Grammar as procedures

Citation for published version:
Kempson, R & Cann, R 2018, Grammar as procedures: Language, interaction, and the predictive turn. in K Turner & L Horn (eds), *Pragmatics, Truth and Underspecification: Towards An Atlas Of Meaning.* vol. 34, Current Research in the Semantics / Pragmatics Interface, Brill, pp. 261-299. https://doi.org/20.500.11820/453b645e-4433-4505-b73b-6921ed60f1c5

Digital Object Identifier (DOI):
20.500.11820/453b645e-4433-4505-b73b-6921ed60f1c5

Link:
Link to publication record in Edinburgh Research Explorer

Document Version:
Peer reviewed version

Published In:
Pragmatics, Truth and Underspecification

Grammar as Procedures: Language, Interaction, and the Predictive Turn

Ruth Kempson and Ronnie Cann

May 28, 2017

1 Grammar, processing and interaction

Building on the trend initiated in part by the insight of Atlas 1977 that linguistic content had to be seen in terms more abstract than is expressible in denotational terms, at least for negation, this chapter argues for a new perspective on language in which language is characterised in procedural "knowing how " terms and emerges as a natural subpart of an overall cognitive system, bringing to it the added dimension of cross-party interaction. This is achieved by bringing together two independent strands of research. On the one hand, we sketch the Dynamic Syntax view of language (Dynamic Syntax: DS Cann et al. 2005), in which the concept of procedure is central: languages are defined to be sets of conditional actions incrementally inducing the building up of representations of content relative to emergent and shifting contexts to yield effects of linearisation (production) and interpretation (parsing). Furthermore, because these procedures constitute a grammar, producer and hearer are expected to use the same predictive procedures for string-content processing, both incrementally building up structured representations relative to context and some anticipated goal; so interaction between conversational agents is directly anticipated.

In then turning to a broader cognitive science background, we find confirming evidence of this perception-as-action perspective in the way this dynamic is shared by the brain itself. The brain is claimed to be a predictive engine using immediate and encyclopaedic context constructively at every step to predict the structure/shape of the incoming sensory array (The Predictive Processing (PP) model of cognition: A. Clark 2013, 2016). Within this model, as in the DS account of language, all activity is seen as process-based and embodied: action and perception alike involve anticipatory projections as to what the input stimulus might be providing, with that stimulus acting as a filter, confirming or correcting those anticipations. Given the consonance of these two models, and adding to the PP framework the dimension of interaction which the combined assumptions of DS and PP yield, language emerges as one of a number of cognitive processes which lead to manifest joint interaction through the embodiment of procedures for interacting with the external world (others being music making and dance: Cross 2012, Pezzulo 2011).

With this move, the competence-performance divide is left behind in favour of a model in which the very "knowledge of language" itself is defined procedurally as involving mechanisms whose effects are directly witnessed in actual processing behaviours (parsing and production). And from this interdisciplinary marriage, we have the promise of a wholly new perspective on language evolution following the pattern of first-language acquisition as grounded in the effectiveness of language as an evolved set of mechanisms for human interaction.

In order to be able to appreciate the significance of this interdisciplinary union, there are two important sub-tasks to address: (i) why the DS model of language is a grammar formalism in its own right and not just a model of some aspect of information growth, parasitic on some static grammar formalism; (ii) what evidence there is for claiming that language is grounded in actions yielding interaction between participants through such growth.

1.1 Dialogue Interactivity

The starting point of our argument is the widespread recognition of endemic context-relativity of both natural language interpretation and language production, so in modelling language processing, both perception and production must be seen as involving not merely the concept of evolving content but also its twin concept of perpetually evolving context (Pickering and Garrod 2004, 2013, and for a useful survey see Clifton et al 2012).

Some of the most striking evidence that modelling language processing involves the articulation of evolving representations of content comes from the data of conversational dialogue. In informal conversational exchanges, notably the only data to which a small child is exposed, the emergent conversations are littered with so-called fragments where speakers and hearers simply assume that in some sense to be made explicit, the context provides all that is needed to see their apparently fragmentary contribution as providing a coherent and consistent add-on to what has just preceded. We can join one another's conversations, becoming a new speaker while the previous speaker shifts into being a hearer, a shifting between speaker/hearer roles taking place repeatedly during the course of the exchange. Indeed, this shift of roles is the sine qua non of informal conversations, monologues being the exception rather than the norm:

(1) A: I need a a
B: mattock. For breaking up clods of earth [BNC]

(2) Jack: I just returned
Kathy: from
Jack: Finland [Lerner 2004]

The take over may take the form of an overt attempt to be helpful to the other party, but this is by no means their only function. One may be simply seeking to extend what the other has said:

(3) A: We're going to
B: Marlborough
C: Marlborough?
B: to see Granny
C: With the dogs?
A: if you can keep them under control
C: in the garden?
B: unless it rains
C: which it always does

Conversations such as these, where there may be covert competition as to who can have the last word, are the bedrock of informal conversation, by no means constrained by any over-arching intentionally held aim of achieving coordination. And, of course, extensions such as these may be uttered by a single interlocutor refining and elaborating an initial sentence, as in (4) where we have a first extension providing the time of the event given in the initial utterance, then a source of the event, then a modifier of the subject:

(4) Mary's back. Late last night. From the US. Tired and frustrated at the delays.

There is an open-ended flexibility in how such interactive construction of dialogues can be achieved. The adding on of a fragment in context may be used for novel effect at the level of the speech act, when apparently dual speech acts can be achieved within a single emergent utterance:

(5) A: Are you left or
B: Right-handed

Such interaction is effective even when the parties may be in sharp dispute:

(6) A: It's obvious from what he says (that)
B: (that) you are wrong

Partial fragments may also be being used to direct one's interlocutor to construct something novel, an ad-hoc context-specific concept that depends on knowledge of the individuals in question:

(7) Bird-expert: I watched birds yesterday all afternoon, in the sunshine
Friend: with your grand-son?
Bird-expert: It was great fun. He did really well . . . for 20 minutes.

Metaphorical utterances can be taken up and extended:

(8) A: How's work?
B: Oh, I shuffled along
A: without getting much done, I imagine.

And even very young language-acquiring children can join in on such interactions, either following up on some proffered frame as in (9)-(11) or in initiating a process of frame construction as in one-word utterances, with no preceding linguistic context, as in (12):

(9) A: Old MacDonald had a farm. A-I-E-I-O. And on that farm he had a
B: cow.
A: and the cow goes
B: Moo

(10) A (carer to each child in turn in the nursery-group): And your name is . . .
B (child) : Mary.

(11) Experimenter: This is (NOT) a . . .
Child: comb! [reconstructed from De Villiers and Tager Flusberg, 1975]

(12) Eliot (2 year old on mum's bike waving at empty mooring on the other side of the canal): Daddy.
Mother: That's right dear, you were here yesterday with Daddy clearing out the boat. [direct observation]

The general dynamic is that each such contribution adds unproblematically to whatever string of words has been set out so far, irrespective of whether or not what precedes it is a full sentence, and whether or not the add-on contributes to what the other party is trying to get across. The effect is one of a rich potential for interactivity between dialogue participants, available to all speakers of a language from very early stages of language development.

1.2 Implications for grammar

The existence of interactive constructing of utterances, and the fact that utterances by a single individual may be only a partial contribution to some gradual build up of information, pose significant problems for all conventional grammatical frameworks, in every sub-area of the grammar. From a syntactic perspective, we need to ask: how are split utterances possible; why are they so common in conversational dialogue; and how is it that the shift from one interlocutor to another can occur at any point in a sentence-construction process? Questions such as these might be, and often are, set aside as performance considerations only. But there is one consideration, overwhelmingly, which resists any attempt to dismiss the data as peripheral – the seamless fluency with which individuals take on or hand over utterance responsibility, apparently distributing syntactic and semantic dependencies across more than one participant. The problem for conventional grammars is the commitment to model an individual speaker's linguistic competence as involving knowledge of a set of rules (however defined) that licenses all and only the strings of the language which that speaker judges to be well-formed; and, moreover, only those strings that are defined as sentences with representations rooted in the largest constituent recognised in the theory: S, CP, VP[+SBJ], type t, or whatever.

Even if the goal of a grammar were weakened to characterise only a subset of the total phenomena displayed by a language, the fact that utterances can apparently split apart every single dependency within a language leads to the result that *every* generalisation articulated within a standard grammar will be incomplete. According to even the most modest evaluation metric, the grammar will fail. This is because, in all cases, there will be split utterance data which, despite being wholly acceptable to individuals in some speech situation, will not be given any characterisation by that grammar in virtue of their apparently fragmentary nature, not only non-sentential in nature but possibly also not even constituting any phrasal type, as in the first utterance in (1), (3), or the second in (2) and (11). In order, then, to retain the grammar as an explanatory tool, such expressions uttered by the individual speaker, independent of any interaction with other speakers, will not be characterised as falling within the remit of the grammar. They will accordingly have to be deemed to be irrelevant, characterisable by some other set of principles. But any such dismissal would solely be on grounds that the distribution of such fragments is inexpressible within the grammar methodology adopted, and so not deemed part of any sort of competence on the part of a speaker (Chomsky 1965). But such fragments are not, by any stretch of the imagination, dysfluent uses of language: to the contrary, they are achieved with total fluency, demonstrating a fine level of sensitivity to the evolving development of the discourse.

