communication occurs when the hearer is able to grasp the proposition(s) which the speaker either has or could have intended, given the context. But this necessitates the presentation of some sentence-sized object, relative to which the hearer is able to reconstruct some intended propositional content. In many cases, this assumption is not sustainable, yet the exchange may remain entirely successful. Consider (15) in which the son is resisting all attempts to get him to be helpful. He is certainly not waiting for the third of the commands even if he has bothered to process the second:

(15) (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 pounds?

Even the presumption of any fixed single intended speech act may be in question, since a single fragment is able to serve more than one such function (16).

(16) Lawyer: Will you choose your son as your attorney or
Client: My wife.

Moreover, these data, far from being dysfluencies which the child has to ignore in order to come to achieve competence in their language, emerge in the earliest stages of language learning.
(17) Carer: Old McDonald had a farm… On that farm he had a cow.

And, as with adults, the need to recognize the other party’s intended content seems inappropriate for almost all carer–child exchanges. Indeed, nursery-carers may rely on the interactive ability of the child from an equally young age:

(18) A (teacher to each child in turn in the class): And your name is …
    B (child): Mary.

In (18), the child is merely completing the template set out by the carer with what is answer to a question as well as its completion, which they can do without any need to identify a given thought as held by the questioner. If these data provide evidence that recognition of the content of other people’s intentions is not a necessary condition on successful acts of communication, as we suggest they do, then the foundational assumption of pragmatics needs to be reconsidered (Gregoromichelaki et al. 2011). However, if the grammar formalism is defined to reflect such incremental build-up, these data naturally fall within the remit of grammar, dissolving these syntactic, semantic and pragmatic puzzles.

9.2 Dynamic Syntax

Dynamic Syntax (DS) is a representationalist model of interpretation of which the core structural notion is interpretation growth relative to context. With the dynamics of structural growth built into the grammar formalism, natural-language syntax is a set of principles for inducing growth of such structures. The syntactic mechanisms, being meta to the representations themselves, are procedures (conditional action statements) that define how parts of representations of content can be incrementally introduced and updated, all such growth being relative to context. Moreover, context is as structural and dynamic as the concept of content with which it is twinned: it is a record of the (partial) structures that represent the emergent content plus the actions used in constructing them.

The general process of parsing is taken to involve building as output a tree whose nodes reflect the content of some uttered formula – in the idealised case of a sentence uttered in isolation, a complete propositional formula as diagrammatically displayed in (19). The input to this task 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. This is the minimal left-hand tree in (19), the requirement ?Ty(t) signifying that this is a goal not yet achieved (the ◊ is a pointer indicating the node under development). The output from the parse of the string John upset Mary to the right of the ↦ in (19) constitutes some final end result: a tree in which the propositional formula annotates the top node, its various subterms appear on the dominated nodes in that tree with their specified type (on the tree in (19) these are given as a pair α:x in which α is the formula of type x). The formula values given here are expressed as though in English, but these are concepts as expressed by instructions encoded in words of the language.1 SPAST : es represents the final Davidsonian event/situation argument projected by tense (see Cann 2011):
(19) Parsing *John upset Mary*.

\[ ?Ty(t), \odot \rightarrow Upset'(Mary')(John')(S) : t, \odot \]

\[ S_{PAST} : e_s \]

\[ (Upset'(Mary'))(John') : e_s \rightarrow t \]

\[ John' : e \]

\[ Upset'(Mary') : e \rightarrow (e_s \rightarrow t) \]

\[ Mary' : e \]

\[ Upset' : e \rightarrow (e \rightarrow (e_s \rightarrow t)) \]

These DS trees are invariably binary, reflecting functor–argument structure, and, by convention, the argument is on the left branch, the functor on the right branch; and each node is identifiable by a tree-node identifier \( Tn(i) \) in a tree (not given here). The parsing task is to use both lexical input and information from context to progressively enrich the input tree to yield such a complete output following general tree-growth actions.