The ever-extendability of a structure across more than one speaker may at first sight seem to be a mere idiosyncracy of informal conversational exchange, where people may allow a single composite sentence to unfold apparently indefinitely as in (3), with each providing their own add-on, one after the other. There are two issues here. First, there is the issue of how fully determined the concept of a sentence boundary is. It is commonplace to observe of languages with which we are not familiar that the sentences "seem to go on for ever" and it is hard to find out where one begins and the other ends. But what data such as (3) illustrate is that this is an attribute of spoken conversation. The assumption that this doesn't occur in well-documented languages is no more than a parochial folk-prescriptive attitude to "correct grammar" as essentially that of written language (Miller and Weinert 1998). It might then seem to be but a short step to see that in conversation, one contributor to the conversation is merely taking over responsibility for some emergent sentential sequence from someone else, with the the activity being in some sense shared, an explicitly joint activity.

So what seems to require building up is not the construction of sentences at all, but rather some structured representation of content. In some cases, indeed, the subparts of the sentence-string which appear to be omitted would seem to be either internally inconsistent or even not well-formed, considered as a single sequence of words:

(13) Eleni: Is this yours or
Yo: yours. [natural data]

(14) {A emerging from a smoking kitchen} A: I've burnt the kitchen rather badly.
B: Have you burnt
A: Myself? No.

In (14), the relevant sub-parts, put back together, yield an ungrammatical string due to the conflict in person between the binder and the reflexive pronoun: **Have you burnt myself.* Yet the exchange is perfectly well-formed, the reflexive pronoun being interpreted locally as the subject of the predicate just constructed from B's utterance, but with the switch of speaker having to be commensurate with attributes of the new speaker, now A. Intuitively, in some sense, these two parts have to go together to determine a whole syntactic structure, even though the result string-wise is not grammatical.

What such data indicate is that these exchanges are not about putting word sequences together to form some string of words constituting a sentence or even a string of sentences in a discourse (see also (3) and (13) above which, when put together into a single discourse are incoherent, despite the interaction being perfectly acceptable). Instead, what we appear to need to characterise such dialogue is a grammar that licenses, not sentences as traditionally construed, but sub-sentential units that put together may, or may not, yield a well-formed sentence as defined within a language's grammatical tradition depending on speech act and other contextual factors. By implication, this yields a view of grammar as providing mechanisms for successful linguistic interaction, rather than a set of rules that define the set of all and only the grammatical sentences taken by speaker-hearers to constitute a language.However, even though it may be that representations of meanings are what is constructed rather than representations of putative structure defined over strings of words, such structures are nevertheless not wholly insensitive to the morpho-syntactic constraints imposed by the words of the language, as the required reflexive form in (14) indicates. This problem cannot be dismissed as a language-particular idiosyncracy. In all case-rich languages, fragments have to adopt a morphological form appropriate for what would be expected in an overt clausal sequence, commonly replicating the pattern provided by the immediately antecedent context. Thus in (15), the form of the fragment pronoun has to be the nominative form *ego*, not the accusative *emena*:

(15)

A:	*I*	*Maria*	*to*	*egrapse*	*to*	*grama?*	*B:*	*Oxi,*	*ego*	*(*emena)*
	the$_{Nom}$	Maria	it$_{Acc}$	write$_{Past,3sg}$	the	letter		No	I$_{Nom}$	(*me$_{Acc}$)

A: 'Did Maria write the letter? B: No, I did (*Me did)'

Furthermore such fragments are required to take a definitive morphological form whether or not there is some overt antecedent in a previous clause to provide the appropriate pattern, again as in Modern Greek (Gregoromichelaki 2012):

(16) [Context: A is contemplating the space under the mirror while re-arranging the furniture and B brings her a chair]

A to B:	*tin*	*(*i)*	*karekla*	*tis*	*mamas?*	*Ise*	*treli?*
	the.ACC	(*the.NOM)	chair	the.GEN	mum.GEN	be.2.SG.PRES	crazy

'Mum's chair? Are you crazy?

Thus, first and foremost, relative to any sentence-based grammar, the split utterance data seem to fall irretrievably outside its remit. But even if the remit of grammar is somehow extendable to include these data, the structured representations which have to be induced appear to have both to be representations of content and yet also sensitive to what morphological constraints the particular grammar may impose on what representations can legitimately be constructed. Moreover such constraints must be expressible as lexical properties of the given words, despite possibly lacking the environment these need for successful completion. There are thus major hurdles for any conventional syntactic framework in addressing these data, if the framework is to have the expressivity needed to incorporate them within its remit.

The problems posed by naturalistic data are barely less for semantics. First is the issue of how it is that almost every word can have its interpretation modified in context with compositionality to be defined over such modifications. The meaning attributable to our words reflects the context to which it is intended to contribute, shifting so seamlessly from one understanding to another that it is often hard to notice even that there has been such a shift – in (17) three different concepts of 'burn' emerge within a single exchange.

(17) A: I'm afraid I've burned the kitchen ceiling. The paper's blackened and part of it has come away.
B: Did you burn
A: myself? No fortunately not. Well, only my hair.

And in (18), an expression is placed in the conversational arena, so to speak, and immediately modified, or at least justified by one particular attribute of stage-managing:

(18) M: He's stage manager [pause]
He's actually first assistant but- he's calling the show. [Hough 2015]

There is then the question of hand-on from one speaker to another. If meanings evolve over the course of an exchange, with shifts which we are barely aware of, how is it that propositional content can nevertheless be seen to emerge incrementally across a group of speakers?

Things are no better for pragmatics. Probably the most basic questions needing an answer are: how can speaking and hearing be so seamlessly interwoven; and why is there such systematic cross-speaker use of non-sentential utterances in dialogue? The reason these questions are fundamental is that the answers to them impact on two of the most common assumptions made in pragmatic theory: that pragmatic inference operates over propositions as expressed by sentences; and that successful communication is only achieved if a hearer can identify the proposition or thought that a speaker intends (or could have intended) to convey (Grice 1975, Sperber and Wilson 1995, Bach 1994). The former assumption is seriously called into question by examples like (3). Here, the conditional *if you can keep them under control* scopes, not over the proposition purportedly expressed by the preceding utterances, *We're going to Marlborough to see Granny with the dogs* but only over the non-sentential prepositional phrase *with the dogs*: going to Marlborough to see Granny is not contingent on C's being able to control the dogs. If inference is defined solely over complete propositions, then any such theory will have to be powerful enough to reconstruct some proposition from the prepositional phrase and further to explain why the scope of the conditional is not over the proposition expressed by the previous utterances taken together.

Additionally, the construction of a speech act, usually taken to be something that only sentences are the vehicle for conveying (Austin 1962, although see Stainton 2006 for a discussion of speech-acts conveyed by sub-sentential utterances), is nevertheless able to be constructed as in (13) above and (19) below in which the first utterances are incomplete interrogatives implying a question, but whose interactive completion yields a statement:

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

With respect to the question of intention recognition, there is plenty of evidence in dialogue that, while building on some estimation of the other person's mind-set is commonly made use of both in parsing and in production, it is not a prerequisite for successful communication (Gregoromichelaki et al. 2011). For example, the clarification request in (20) is made before B could possibly have recognised any intended meaning by A:

(20) 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 1005-1008]

In non-co-operative examples like that in (6) above and (21) below the second speaker completes the previous utterance without any necessary consideration of the intentions of the first speaker. In (21), the son is certainly not waiting for the third of the commands even if he has bothered to process the second, but nevertheless he has responded wholly appropriately

(21) Mother: This afternoon first you'll do your homework, then wash the dishes and then
Son: you'll give me 10 pounds?

And, most strikingly of all, such split-utterances are employed in interactive exchanges with very young children, even when a compound speech-act effect is achieved as in (10) parallel to (19). Indeed they are one means for detecting complex conceptual abilities like negation as in (11). This strongly suggests that recognition of the content of other people's intentions is not a necessary condition for acts of communication to be successful and, more generally that acquisition need not make reference to higher order reasoning (Gregoromichelaki et al 2011).

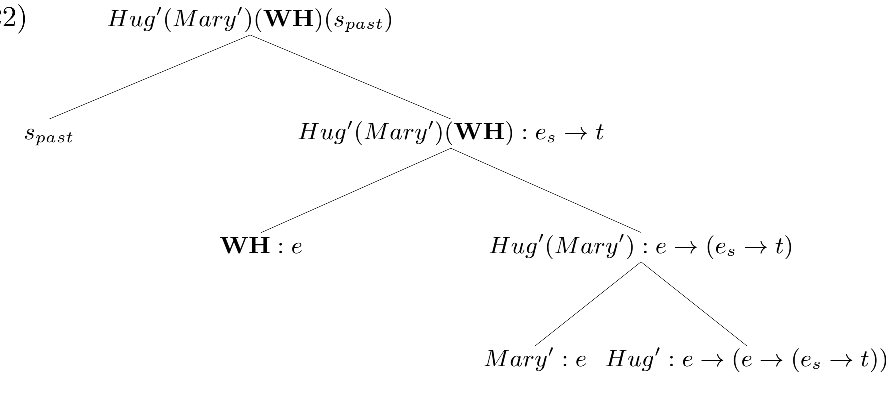
Taking a step back, what appears to be needed to model these data is a framework in which strings, content, context and even speech-act determination can all be seen as evolving during the course of a conversational interaction. This involves a radical shift of assumptions. Instead of presuming on a methodology which isolates the capacity for language from all attributes of its use, the concept of the process of building up structured representations needs to be taken to be central. Instead of such a system being assumed to be articulated in a vocabulary wholly encapsulated from general cognitive considerations, the emergent representations will need to be conceptual representations, so in principle allowing free interface between such grammar-internal specifications and general context considerations. Any such level of representation will nonetheless have to be sensitive to whatever morpho-syntactic constraints the forms of the language may impose, so the articulation of morpho-syntactic constraints will have to be in terms of constraints on the mappings from the morphological forms onto such conceptual representations. In short, we are turning to the exploration of a view of language which breaks out from the stranglehold of the competence-performance division in two ways, putting in its place a methodology for defining language as a set of actions for coordination between interlocutors. The concept of *sentence* will no longer be central to determining the remit of grammar; and the structural properties of any natural language will be defined in terms of conditional actions mapping words and word sequences onto representations of content. So the system will no longer be an encapsulated set of properties defined over the strings of words and lacking any interface with general cognitive restrictions (as in conventional frameworks), but rather will be defined in a domain-general vocabulary and able to function as a proper subpart of the general cognitive architecture. Principles underpinning structured strings are thus transformed into procedures for incrementally building structural representations of content, a stance which enables language acquisition to be seen as a continuous developmental process rather than a sequence of different grammar formalisms with the child having to hypothesise the chain of grammar improvement. With this transformation, the concept of mapping expressions onto emergent contents and contexts as these become available lies at the centre of what has to be expressed, so the concept of procedure becomes central, together with the twinned concepts of initial underspecification and subsequent update.

2 Modelling incrementality and context-dependent interpretations

The framework we adopt to model this phenomenon, is Dynamic Syntax (DS), which defines grammar as mechanisms for incremental growth of interpretation: the concept of underspecification plus update is central to the grammar (Kempson et al 2001, Cann et al 2005, Kempson et al 2015, 2016). DS is a representationalist model of interpretation of which the core structural notion is interpretation *growth* relative to context, so a process-based model. What has gone in this approach to language modelling is any concept of static, purely syntactic structure, inhabited by strings so a fixed pairing. What is retained is a DRT-like concept of building representations of content relative to context (Kamp and Reyle 1993). The formal devices are defined to model how to pair emergent interpretations with words being uttered in sequence. The system defines actions that give rise to expectations of further actions, all involving incremental updates towards some overall goal. On this view, the setting out of such dynamics constitute the grammar: syntax just IS a set of principles for inducing growth of such structures. The syntactic mechanisms are procedures that define HOW parts of representations of content can be introduced and updated, all such growth being relative to context. Moreover, context is as dynamic and structural as the concept of content with which it is twinned, a record of the emergent structures that represent the unfolding of interpretational content plus the actions used to develop this incremental process (Purver et al. 2006). So the overall picture is a shifting one: shifting goals, shifting contents, and shifting contexts, as each established content becomes part of the context for the next step in the process.

These goal-driven sequences of actions are taken to involve building partial trees, with, as output, a tree whose nodes reflect the content of some utterance, as in (22) which represents the content of an utterance of *Who hugged Mary?* in binary tree form.

(22)



The display in (22) indicates only the output representation of the utterance - a static structure. What is not shown is the way the process giving rise to this output is initiated and developed. It is thus the process of tree growth and not the structure of the output trees (mere reflections of functor-argument structure) that provides the explanatory core of the framework.

The point of departure for the analysis of *Who hugged Mary?* is some partial tree whose only decoration gives a skeletal indication of the goal to be achieved, a propositional formula, otherwise unspecified, i.e. a single node with the single decoration a goal requirement to derive a formula of the specified type, $?Ty(t)$:[1]

(23) $Tn(0), ?Ty(t), \diamondsuit$

What has to be built to meet this goal requirement is a formula matching this type, a result which is achieved on the basis of what the context already provides, what general strategies for progressive growth of structure are available, and information provided by words used. In any complete tree, each node has a content formula, shown to the left of the colon ":" and a type specification to its right (e for entity, e_s for event, $e \rightarrow (e_s \rightarrow t)$ for predicates, etc.). The trees are binary, with, by convention, the functor node on the right, and the argument on the left. Individual terms are invariably of type e, presuming on an account of quantification, which we will sketch later. There is a particularised type e_s for event terms, only schematically represented here (Gregoromichelaki 2006, Cann 2011). The dynamics of growth is partly top down – as general strategies open up possibilities – and partly bottom up – as actions induced by the words provide terminal decorations which provide the input to a bottom-up process yielding the effect of compositionality in the resultant tree.

To flesh out this concept of growth, we need first a vocabulary for defining trees, and then a vocabulary for defining how such trees can be seen to grow. The modal logic of trees of (Blackburn and Meyer-Viol 1994) provides the language of tree description, with two basic modalities. $\langle\downarrow\rangle$: '$\langle\downarrow\rangle\alpha$ holds at a node if α holds at its daughter (one node down)' for which there are variants $\langle\downarrow_0\rangle$ and $\langle\downarrow_1\rangle$ for argument and functor daughter relations respectively. And the inverse $\langle\uparrow\rangle\alpha$: '$\langle\uparrow\rangle\alpha$ holds at a mother node, one node up, if α holds at its mother', equally with argument and functor variants. There are also Kleene star operators which yield concepts of underspecified *dominate* and *be dominated by*: $\langle\downarrow_*\rangle Tn(n)$ holds at a node when a node $Tn(n)$ is somewhere below it (along an arbitrary sequence of daughter relations), $\langle\uparrow_*\rangle Tn(n)$ holds at a node when a node $Tn(n)$ is somewhere above it.[2]

There are three principal kinds of underspecification that require updates. First, there are partially specified trees as some nodes or formulae remain yet to be developed. Secondly, there may be underspecified content formulae which are assigned a type but as yet with only a place-holding metavariable as formula, itself needing update to some specified formula value. Of these, anaphoric devices are the familiar cases of content underspecification, encoded as projecting such a metavariable, possibly with constraints on substitutions for them: for example, $\mathbf{U}_{Female'} : e$ for *she/her* or $\mathbf{U}_{3Plural'} : e$ for *they/them*. Thirdly, the concept of underspecification is extended also to structural underspecification, in which it is a tree-relation that may be incompletely specified, so that a node introduced into a tree may lack a fixed site, hence referred to as an "unfixed node", defined by the Kleene* operator $\langle\uparrow_*\rangle$ and its variants. This is the mechanism which underpins long-distance dependencies and local scrambling.

Underspecification of any parameter (type, formula or position within a tree) is invariably twinned with an associated *requirement*, symbolised as $?X$, for some annotation X. So we have type requirements $?Ty(\alpha)$ to find a formula of some type α; $?\exists\mathbf{x}.Fo(\mathbf{x})$ for a requirement to find a value for a metavariable; and $?\exists\mathbf{x}.Tn(\mathbf{x})$ to find a fixed treenode address for an unfixed node. Requirements may be modal as well as non-modal, with modal requirements imposing constraints on how a structure may be developed. For example $?\langle\uparrow_0\rangle Ty(e_s \rightarrow t)$ is a requirement expressing an output constraint imposed by nominative case-marking that a node so decorated be the highest

[1]The label $Tn(0)$ gives the address of the node in the unfolding tree, here address 0, the rootnode. $\diamondsuit$ is the pointer indicating the node currently under development.

[2]Various types of locality can be defined over dominance relations, inducing different domains in which unfixed nodes can be fixed. See Cann et al (2005) for details.

individual argument of a predicate, still needing an event term incorporating temporal specification with which to combine. Requirements are essential for reflecting the incrementality involved in progressively building of trees, providing goals to be achieved that may be satisfied immediately or not until substantially later in a derivation.

2.1 Actions for Tree Growth

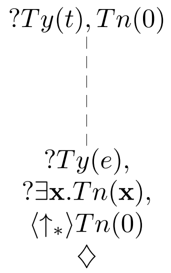
The concept of underspecification is itself a static notion, a property of a given tree configuration. To define the *process* of tree-growth, we adopt a language for formulating actions that determine the (monotonic) transitions between partial trees. There are four primitive actions: `make`(X) for constructing a new node; `go`(X) that moves the pointer to a node; and `put`(Y) that annotates a node with some information. There is also an `Abort` action, which terminates an action sequence. These are used to define a range of action sequences.

Computational Actions constitute generally available strategies for tree-growth, either inducing the unfolding of an emergent tree on a top-down basis, or inducing bottom-up processes which, once appropriate terminal nodes are decorated, lead to the decorations for all non-terminal nodes. For example, as one of a very restricted number of initial steps in a derivation starting from a propositional requirement, $?Ty(t)$ as in (22), a computational rule licenses the construction of a node characterised solely as being unfixed within a given tree domain, as expressed by decoration $\langle\uparrow_*\rangle Tn(a)$ which indicates that "somewhere above me is the Treenode a'.[3] Rules are given in standard $\langle IF\ldots, THEN\ldots, ELSE\ldots\rangle$ statements. The example in (24) shows the actions and output tree that define the construction of such an unfixed node.[4]

(24)

Actions:		
IF	$?Ty(t), Tn(a)$	
THEN	IF	$\langle\downarrow\rangle\langle\downarrow_*\rangle\top$
	THEN	Abort
	ELSE	$\texttt{make}(\langle\downarrow_*\rangle); \texttt{go}(\langle\downarrow_*\rangle);$
		$\texttt{put}(\langle\uparrow_*\rangle Tn(a), ?Ty(e),)$
		$\texttt{put}(?\exists x Tn(x))$
ELSE	Abort	

Output tree:



Words also invoke actions of the same sort: making and annotating nodes and moving the pointer around a tree. The simplest of such macros of actions is when nodes are merely annotated with formulae as in a parse of the name *Mary* as in the entry in (25):[5]

(25) *Mary*

IF	$?Ty(e)$
THEN	$\texttt{put}(Ty(e), Fo(Mary'))$
ELSE	`Abort`

Verbs induce more structure than names, triggering conditional actions that induce all the nodes that define their predicate-argument structure, including an event argument where lexical

[3]The tree displays diagrammatically the unfixed node, using a dashed line. Fixed tree relations are shown with a solid line. Note that such tree diagrams record only the effect of the rule, and not conditions on its application.