### 9.2.1 Formal properties of trees

Though the decoration on trees may be determined idiosyncratically and in part through lexical itemisation, the trees themselves and the mechanisms for their growth are language general, defined by the formal meta-vocabulary for grammar writing. In order to explicitly formulate how such structures grow, trees have to be formally defined, together with a vocabulary for describing actions that induce the requisite tree growth. Following Blackburn and Meyer-Viol (1994), DS adopts a modal logic with two basic modalities. There is \( \langle \downarrow \rangle \): \( \langle \downarrow \rangle \alpha \) holds at a node if \( \alpha \) holds at its daughter (with variants \( \langle \downarrow_0 \rangle \) and \( \langle \downarrow_1 \rangle \) for argument and functor daughter relations respectively). There is the inverse \( \langle \uparrow \rangle \alpha \) which holds at a node if \( \alpha \) holds at its mother, equally with argument and functor variants indicative of the status of the daughter–mother relation so identified. Based on these, there are defined Kleene star operators which yield concepts of *dominate* and *be dominated by*: \( \langle \downarrow \rangle Tn(n) \) holds at a node when a node \( Tn(n) \) is somewhere below it, \( \langle \uparrow \rangle Tn(n) \) holds at a node when a node \( Tn(n) \) is somewhere above it. There are then the analogue \( \langle \downarrow_0^* \rangle \) and \( \langle \uparrow_1^* \rangle \) operators defining a functor spine within a tree, hence concepts of *locally dominate* and *be locally dominated by*. There are also so-called *linked* trees for the modelling of adjuncts (Cann et al. 2005).

The core pair of concepts driving forward the tree growth process is that of underspecification with an attendant *requirement for update*: both are essential for reflecting the time-linearity involved in progressively building up (partial) trees. There are different types of underspecification: of some putatively final tree, of formula content or type, of type of node, and even of relation of that node to others in the tree. All aspects of underspecification are twinned with a concept of *requirement*, \( ?X \), for some annotation \( X \); and these are constraints on how the subsequent parsing steps must progress. Such requirements apply to all types of decoration, including modal requirements expressing future developments, for example \( ?\langle \uparrow \rangle Ty(e_s \rightarrow t) \) capturing nominative case-marking as an output requirement that a node so decorated be immediately dominated by a node of propositional type. Requirements drive subsequent tree-construction processes: unless they are eventually satisfied, the parse will be unsuccessful.

Actions for tree growth determine procedures for building such tree relations, defined as sequences of *make*(*X*), *go*(*X*) and *put*(*Y*) operations, where *X* and *Y* are tree relations and node-decorations respectively. *Computational actions* constitute generally available
strategies for tree-growth, inducing growth of these relations via strategies for unfolding an emergent tree on a top-down basis, and then bottom-up processes compiling decorations for all non-terminal nodes: an overall resulting formula decoration at the top of a tree is thus finally derived, achieving the overall goal of establishing a representation of propositional content.

9.2.2 Lexically induced growth of trees

Lexicon-internal specifications express language particularities in the same formal language as computational actions, differing only in being lexically triggered. The update actions specified are conditional in form: given a certain condition, a macro of tree-growth actions induce some update yielding a distinct partial tree. For example, from a requirement of the form ?Ty(t) as the triggering condition, verbs project a skeletal propositional template projecting a node for a predicate and nodes for its attendant arguments as required by its type. These argument nodes are decorated as part of the action sequence with either the requirement ?Ty(e) or with a typed place-holding decoration exactly in the manner of anaphoric expressions. And this is where there is an array of options available across different languages. Like other Germanic languages but unlike many other languages, the argument nodes projected by English finite verb forms are associated with a requirement of the form ?Ty(e) which ensures that in each case, there has to be some further step of language-input processing in order to satisfy type requirement and project some concept formula of that type. The lexical specification of the verb induces the entire propositional template and such attendant decorations through actions of make(X), go(X) and put(Y):

```
upset

IF ?Ty(t) THEN make(\downarrow_o); go(\downarrow_o);
put(Ty(e), Fo(U_{past}), ?\exists xFo(x); go(\uparrow_o)
make(\downarrow_i); go(\downarrow_i);
put(?Ty(e\rightarrow (e\rightarrow t));
make(\downarrow_1); go(\downarrow_1);
put(Fo(Upset'), Ty(e\rightarrow (e\rightarrow (e\rightarrow t))))
go(\uparrow_i); make(\downarrow_o); go(\downarrow_o);
put(?Ty(e); go(\uparrow_i)); go(\uparrow_i); go(\uparrow_i));
make(\downarrow_0); go(\downarrow_0);
put(?Ty(e)) ELSE Abort
```