[4]There is a condition on this action that it can only take place if no other node from this point of departure already exists within this newly emergent tree. This is the second condition preceding the sequence of actions, aborting the action if there are already constructed nodes within this emergent tree. There is a distinct local variant of this action inducing underspecified substructure within individual proposition-templates, but for simplicity we omit this here.

[5]This is an oversimplified characterisation omitting all discussion of how or whether names are associated with a uniqueness constraint or what form that constraint should take.

aspect information may be added and tense is specified by finite verb forms. Languages vary as to what options for the argument node decorations they license. In English, being not a "pro-drop" language, verbs specify argument nodes as open type requirements of the appropriate sort ($?Ty(e)$ for individual term arguments, $?Ty(t)$ for propositional complements, etc.) which ensures that there has to be some further step of linguistic-input to satisfy the type requirement. Languages also vary in word order, with English finite main verbs unable to appear sentence initially and necessarily appearing after the subject and before a direct object (modulo discontinuous dependencies). These constraints are guaranteed by the set of actions in (26) where the second condition requires some structure to have already been built, while the first two actions ensure that this is identified as the logical subject (for non-passive forms). The remaining actions build an event argument which is annotated with lexical aspectual information and tense (for finite forms) and then construct the remaining predicate-argument structure, leaving the pointer on the open logical object node, thus ensuring that some object term be identified next. This complex of conditions and update actions are illustrated in the lexical specification of the verb *hugged* in (26):[6]

(26) *hugged*

$$
\left|\begin{array}{lll}
\text{IF} & Tn(a), ?Ty(t) & \\
\text{THEN} & \text{IF} & \langle\downarrow_*\rangle\top \\
 & \text{THEN} & \texttt{go}(\langle\downarrow_*\rangle); \texttt{put}(?\langle\uparrow_0\rangle\langle\uparrow_1\rangle Tn(0)) \\
 & & \texttt{go}(\langle\uparrow^*\rangle Tn(a)) \\
 & & \texttt{make}(\langle\downarrow_0\rangle) : \texttt{go}(\langle\downarrow_0\rangle); \\
 & & \texttt{put}(Ty(e_s), Fo(U_{PAST}), ?\exists x Fo(x)); \texttt{go}(\langle\uparrow_0\rangle) \\
 & & \texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle); \texttt{put}(?Ty(e_s \rightarrow t)); \\
 & & \texttt{make}(\langle\downarrow_0\rangle); go\langle\downarrow_0\rangle; \texttt{put}(?Ty(e)); \texttt{go}(\langle\uparrow_0\rangle); \\
 & & \texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle); \texttt{put}(?Ty(e \rightarrow (e_s \rightarrow t))); \\
 & & \texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(?Ty(e)); \texttt{go}(\langle\uparrow_0\rangle) \\
 & & \texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle); \\
 & & \texttt{put}(Fo(\text{Hug'}), Ty(e \rightarrow (e \rightarrow ((e_s \rightarrow t))))) \\
 & & \texttt{go}(\langle\uparrow_1\rangle); \texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \\
 & \text{ELSE} & \text{Abort} \\
\text{ELSE} & \text{Abort} &
\end{array}\right.
$$

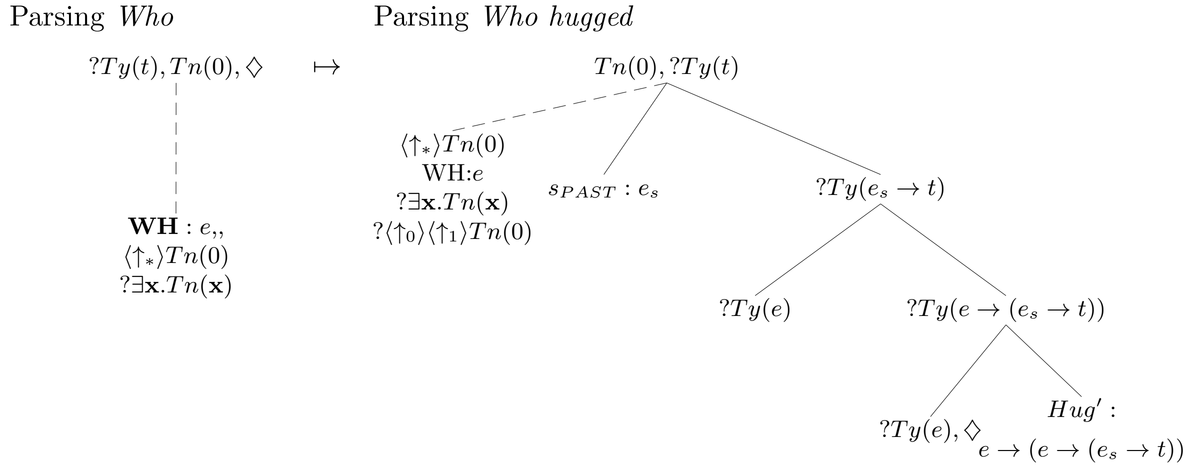
Relative to the prefixed condition, the subsequent sequence specifies step by step how to build the necessary tree structure, for each relation between nodes in such a structure, building that relation, going along it to the new node so constructed; decorating it, coming back to some more central point, and repeating this sequence of actions as many times as there are argument and functor nodes. The result is a skeletal propositional template associated with the predicate Hug', together with a partially specified event term (only sketched here), notably with only that predicate node having some content specification.

2.2 Combining actions in sequence

With this minimal sketch in place, we can see how structure is built up for (22). The first step listed is to induce an unfixed node from the underspecified tree in (23) by the actions specified in (24) yielding the tree specified there. This provides the context (an open term requirement) for the parse of *who* which annotates the node with a particularised metavariable, **WH** and a type specification. With the pointer returned to the topnode, the trigger for the processing of the verb *hugged* is satisfied, inducing the sequence of actions given in (26). These steps are shown in (27):

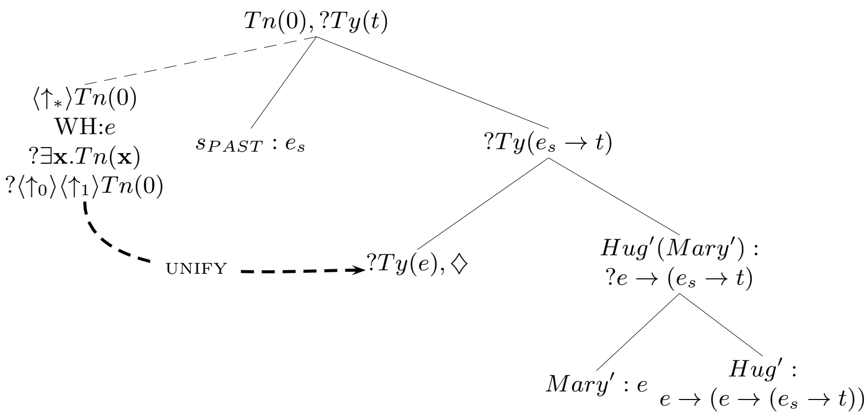
[6] These actions have been slightly simplified for expository reasons. See Cann (2011) for details and discussion.

(27)



This update leaves the pointer at the logical object node, ensuring that the syntactic object, *Mary*, can be parsed next to decorate the relevant node with type and formula information as in (25). One of the effects of the actions induced by the verb is that the unfixed node is required to be the logical subject given the position of the *wh* expression immediately before the verb, and this is achieved by a process of unifying the unfixed node with the open term node to yield (28) satisfying all outstanding requirements on the unified nodes. From this interim result, a simple sequence of applications of the computational action of modalised function-application is all that is needed to successively satisfy all nonterminal node requirements up the functor spine of the tree, giving (22) above:

(28)



Of course, there is more to be said, in particular because these actions build relative to context, with both hearer and speaker following the same sort of activity. The only difference between the dynamics of parsing and production is that, in parsing, there is only the relatively weak goal of trying to build some propositional formula relative to the utterance and context, while, in production, there is some more particular goal to be achieved, some (possibly partial) tree representing something to be communicated. So, for the speaker, every update associated with a putative word must be extendable towards that goal tree to allow that word to be produced. The only difference between being a speaker or a hearer is thus the checking by the speaker of the subsumption relation between the tree under construction and some entertained (partial) goal tree.[7] So from this perspective the shift from hearer to speaker and back again is seamless and unproblematic: both interlocutors entertain a partial tree representing the content of the utterance so far. To start

[7]The hearer may well also have expectations as to what is to follow in the form of some hypothesised extension of what they are currently parsing, but such anticipatory steps are not an essential accompaniment.

generating an utterance all a hearer has to do is extend the current parse state to act as a new goal tree and then start speaking, again checking output parse with the goal tree (Gregoromichelaki et al 2013, Gregoromichelaki and Kempson 2015). And this may go on as long and as often as the interlocutors want to. In the same way, a hearer does not have to recognise the speaker's intended communication before starting to speak: he merely has to decide to extend what has already been said into something he wants to say, whether or not he has anticipated what the speaker was intending to utter. In consequence, the model is entirely flexible with respect to our intentions to interact with each other and how big a role such intentions might play.