In the formulation in (20), first an event term is taken to be projected by the form upset (in fact as an anaphoric device, see §9.2.2.1 and, for detailed specifications of English tense, Cann 2011). Then the predicate internal structure is constructed along with the concept associated with upset, with its two attendant arguments plus that of the event term. As this lexical specification indicates, words project very much more than just a representation of some concept: they project a sequence of actions of arbitrary complexity; and such complex idiosyncratic structural projections may be the emergent result of action sequences that get consolidated over years into composite macro routines.
9.2.2.1 Triggering context-dependent construal

Despite idiosyncrasies that may thus arise, some commonalities run right across languages. Anaphoric expressions, for example, encode an underspecified form of content, being dependent on something external to themselves for interpretation, to be provided either from the general utterance context, or from elsewhere in what is being said. Pronouns provide a familiar case; but anaphora is of several types: nominal, hence 'pronominal', pro-predicate, pro-propositional, pro-temporal, and so on. For example, English auxiliaries can be anaphoric, with indexical uses as seen in (21)–(22).

(21) Mother shouting to her toddler reaching for the saucepan on the stove:
Don’t.

(22) Mother to her teenage son (with surfboard in hand), both looking out to sea at some perilously large waves:
I wouldn’t if I were you.

Indeed the lexical specification of upset reflects this tense anaphoricity as we now see. To model what these triggers for context-based construal encode, we assume that anaphoric, elliptical and tense expressions project a content-place-holding device, together with whatever form-specific constraints there may be on how the relevant interpretation is selected. They are defined as projecting a formula metavariable ($Fo(U)$), with accompanying requirement for formula update, $\exists xFo(x)$. There is undoubtedly cross-linguistic variation in anaphoric and elliptical forms (the English auxiliary system is notoriously idiosyncratic in this respect, with few other languages licensing elliptical use of auxiliaries as in (21)–(22)); but there are general patterns. Context-dependencies are resolvable either from the larger context, as here, or via identification within the construction process. The former case is familiar. With the pointer being at the node at which the expression in question is parsed, such update is normally identified at that relatively early stage in the construction process from what is independently available, possibly from the utterance scenario. However, the parsing process may proceed without any identification of value for the metavariable until a second point in the construction process when its value has to be combined with that of its sister node to yield some value for their shared mother node in order to yield the emergent compositionality of overall content: at this point its formula requirement must be satisfied. So-called expletive pronouns display this property, anticipating and being associated with late identification of the value for the node they decorate, which takes place after their sister predicate node has been constructed as in (23)–(24).

(23) It is unfortunate that Mary smokes.

(24) It worries me that Mary smokes.

Such anticipatory uses are normally treated as wholly distinct phenomena; but, as we shall see, an underspecified value can be resolved either from context or, as with expletives, from within the utterance process itself, once that process has provided the context which makes its update possible.
9.2.2.2 Interplay of lexically and syntax-driven updates

With all lexical and computational actions defined in tree-growth terms, seamlessness between lexicon-internal and lexicon-external processes becomes unproblematic. Concepts of locality, for example, can be identified either word-internally or as a structural constraint. So it is straightforward to identify the requisite concepts of locality constraining antecedenthood for both reflexives and pronouns, enabling a natural feeding relation of such specifications into computational actions of various sorts; and we expect such locality effects to be demonstrable in other areas too. Arguments local to a given predicate can all be defined as meeting the characterisation \( \langle \uparrow_0 \rangle \langle \uparrow_* \rangle Tn(a) \) (‘\( Tn(a) \) holds up one argument-relation plus a possibly empty sequence of function-path relations’). Accordingly, reflexive anaphors can be characterised as projecting the action specified in Figure 9.1; and, conversely, the substitution process of pronominals excludes as antecedent any formula decorating a node standing in such a local relation.