3 Cross-turn syntactic licensing: split utterances

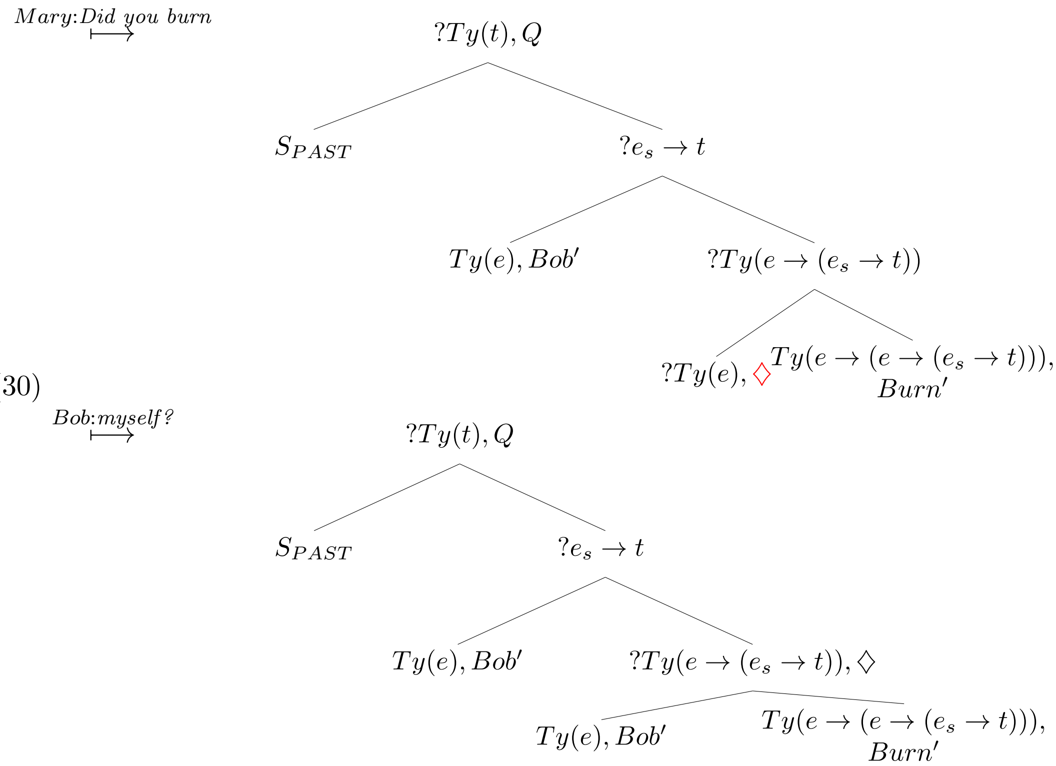
It is this free switching of rules which constitutes the split utterance phenomena, the most striking aspect of which is how the shift of roles from speaker to hearer and conversely can take place apparently splitting dependencies of all sorts. By way of illustration why such phenomena can, and must, be licensed cross-turn and distributed across speakers, we take morphosyntactic dependency between a reflexive anaphor and some locally available antecedent:

(29) Mary: Did you burn
Bob: myself? No.

(30) displays the partial tree upon Mary's utterance having been processed. At this point, with *burn* as well as *you* processed, Mary's utterance of the indexical pronoun *you* will have served to annotate from context the subject node with a conceptual representation of the current addressee, Bob. Given this as a context, Bob is now licensed to complete the utterance using the first-person reflexive, because this expression, by definition, induces a copy of the formula residing in a local co-argument node just in case that formula satisfies the conditions set by the person and number features of the uttered reflexive, i.e., the current speaker:[8]

[8]Bob's utterance here serves simultaneously as a completion and a clarification request reinforcing the DS claim that all such phenomena be treated uniformly by elaborating subpropositional types, for more details see Kempson et al (2007); Gargett et al (2008); Purver et al (2010); Eshghi et al (2015); Gregoromichelaki and Kempson (2015), Gregoromichelaki (2016).

(30)



The fact that the structure may have been induced in part by someone else is irrelevant. This phenomenon, as here modelled, is specific to the construal of the reflexive pronoun, said to be subject to a Principle A constraint (Chomsky 1986); and the associated "Principle B" constraint imposed on a nonreflexive pronoun debarring it from picking up an antecedent from this domain is equally simple to express in the same terms (Cann et al 2005; Gregoromichelaki 2006, 2013). In both cases, the morphosyntactic restrictions are expressible as constraints circumscribing the goal-directed search for achieving interpretation of linguistically-underspecified elements. Though this seamlessly achieved splitting of utterances can apparently separate off any expression from the syntactic environment it needs for its wellformedness, this effect is achieved because speakers and hearers both use the defined tree-growth mechanisms to construct a representation of what is being said taking the immediate context as input. The only difference between speaker and hearer is the additional requirement on speakers that the construction process has to be commensurate with some more richly annotated (possibly incomplete) structure corresponding to what they have in mind to say. This dynamic predicts that switching from parsing to production, and conversely, will be seamless, yielding the effect of in-tandem construction without needing to invoke higher levels of inference (Poesio & Rieser 2010) or superimposed duplication of the one type of activity upon the other (Pickering & Garrod 2013): each individual will simply be constructing the emergent structure relative to the context which they themselves have just constructed in their other capacity (Gregoromichelaki et al. 2011, 2013), so there is in effect no splitting of the dependency environment. Given the bidirectionality of the grammar, parser-generator switches are seamless and the resulting interlocutor coordination is secured by the subpersonal processing mechanisms of the grammar directly without needing to invoke high-order inference. Such shifts of roles from speaker to hearer are displayed even in supposed hard-core cases of syntactic dependencies, such as the classic "movement" cases of long-distance dependency:

(31) A: Mary, John …
B: trusts?

A: Well, she clearly likes him.

In this case what A and B severally as speaker and hearer have built at the point of switch is a partial structure. Following the switch, whether as speaker or as hearer, the particular form of the verb, here the finite form *trusts*, will fix the interpretation of $John'$ as logical subject (in virtue of the property of English finite verbs determining that the most local unfixed node already in the structure under construction must be construed as logical subject: see Kempson et al 2016 for details). The form of the verb will also fix the interpretation of $Mary'$ as object in virtue of the step of unification immediately after the processing of the verb *trusts*. Such restrictions will be applied by both speaker and hearer within the partial trees they have individually constructed within their respective activities whether as speaker or as hearer.

4 Extending content: Building relative to context

There is of course much more to add to have even a minimal indication of why DS constitutes a grammar formalism. In particular, the sketch so far has placed relatively little emphasis on the context. The exegesis in the main here presumes some propositional requirement leading to some output of type t. But, as these split utterances vividly demonstrate, speakers and hearers can start these activities from partial structures of whatever sort, so the concept of growth itself needs to be defined in these terms, for both content and context. Generalising then to partial trees as input or output from any task, we need to model how speakers may start out with only a partial thought in mind (with a partial tree as goal), may intervene with some partial contribution to the emergent structure (with partial trees both in context and as goal) and, equally, hearers may start with some partial tree as context, updating it as appropriate driven by the utterance of incoming words relative to that context. The minimal context for both parsing and production thus must include partial trees under construction,[9] the sequence of words so far uttered, and the sequence of actions so far used.[10]

The status of this composite context as a record of tree growth within the grammar is confirmed by the fact that languages all contain explicit devices for indicating links with some constructed context, more familiarly known as the head-adjunct relation. Examples include

(32) John, a lovely guy, is becoming Chairman of the company.

(33) John, who I admire, is becoming Chairman of the company.

The process of adjunction, by definition, involves the adoption of some term as a context from which to articulate some ancillary add-on – the adjunct – so that the result is the extension of some partial tree by additions which in some sense preserve the context provided by that simple primary tree, nonetheless adding additional "adjunct"-provided information to it. As (32)-(33) show, such adjuncts can be of different types, here a noun phrase and a relative clause. DS models this directly as a process of building so-called *linked* trees in which from any node once its requirement is satisfied, a LINK transition onto a newly emergent tree is licensed with a requirement that the result is a node sharing a term with the node at point of departure – the context relative to which this transition is licensed. The canonical case of these is relative clauses such as (33) which can be added on at any point in the overall tree being developed,

[9] The context will characteristically also include a wider set of structures, by assumption also modellable in tree-theoretic form.

[10] In Eshghi et al 2011, this is modelled via a mapping onto a Type Theory with Records formulation (TTR), but we suppress these details here: see Sato 2011; Hough and Purver 2014; Eshghi et al 2013; Hough 2015

(34) I walked up to John, who was ignoring me, with a gift for his wife.

(35) John, and not Mary, is about to be voted in as Chairman, who I warned about this some weeks back.

(36) I approached a man who had been annoying people at reception, with a request that he should leave.

(37) Someone came in to the lecturehall, who had been annoying people at reception.

On this analysis, such structures, which can occur anywhere in a string in English as long as relatable to that head, are modelled preserving Jespersen's insight of relative pronoun being an encoded quasi-anaphoric device (Jespersen 1949) ; on the DS view these add-ons are projected as independent structures through the addition of a LINK transition onto this newly initiated structure (Cann et al 2005), the relative pronoun in effect projecting a constraint on a transition across from one partial structure to another, which its actions induce.

Though as with relatives, the new structure is propositional in type, this is by no means a general constraint on adjunct structures, as witness (32) and (38), for which the requirement of a shared term at the content level of DS tree-construction is preserved, even though notably not observable from the sequence of words themselves:

(38) A_1: Mary's just returned.
B: From Germany?
A_2: Very tired.