9.2.3 Structural options for tree growth

With the concept of underspecification and update being the core notion on which the grammar specification turns, there is a structure-building analogue of the content underspecification familiar from anaphoric specifications. Amongst the computational actions are processes inducing underspecified structural relations, local and non-local. These weak structural relations are defined using the Kleene star operators: all such weak tree relations have an associated requirement for future provision of a fixed tree-relation, i.e. a fixed tree-node address: \(?xTn(x)\). For example, \( \langle \uparrow \rangle Tn(a) \) is defined as a decoration on a node indicating that there is at least one future development in which the node with address \( a \) bears a sequence of mother relations to the present node. This relatively weak tree-relation is taken to express long-distance dependency effects in terms of structural underspecification and update. For example, when first processing the word \( Mary \) in Figure 9.2, which is initially construed as providing a term whose role is not yet identified, the parse is taken to involve the application of a computational action which introduces from the initial root node decorated with \(?Ty(e)\), a relation to that top node which is underspecified at this juncture, identifiable solely as dominated by the top node (indicated by \( Tn(0) \)), and requiring type \( e \), specified by a \(?Ty(e)\) decoration. This enables the expression \( Mary \) to be taken to decorate this node: this is step (i) of Figure 9.2. The accompanying requirement for a fixed tree-node position \( ?x.Tn(x) \) induces the update to this relatively weak tree-relation which in this derivation takes place after processing the subject plus verb sequence to yield the two-place predicate structure in step (ii) of Figure 9.2. Provision of a formula decoration for the object argument node and update of the unfixed node initially introduced is given by the unification step indicated there, an action which satisfies the update requirements of both depicted nodes.

\[
\begin{align*}
\text{IF} & \quad ?Ty(e), Fo(\alpha) \\
\text{THEN} & \quad \text{IF} \quad \langle \uparrow 0 \rangle \langle \uparrow_* \rangle Tn(a) \\
& \quad \text{THEN} \quad \text{put} \ (Fo(\alpha), Ty(e)). \\
& \quad \text{ELSE} \quad \text{Abort}
\end{align*}
\]

Figure 9.1 Action for reflexive anaphora
Accounting for left-peripheral expressions in these dynamic update terms is not contentious as a parsing strategy: what is innovative is its incorporation within the grammar-mechanism as the basic underpinning to syntactic generalisations.

### 9.2.3.1 Parallelism of content and structural updates

The parallelisms between content and structural modes of underspecification can now be brought out, in two ways and two directions. First, resolution of left-peripheral expressions at some later point parallels that of expletive forms of pronoun construal: both involve some underspecification in their value that is resolved at some later stage in the interpretation process. Second, such left-peripheral expressions can be interpreted as it were indexically from the previous context, from which actions can be retrieved/constructed from context to provide the basis on which the fragment is to be understood to contribute to some larger structure. That is, initial noun-phrase expressions can occur, on their own as it were, at some left periphery with nothing following, hence constituting an incomplete fragment decorating an unfixed node:

(25) Sue, I very much admire. John too.

(26) A: John makes the most amazing cakes.
    B: Mary too.