We omit all details here, but merely note that the requirement of shared content is met by B's utterance having to be construed as the predicate 'returned from Germany' and A's follow-up utterance providing a new adjunction structure modifying the subject 'Mary'. An important further reflection is what this direct modelling of adjuncts allows. Because these loosely connected structures are *not* structures defining hierarchical and linear dependencies between words, but rather expressing links between structured concepts, they provide a vehicle for modelling a whole range of possibly covert inferences – ranging over propositional attitudes, politeness indications about either speaker/hearer or individuals described, and more generally the whole class of implications broadly characterised as implicatures (see Cann and Kempson forthcoming). The overall effect thus is that structure content and context all evolve, step by step. "Content" associated with words involves very much more than a mere itemising of conceptual content: to the contrary these macroed-sequences called up by a word characteristically project local structural update but yet invariably less than a fixed denotational content within a fixed structure.

The evolving nature of both context and content is modelled directly in DS through defining a composite evolving structure represented as a directed acyclic graph (DAG: Eshghi et al 2011, 2013, 2015; Hough 2015). This keeps track of choices available and selects amongst them as soon as possible, a device which, depicting transitions between processing states T_i graphically as in Figure 1, shows how burgeoning choices may yet serve to narrow down derivational choices as these become available. We take only the simplest sequence of steps here, that involved in processing the sequence of an auxiliary plus *you* construed as subject of the transitive verb *burn*, but even in such a short three-word sequence, additional options can be seen to open up but be swiftly closed off. In the case of split utterances then, as with all other utterances, it is the emergent DAG which serves as recording for speaker and hearer alike the shifting triples of context, words and content.

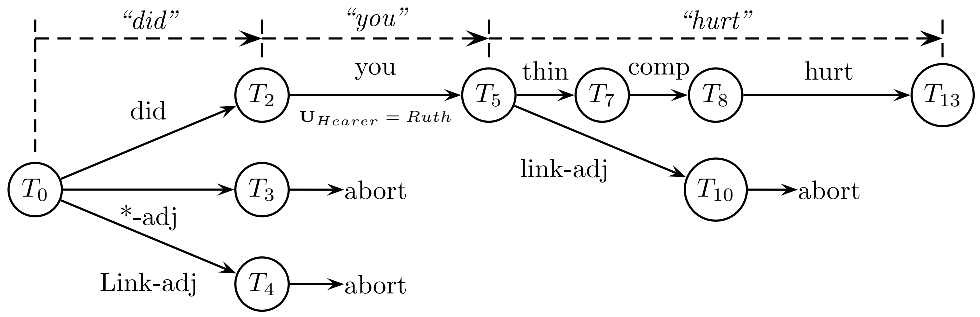


Figure 1: DS context plus content as a graph

A further bonus of incorporating the DAG within the formalism itself rather than as some external "performance-only" add-on, is that it provides a basis for self-correction, allowing local backtracking in the face of inconsistency to some first compatible point in the path, displaying precisely the fine-grained incrementality needed to model the flexibility of on-line production and parsing both in what are manifest corrections ((39),(42)), and others which may get identified as "apposition" structures, rather than the online revisions which they often serve as ((40)-(41)):

(39) The yell-uh purple square

(40) The yellow square, the one with a black dot on it,...

(41) The proper conclusion, the conclusion to be drawn from these arguments, is ......

We thus now have the basis for anticipating encoded lexical specifications for expressions such as *or rather* in (42) which are defined to make reference to local back-tracking:

(42) The yellow or rather orange square

Standing back from the details, what we now see is an action-based grammar, and a non-representationalist view of "syntax", syntax being transformed into a set of conditional update actions that induce or develop partial representations of content relative to context. These are prediction-driven strategies, inducing structure-developing options that get progressively developed or locally corrected, with both local and universal grammar-defined procedures that have the effect of inducing interaction and coordination effects. Furthermore, though there appears to be rich interaction between minds with clear capacity for inference and, at least, in principle sensitivity for other people's needs, successful communication does not need the capacity of either mind-reading or recognition of the other party's intentions before successful communication can be said to have taken place (contra Breheny 2006). To the contrary, the effect of interaction takes place simply because parsing and production use the same processes, as definitive of the DS formalism: no separate parsing and processing modules have to be co-defined explicitly for such interactive effect. And with grammars defined as mechanisms for information/action coordination, languages can be seen as tools for interaction.

5 Language as actions

In drawing such linguistics-internal arguments together, we note ongoing research in cognitive psychology, where embodied representations have been explicitly argued for as a means of explaining the tight physical coordination that exists between individuals engaging in manifestly interactive behaviour, where there is striking parallelism between the coordination demonstrably sustained in split utterance exchanges and these more general actions. In studying activity such as ballroom dancing, for example, Pezzulo and colleagues (Pezzulo 2011) have argued against invocation of higher order reasoning as the basis for such cases of rapid cross-modal coordination: to the contrary, they provide evidence that both action and perception involve low-level procedural mechanisms which determine agents' actions, being both predictive and simulating. The general framework for this embodied view is put forward by Andy Clark (Clark 2013, 2016), who argues that all cognitive behaviour involves constructive actions, so that perception and action both involve anticipations constructed relative to immediate and more long-term context in advance of the designated activity itself, the input stimuli merely acting as a filter on such anticipatory guesses, rebutting or confirming the expectations with which we greet the world. With even half an eye on this neighbouring research activity, the parallelism with the language case is striking: taking heed of the world of cognition as espoused by Pezzulo, Clark and colleagues has become an imperative.

5.1 Extending The Predictive Processing model

In the so-called Predictive Processing (PP) model of Clark 2016, the bottom up flow of information of the input system model advocated by Fodor (1983) and Marr (1982) is thus replaced in total by a generative probabilistic model of cognition, in which cognitive assumptions are created on the basis of the context, anticipations which constitute 80 percent of the explanatory load. Unlike the Fodor/Marr model in which the input stimulus is the point of departure, on the PP model, the incoming stimulus functions as a filter that checks for prediction errors to which that anticipatory input might give rise, contributing a mere 20 per cent load of the explanation. According to Clark, this is the pattern for the whole cognitive system. Perception, action (and imagination too) are all said to construct guesses about incoming signal: what they differ in is the response to any recognised inconsistencies. Perception involves a filter weeding out impossible and unlikely construals. Action involves physical action to make true some hypothesis, in order to avoid what would otherwise be filtered out on the appropriate anticipatory hypothesis as transparently not matched by the input stimulus. And imagination involves a suppression of the inconsistency filter itself that would otherwise involve either rejection or action. In such a system, as in DS, context is everything. Brains are not the passive modular input systems of Marr (1982). To the contrary, they are predictive engines using their own immediate and encyclopedic context at every step to guess the structure/shape of the incoming sensory array, whether for perception or action. The status of the incoming signal, which changes moment by moment, is on this view substantially reduced from that attributed to it in the static multi-level hypothesis of Marr. Instead, there is an affordance competition hypothesis whereby the brain specifies in parallel several potential actions which compete against each other for further processing relative to probabilistic weightings (these determined by iterative experiences, cognitively or normatively culturally imposed), until, standardly, all but one of these is winnowed out. While the model Clark depicts is richly articulated, being a claim about cognition in general, his view is strikingly a generalisation of the DS stance: language, like other cognitive activities, on this view involves the active construction of assumptions in real-time as the grounding for all language-based activity.

In the light of this parallelism, the relatively weak characterisaton of language in Clark (2016) comes as a surprise. He draws current research on language processing together to obtain a conclusion that is both too strong and too weak. It is too strong because Clark argues against the necessity of invoking a so-called efference-copy view of perception and its counterpart action. Nonetheless, Clark cites with approval the Pickering and Garrod (2013) efference-copy view of the shared activity involved in split utterance processing.[11] Furthermore, despite rejecting inferentialist views of perception/action in favour of his own PP model in which embodied mechanisms carry the explanatory load of how cognitive activity takes place, he nevertheless suggests that language enables top-level "script sharing" and "brain to brain coupling", along lines of the "shared brain" view of Frith 2007 (though see the modified view of Friston and Frith 2015);[12] but these variants are all solidly inferentialist, being grounded in Gricean assumptions of meaning, with the additional twist for the associated explanation of language acquisition of requiring that the capacity for such higher-order mind-reading capacity be in place prior to the very earliest stage of language acquisition, as a pre-requisite for language learning to take place at all, a view which both Tomasello (2008) and Christiansen and Chater (2016) embrace. So the view of language which Clark promulgates has all the richness which his own PP model signally does not endorse. On the other hand the characterisation of language processing in Clark 2016 is too weak, in that it fails to capture the coordinated social interaction endemic to language use because the language perspective he presumes upon preserves the view of language as a statically available body of knowledge. This means that he misses the parallel between language and cognitive architectures, missing the potential for replacing top-down inferential accounts with emergent systemic effects. The answer to this conundrum is straightforward: what the Predicting Processing perspective lacks and needs is the rounding out of the PP model to incorporate a model of language as evolved for interaction, such as DS.

5.2 The DS/PP Model

Bringing the DS characterisation of language into the PP landscape, the unitary characterisation of language as a sub-part of a Predictive Process cognitive system that is specialised for its interaction effects emerges wholly naturally, and the anomalies of the Clark 2016 discussion of language are resolved. Both the DS model of language and the PP model of cognition are fundamentally action-oriented, involving predictive moment by moment processing, evolving context dependency at every level with continuous intermingling of production, thought, encyclopedic and immediate context, with choices made relative to the individual's own context. In this system, choices between competing hypotheses are made with respect to an individual's own context. There is no cross-individual replication or "mutual knowledge" constraint: the dynamic is hypothesis-construction relative to context plus error-correction narrowing down possible alternatives, a dynamic by claim underpinning all cognitive activity. Despite close parallelism between complementary activities such as perception and action, no "efference-copy" mechanism is required to duplicate the one constructive process relative to the other, nor, as Clark explicitly notes, is any higher-order inference required by either party conjecturing another's mental states. What DS brings to the Predictive Processing perspective is that the very same PP dynamic extends to cross-individual exchanges simply in virtue of the construal of language as a mechanism for information growth, whether in parsing (perception) or production(action). The hearer and the producer in any utterance exchange are both engaged in the action of structure building and it is for this reason that at any point in the utterance sequence, they can exchange roles – in both cases, the process of structural growth will simply be proceeding from where they had got to. Following the Clark parallelism between perception and action, parsing can now be seen as hypothesis formation constrained by an associated acoustic filter eliminating impossible interpretations; and production can be seen as

[11]It is notable that their account fails to provide more than a partial account of the phenomenon, capturing as it does only the "helpful completion" type of data where the interrupting participant is actively seeking to construct the completion their collaborator seems to need help with.

[12]Friston and Frith (2015) retract from the Friston "social brain" view to one in which co-ordinated behaviour involves an abductive step of self-other identity.

hypothesis formation in a richer, plan-driven, enterprise, involving transmission to the realisation of bodily actions to make true a correspondence between some assumption and its outcome. So though the two counterparts may be seen as feeding into distinct physiological realisations, in other frameworks taken to constitute self-evident grounds for the disjointness of these two inverse activities (Hendriks 2014), on Clark's view they are both driven by the same pressure to make true the pairing of structure and output, and so essentially integrated through a common mechanism. Hence, on this joint DS/PP view, syntax is an embodied skill consisting of coupled interlocutor actions for incremental predictive processing in context. [13] The effects of interaction so transparently displayed in conversational dialogue can thus be seen as emergent from the system with no essential high level inference or mind reading, nor any risk of implicitly invoking an invisible agent, this the well-known homunculus charge of Putnam (1967).

There is a further commonality. The specifically probabilistic generative model, with shifting probabilities at every level as filtered by the carrying forward of putative prediction errors, which might seem to distinguish the PP model from DS, is to the contrary, reflected in the most recent DS implementation (Eshghi et al 2013, 2015; Hough and Purver 2014; Hough 2015) by the addition of probabilistic weightings into the language system as part of the macro sequences of actions stored in the lexicon, completing the parallelism. As Clark's perspective would require, the DS system has thus been developed to incorporate a built-in self- and other-correction mechanism which yields step-wise weeding out of options that are probabilistically implausible as well as categorically impossible (see Bickhard 2009 who claims this is diagnostic of an interactive system). Such weightings derive from the accumulated weight of iteration, personal and even social experience through exposure to cultural norms that dictate perspectives for understanding phenomena in context.[14]

Finally there is the constraint imposed by the cognitive system itself, driven by economy for some given effect, a property often invoked by pragmatists as determining communicative success, but here seen as the natural emergence of the way the brain has evolved to determine success in its task of providing a filter on the wealth of incoming information to determine that its task in real time is manageable (what Clark refers to as the lazy brain hypothesis, citing Herbert Simon 1956 for the original idea). No longer do linguists have to presume that the performance factors involved in communication of necessity involve steps of higher-order inference constrained by assumptions of normativity (the Gricean perspective: Grice 1975) or by cognitive constraints imposing limits on such inferences as stored in some pragmatics module (the relevance-theoretic perspective: Sperber and Wilson 1995, Sperber 2002). On the new view such constraints are the constraints of the cognitive system in general, imposed on all cognitive activity, being implementations of system-internal interdependencies (Clark 2016 among others). On the combined DS/PP perspective, the phenomenon of a language and its processing, at any one time are, indeed, not essentially distinct: both are ephemeral ever-changing constructs. What is relatively constant is the capacity for language, and the cognitive system more generally, this being the predictive device which constructs hypotheses for action/perception, systematically able to correct any such puta-

[13]The existence of "prosodic words" which are signally not noticed by either producer or hearer, suggest that the mechanisms for realisation are a different level of abstraction than the dynamic core of language itself. In claiming that 'she's a genius' by uttering the prosodic word /ʃiːzə/ the speaker will be building a structure commensurate with the thought that Mary, let us say, is a genius, without any hiccough caused by the fact that his articulation is at best apparently nonoptimal. On the other hand, the hearer, processing the speaker's utterance only has to be able to retrieve the appropriate interpretation from the very high probability, given its very common occurrence, that this word comprises the sequence of 'she' plus 'is' plus 'a', so both parties will be successfully building up the thought that Mary is a genius without the lack of the supposedly indicated lexical entry causing any impediment to the success of the exchange.

[14]See the neuro-science perspective as reported in Rose and Abi-Rached (2013) in which the brain is demonstrated to be responsive to iterative cultural processes and so subject to both short- and long-term context-relative change.

tive actions/judgements relative to its inbuilt correction device, hence enabling learning, ongoing interaction and modification.

6 Evolution: Surfing Uncertainty together?

With language now seen as a tool for interaction, the perspective from which to develop the view of natural language as an evolved adaptive system shifts into a less inferentialist perspective while remaining broadly functionalist. Natural languages are perhaps the quintessential case of adaptivity, and the view that language is an adaptive system was first suggested by Darwin. The adaptationist claim is disputed by some (Fitch, Hauser and Chomsky 2005, Lightfoot 1999) and strongly advocated by others (Christiansen and Chater 2016 among others). According to the non-adaptationist view, the architecture of the mind/brain determines the form of natural language in virtue of domain-specific attributes of a language-specific module nested in the brain determining the limits shared by all natural languages, which could not be subject to any gradualist account. Pinker and Bloom (1990), despite defending an encapsulated language module nonetheless defend an adaptationist stance. Others claim to the contrary that language is a product of cultural evolution; and that languages have developed in the light of pressure from cultural and pragmatic (inference-driven) pressures to match the needs of our brain, e.g. evolving in the direction of progressive lowering of cognitive costs in order to meet the pressures for fast processing which are imposed by a system evolved to make very rapid choices within a very narrow and endlessly shifting window of time (Christiansen and Chater together with Hawkins (2004), Bybee (2010) and other construction theorists). These latter proposals all pursue an essentially functionalist form of explanation - language as shaped by the brain.

There are many intermediate positions in this debate; yet these varying stances characteristically leave essential questions unaddressed. Though they address (and disagree about) the nature of human language in the current evolutionary window of time, abducing from that to some pre-language state, by and large they do not address the “how questions”: how the properties particular to language such as its notable signal-content correlations came into being; how these attributes evolved alongside the emergent general human potential for context-relative inference, before, after, or in tandem; and so on. Yet, as Kirby et al (2008) rightly stress, essential to any characterisation of language as an adaptive system is a specification of the nature of its interdependence with other adaptive systems with which it interacts, notably the human biological system, the human rich cultural capacity, and the human capacity for learning which gives rise to the adaptivity of language through use. As Scott-Phillips et al (2009) have emphasised, studies of language evolution have to address such questions as how signalhood itself evolved, as they put it “How does a signal signal its own signalhood”. To address these questions, they develop a significant methodology for experimental testing, via both computational and cognitive-psychological test methodologies.

Despite often marked differences between these various stand-points, two assumptions nevertheless have remained common to all: (i) that establishing signalhood, i.e. string content correspondences has to involve assigning structure to the natural language strings themselves with compositionality to be defined over those strings; (ii) that successful communication requires establishing some fixed signal-content correspondence as intended by the speaker to be recovered by the hearer.[15] The new dimension that DS brings to the debate is the claim that it is interaction which is critical to successful communicative exchanges rather than agreement of the necessity to recover some priorly fixed signal-content correspondence which the speaker of the utterance might

[15] Kirby et al modify this assumption in providing evidence that in cross-generational acquisition, positing of polysemies enhances rather than disadvantages language learning/change (Kirby et al 2008).

have in mind, and the attendant claim that language is not so much a set of string-structure pairings which have to be subject to interpretation but a set of processes for building up structure-content correspondences relative to each participant's own (structured) context.

The starting point for exploring this shift in perspective is the very familiar observation that all utterance exchanges will, on DS assumptions, necessarily and visibly involve moment by moment interaction between participants in the communicative activity as they severally build up structured representations of content. But we can go further than this by extrapolating from first language acquisition exchanges between carers and infants the evidence for activities that can be said by abstraction also to be available to human agents at an evolutionary stage immediately previous to language emergence. Before any of the emergence of language, higher primates were displaying some potential for interactive behaviour, often but not exclusively associated with females with the nurturing of offspring to maturity, grooming, etc (as witness the many tv programmes we are currently inundated with). We can thus legitimately presume that the capacity for interaction itself was in place at the time of emergence of humans as a species. What was missing and had to be secured, was the recognition by the parties of the manifestness of such interactivity.