The resolution of the overall contribution of these fragments is resolved backwards as it were, by reiterating actions taken from the preceding context, with the entire context-provided reconstruction of structure then providing the update to the fragment, essentially like a freer variant of VP ellipsis. This completes the parallelism between structural and content updates, for underspecification of structure in these cases is construed in like manner to indexically construed pronouns, that is from the context provided. This is characteristic of many fragments, particularly those which provide some add-on. And, as we might now expect, there is free interaction between lexical processing of a fragment whose actions can feed into a sequence of actions retrieved from context which in their turn can feed into yet a further step of lexical processing:

(27) A: John’s handing in his assignment.
    B: Bill will be too, tomorrow.
The significance for the modelling of ellipsis in assuming that context keeps a record of actions used is that we now expect that fragments can be construed not merely by retrieving representations of content from context, but also by reiterating actions previously used (Purver et al. 2006). In (27), for example, the actions chosen from context are those employed in the previous processing of handing in his assignment, taken from the preceding utterance and re-applied to the new fragment, prior to the processing of the final tomorrow. To reflect this notion of context formally, what has to be assumed is that, in evolving along with content, the context of utterance keeps track not merely of representations of such content, but also a record of the sequence of actions used. More precisely, context is modelled as the triple of words, sequences of partial emergent structures, and a record of actions used to build them.

Reliance on such contexts as a basis for interpreting elliptical expressions is a mechanism used by very young children, who in the so-called one-word utterance stage rely to a very large degree on the adult with whom they are being able to reconstruct some content from their utterance. One case was observed by Alex, cycling along a canal path with two-year-old son Eliot on a bike pannier behind her, when Eliot waves across the canal to an empty spot on the other side, giving rise to the exchange in (28).

(28) Eliot: Daddy.
    Alex: That’s right, you were here with Daddy yesterday clearing out the boat.

This exchange took place the day after the boat had been emptied of the family possessions and taken back to a central mooring, leaving the empty space pointed to. The child is constructing the predicate to be applied to his fragment from memory, presuming the mother can do so too (though she was not actually there). In such cases, the child is relying on his ability to construct actions necessary to build up content from memory and not from a linguistically presented antecedent utterance – an indexical use of elliptical fragments. Thus, both content-underspecified forms projected by anaphoric expressions, and structural underspecifications in the characterisation of ‘unfixed’ nodes in a structure, may have that underspecification resolved either through the process of content construction, or from the context, that is, from interaction.

There is much more to be said for a full DS account of ellipsis. But the overall picture this has been used to illustrate is that both content dependencies (the provenance of semantics) and structural dependencies (the provenance of syntax) are taken to constitute core syntax, and yet be mechanisms for building interpretation. Remember in this connection from §9.2.1 the problems posed for semantic accounts of ellipsis. In this framework, these are unproblematic, as with nominative case expressed as a tree-update constraint, the model of build-up of interpretation will provide the granularity required to characterise restrictions on construal imposed by case-bearing fragments, as in Greek (§9.1.1). Morphosyntactic particularities, like other more general structural options, are expressed as constraints on growth of representation of content; and these can be recalled exactly as semantic forms of update. Furthermore all updates, whether syntactic, morphosyntactic or lexical, can be manipulated as mechanisms for interaction. Even long-distance dependencies, thought for decades to be the lynchpin of arguments that natural-language syntax is irreducibly encapsulated, can be seen in this light; for these are now seen as inducing an underspecified structural relation and resolving that structural indeterminacy by unification with some independently constructed structure.
Not so fast, you might be pleading: is not this account misplaced as a grammar formalism because it is skewed towards parsing? The answer is indeed that the system is grounded in parsing. But the lack of reference to parsing is little more than a matter of exegesis. Production (generation of pairs of linearised string and tree) is presumed to follow the same process of tree construction with one further assumption: at every step in production, there must be some richer tree, a so-called goal tree, which the tree under construction must subsume in the sense of being able to be developed into that goal tree by rules of the system. This is a direct reflection of the intuition that in production, the speaker must have some richer concept in mind, even if not fully complete. Formally this is the assumption of some goal tree relative to which the emergent tree-growth structure is checked (Purver et al. 2006). So parsers and producers alike use strategies for building up representations of content. In sum, the familiar array of syntactic phenomena – discontinuity in syntactic dependency, expletive pronouns, binding specifications for anaphoric construals, agreement constraints on such construals, and so on – can all be expressed by shifting perspective to grammar formalisms articulating concepts of real-time growth of semantic representation. Syntactic dependencies can be redefined as the effect of manipulating mechanisms licensing various forms of underspecification whose update processes then interact with both general and idiosyncratically imposed constraints to incrementally determine an array of emergent partial structures reflecting content more or less transparently. All that is then needed to express arbitrarily complex structures is the assumption that any one of such locally induced predicate argument structures can be extended by the creation of composite pairs of local structural domains through mechanisms of variable sharing. I will not go into this here, but it forms the basis of accounts of relative clauses, adjuncts, appositions, coordination. Overall, the result is a rich basis from which to set out novel solutions to syntactic puzzles (Cann et al. 2005; Kempson et al. 2011 are representative).

9.2.4 Split utterances: an English case study

An immediate bonus for the DS perspective is that the puzzle of how to characterise the phenomenon of split utterances that this chapter opened with is directly solvable for they follow as an immediate consequence of the general dynamics presumed by the model. Generation, recall, makes use of the same tree growth mechanisms as parsing, with the same commitment to incrementality. This means that, in all exchanges, parties to the exchange will be building up a tree to serve as the basis for both the parsing and production process. From the perspective of modelling the point of switch of roles, two properties of the generation mechanism are pertinent. First, there is nothing to prevent speakers initially having only a partial structure to convey, i.e. the goal tree may be a partial tree. This is unproblematic, as all that is required by the formalism is progressive tree growth in any partial-tree sequencing (=monotonicity of tree growth), and the subsumption check is equally well defined over partial trees. Indeed, the system is set up to project partial trees as context, hence as input to some next point of departure. Second, the goal tree may change during generation of an utterance, as long as this change involves monotonic extension: because speaker and hearer essentially follow the same sets of actions, each incrementally updating their semantic representations, the hearer can effectively continue to mirror the same series of partial trees as the producer after the switch, albeit no longer knowing in advance the content of the unspecified nodes, no matter what extension the producer may provide. For example, for dialogues such as (29), Mary as the speaker reaches a partial tree of what she has uttered through successive updates, while Bob initially as hearer, follows
the same updates to reach the same representation of what he has heard: they both apply the
same tree-construction mechanism which is none other than their effectively shared
grammar.2 This provides Bob with the ability at any stage to become the speaker, interrupting
to take over Mary’s utterance, repair, ask for clarification, reformulate, or provide a
correction, as and when necessary. What, then looks like distributed construction of some
structure across distinct participants is in fact no more than switch within the tree-
development process from parsing without any goal tree for subsumption purposes to the
very same tree construction process except that with the switch into production there will be
a goal tree providing the basis for the further subsumption check. Yet the dependencies
thereby apparently so distributed remain constructed and resolved within a single locally
definable tree domain for an individual agent. So, according to DS assumptions, repeating
or extending a constituent of someone else’s utterance is licensed only if the hearer now
newly turned speaker, entertains a message to be conveyed (a new goal tree) that matches
or extends in a monotonic fashion the parse tree they themselves have constructed from
what they have just heard. Let us take this up step by step on (29), using a simplified variant
of (13).

(29) Mary: Did you burn

Despite the fact that the string *Did you burn myself? is unacceptable, under DS assumptions,
with representations only of informational content, not of structure over strings of words,
the switch of person is straightforward and leads to a well-formed result. Figure 9.3 displays
the partial tree induced by parsing Mary’s utterance Did you burn which involves a
substitution of the metavariable projected by you with the name of the interlocutor/parser
(setting aside how to capture questionhood). At this point, Bob can complete the utterance
with the reflexive as this expression, by definition, copies a formula from a local co-argument
node onto the current node just in case that formula satis-

This illustration is only of the simplest type of split-utterance ellipsis, but the point is
general. These seamlessly achieved split utterances can apparently separate off any
expression from the syntactic environment it needs for its wellformedness because both
speaker and hearer make use of the same mechanisms. So the same individual whether as
speaker or as hearer will invariably have a partial structure on which to rely at the point of
participant switch: the use of structure retrieved from context is just one type of information
stored in context. Context has the same granularity of information as the lexical speci-
cifications given by the words. What is recorded is: (1) the words so far expressed; (2) the partial
trees so far built, up to and including the one containing the node indicated by the pointer; (3) the
actions used to build up that emergent partial structure. And since it is actions making
up the process of interpretation which constitute the basis of syntax, it follows that we shall have all the granularity needed to characterise the full range of ellipsis effects. Of course the justification for this claim will turn on the detailed specifications of individual cases, but, in principle, the DS framework articulates the concept of context that is needed to provide an integrated account of ellipsis.

Stepping back from the details, note that the account of ellipsis in dialogue rests on the jettisoning of a conventional ‘syntactic’ level of representation in favour of syntax as a set of mechanisms for incrementally building semantic representations, with these mechanisms shared by both speaker and hearer. The incrementality of parsing and production is the key to the simplicity of the story: it is this which allows the direct successful integration of interpretation provided by the spliced-utterance fragments. Both speaker and hearer are continuing from where they have got to in the utterance processing task, intuitively and formally using their own context (the Alex–Eliot exchange is notable in this connection). In contrast to the immediacy of DS applicability, the phenomenon of speaker switch in the middle of a dependency under construction is a major challenge for sentence-based grammar frameworks. These fail to provide the appropriate units: all dependencies are articulated grammar-internally as defined exclusively over hierarchically defined sentential structures which in principle do not reflect time-linear emergent growth. From that sentence-based perspective, successful communication is then seen as establishing some propositional content from sentences that the speaker could have intended to express, with essential mind-reading. On the DS account, no such recognition of intended content is required as a definitional property of communicative understanding, for dependencies are defined over some partial structure as yet still under construction, and specific to the individual. The data which are so recalcitrant for these frameworks are thus predicted as a direct consequence of the way DS mechanisms are defined.

To substantiate the claim that DS assumptions can express the whole range of syntactic phenomena of course needs very much more than one limited case study (Cann et al. 2009). But because ellipsis requires syntactic and semantic reconstruction, the point is general. Because syntax involves the emergent projection of representation of content in combination
with lexical tree growth actions, we can express what is general to structure and interpretation in language, yet the variation also. A considerable amount of DS-related work has been done on cross-language, cross-dialectal and historical syntax substantiating these claims. There is work on tense and aspect articulating event-term specifications (Cann 2011). There is work on left–right asymmetries (Cann et al. 2005 among others), and on verb-final languages, which receive natural syntactic analyses only when syntax is defined as incremental building of semantic structure (Kempson and Kiaer 2010). There is work on pronominal clitics, whose obscure syntax–semantics mappings have been successfully modelled in terms of concepts of routinisation of action update sequences, a move which provides an evolving basis for both diachronic change and synchronic variation (Chatzikyriakidis and Kempson 2011; Kempson et al. 2011). There has also been exploration of issues of intentionality (Gregoromichelaki et al. 2011): and there is a host of further issues, both linguistic and philosophical, waiting in the wings.

Acknowledgement

This chapter is grounded in work on The Dynamics of Conversational Dialogue (ESRC-RES-062-23-0962). I am immensely grateful to my co-workers on this project and too many others to thank who have contributed to the fine landscape which this work has opened up. Thanks in particular to Ronnie Cann for comments on an earlier draft of this chapter.

Notes

1 Possible complexities of lexical decomposition are ignored here. Quantification is taken to be name-like, hence of low type: see Kempson et al. 2001; Cann et al. 2005.
2 A completely identical grammar is, of course, an idealisation but one that is harmless for current purposes.

References