Without seeking to trivialise the leap involved in attribution of content to signals (and from there to sign and symbol recognition), nonetheless, we can glean a clue from the peekaboo games which pre-linguistic children so much enjoy, where there is no attribution of content. These illustrate the fluency with which pre-language children can join in on interactional exchanges despite there being no goal of estabishing shared content, the sole goal being the pleasure of the interaction as the much-enjoyed effect manifestly displayed by both participants.[16] We can thus relatively safely assume that interaction and manifest pleasure in such interaction were in place before the emergence of language. There is another clue from acquisition, for, in the next phase, building on the joint manifest activity and the pleasure in doing so continues, despite asymmetry between participants. As reported by Hilbrink et al (2015) and Clark and Cassillas (2016), in the pre-language stage of the infant, the carer characteristically uses her language fluently, even though possibly self-consciously simplified, invariably interpreting the non-language-based verbalising behaviour of the child within her own conception of what this activity leads to, or if the babbling infant is the instigator of some co-activity by providing some take-up which she as carer can construe as some opening request/query to which she provides a response (Clark and Cassillas 2016 report rich exchanges taking such forms between a Korean carer and her child). The invariable result is an interactive effect between the child and carer even in the absence of any expectation of the child's understanding of the exchange - the sole reward of this joint activity for both being the rich bonding achieved by this interactive behaviour.[17]

Out of this interactive behaviour, the emergent language speaking capacity can be seen to emerge, not upon the necessity of having already in place an established rich theory of mind, as supposedly needed to secure agreed recognition of some intended content (Tomasello 2008 and elsewhere), but from the reiteration of these exchanges over the course of the days in which these semi-ritualised shared activities of feeding, nappy- changing etc take place. Events for a pre-language-speaking child follow a regular pattern, reiterated at regularly recurring intervals each day; and for each such case, both child and adult build their own private representation of what is offered them and what they themselves engage in as overt actions. For the child this will mean either

[16] An anecdotal report of one such exchange witnessed by RK confirms the speech-act efficacy of such exchanges. A pre-language infant in her push-chair was persisting with great pleasure in responding in a peekaboo exchange with an older language-controlling child. After several such exchanges, the child broke off saying "Stop bullying me" and rushed back to his mother.

[17] Communication in language exchanges between adults in highly asymmetric native-speaker and struggling second-language participants displays a similar pattern, as does the exchange of fluent speakers in speaking with people with language difficulties caused by brain trauma.

providing some completion to a schematic representation just provided by the adult's utterance, or if if the child is initiating some exchange with such a fragment it is on the expectation that the adult will then develop it (as in (5)-(6) above). In the earliest occurrences of language, both producing and processing such a construction would, following the anticipatory Clarkian dynamic, involve constructing assumptions or putative actions filtered by the input stimulus. In recurrent occurrences of such scenarios, all these actions will thus invariably be buttressed by the fact that the embodied language capacity necessarily coordinates and in this sense duplicates speaker and hearing actions. So no matter how disjoint or asymmetric the construals of these events by the several participants may be in these exchanges, the duplicative effect of participant actions as either parser or producer will determine regular reiteration by the individual of the pattern of activity, thereby familiarising the language-acquiring child with the language input of the carer, with the invariable reward of achieving the emotionally bonding interactive effect. And, by the very same pattern, each further addition of a word plus the update it encodes will multiply the possible cross-individual interactions. In all these cases, at each stage, high-level inferences or other-self abductions do not need to be invoked (contra Sperber and Wilson 1995; Frith 2007; Friston and Frith 2015). Learning can moreover still take place without clarification requests or corrections (though see Saxton (2010) for their efficacy in the learning progression). Acquisition and evolution alike can thus be seen to emerge on a gradualist approach without recourse to either rich innate language structure or an a priori rich mind-reading capacity.

Furthermore, with the parse DAG presumed to be integral to the DS system, the language system is defined as having an inbuilt correction mechanism, which, as Bickhard 2009 notes, is quintessential to learning. We use language throughout our day to day activities, so iteration of such interactions is ensured as needed for consolidation. Such iterations of language continue month after month, and, with successive reiterations, routinised sequences may then become stored as a macro unit sequence of actions. The triggers for these sequences of action will often be stored economically in reduced phonetic form, and cued by intonation/consonant clusters and other prosodic devices for ease of retrieval.[18] With all such cross-party reiterations, group effects can be seen as emerging without having to invoke the normativity which is implicit in Gricean accounts of communication and overtly postulated in Bickhard (2009) as an external normatively imposed pressure.[19]

It is important to stress however, that none of these exchanges *preclude* putative mind-reading activities. To the contrary, such tightly coordinated activity between pairs/groups of people undoubtedly buttresses the manifest group-consolidating effect of all language, for it consolidates groups of individuals who give evidence of sharing a system sufficiently closely aligned that the lack of perfect correlation may well go unnoticed, while nevertheless excluding from such groupings all those who fail to command such a set of procedures. But, importantly, what the communicative effectiveness of small children and the split utterance data demonstrate is that such mind-reading is not integral to the core of the communicative ability which the language device so signally successfully drives. In consequence, the evolutionary system notably does not need to have in place any theory of mind prior to the emergence of language (contra Tomasello 2008). To the contrary, the very activity of language processing, because it involves an intrinsically interactive exchange, will

[18] In this, the diachronic emergence of clitic clusters is illustrative - see Bouzouita 2008, Chatzikyriakidis and Kempson 2011.

[19] In this paper, we leave on one side the arguments of Sober and Sloan Wilson 1998, Sloan Wilson 2002 in which they argue that groups form adaptive units whose status competes with adaptivity at the individual level, with optimal evolution involving a trade-off between group level and individual-level adaptivity, successful groups evolving to determine control over individual behaviours in order to ensure optimal group-level adaptivity. However we note in passing that human language constitutes a striking counter-example to the need to temper the normal "selfish" pattern for survival at the individual level with intra-group mechanisms of control, as, both at individual and group levels, capacity for language is strikingly advantageous.

guarantee such group effects without any prior aspiration of altruistic group-consolidating coordination. All that is required to trigger the process is the recognition of the behaviour as a manifest signal of interaction. The only shift that needs to have emerged is the means of making manifest such desire for interaction (unlike Scott-Phillips et al 2009).[20] In this, our account broadly follows Christiansen and Chater 2016 in seeing language as an evolving device emerging along the same pathway as language change. However, unlike Christiansen and Chater, there is no commitment to higher-order inference mechanisms underpinning communication having to be in place prior to language development, a stance notably confirmed by the research reported here: the need to invoke shared agreed content either as the result or as the grounding for successful communication has been bypassed.[21] And as Kirby et al (2008) demonstrate, the emergence of structure and compositionality strikingly emerge without any intentional recognition of its efficacy, emerging simply through re-iteration across generations.

So finally, language enables us to "surf uncertainty together", to adopt Clark's revealing metaphor, even though characteristically building representations of content on the basis of our own structured context. On this view, we have left behind the necessity of invoking a rich set of unexplained innate properties. To the contrary, innateness is reduced to an inbuilt tendency towards interaction, a minimal Baldwin effect confirmed by the manifest success of group-forming activities to consolidate their evolutionary status. No hidden homunculus threatens, or any rich in-built structural-design feature. We have also left behind the need for a sudden-switch explanation of how language evolved, for language evolution is seen in the same terms as acquisition, an essentially interactive process (Clark, E. 2009; Clark E. and Estigarribia 2011). And there is no recourse to a problematic attendant "social brain" hypothesis as invoked in passing by Wilson 2002. Finally, we have now no need to invoke language evolution as being in principle different in kind from that of historical change: all three can be analysed in the same gradualist terms. The result is that we can now see language itself as an evolving subsystem of the general cognitive system, from whose activities group effects can emerge without ratified common-ground or any concept of a shared world view.

The explanation we have provided is, it must be declared, manifestly functionalist; but it is not circular (the charge of Sperber 1996). Indeed, we urge that it is now high time to set aside the dismissal of functionalist arguments out of hand that have been common-place in linguistics (Berwick and Chomsky 2016, Newmeyer 1988), as in many other disciplines (Sober and Sloan-Wilson 1998, Sloan-Wilson 2002). The closing hope is that in now seeing language as intrinsically a tool for interacting in real time (Kempson et al 2016, 2017a, 2017b), we have the chance to explain why and how the language capacity might have developed inexorably.

[20]The experiment in Scott-Phillips et al (2009) presumes that successful interaction determines some agreed content, conflating the determination of interaction between contributing parties with the determination of attributable content.

[21]It is notable that discussion of binding principles aside, for which Christiansen and Chater give a pragmatic account following Reinhart 1983 (which in being Gricean would be counter-intuitively underlyingly normative), their exegesis of their own Darwinian account of evolution of language as having adapted to our brains makes no further reference to pragmatic considerations, despite the accumulated evidence they provide of how processing-cost constraints have affected language over a long time-scale, a null return which itself suggests that the specifically Gricean form of communicative constraints underpinning their own account plays little or no role in the explanation. And, with respect to the anaphoric binding principles, an alternative account might take the form of emergent encodings of local boundary demarcations, themselves having the functionality of determining the phrasal chunks which participants need to construct during the utterance process, an account fully commensurate with the lazy brain hypothesis (see Cann et al 2005, Kempson et al 2016 for a DS formulation of reflexive and pronominal anaphora).

References

Atlas, Jay. 1977. Negation, ambiguity, and presupposition. *Linguistics and Philosophy* 1.3 21-36.

Austin, J.L. 1962. *How to do things with words.* London, Oxford University Press.

Bach, Kent. 1994. Conversational impliciture. *Mind & Language 9(2)*, 124-162.

Berwick, Robert & Noam Chomsky, 2011 *Why Only Us? Language and Evolution* MIT Press.

Bickhard, Mark. 2009. The interactivist model. *Synthese* 166. 547-591

Blackburn, Patrick & Wilfried Meyer-Viol. 1994. Linguistics, logic and finite trees. *Logic Journal of the Interest Group of Pure and Applied Logics*, 2(1):3–29.

Bouzouita, Miriam. 2008. *The Diachronic Development of Spanish Clitic Placement.* Ph.D King's College London.

Breheny, Richard. 2006. Communication and folk psychology. *Language and Mind* 21. 74-107.

Bybee, Joan. 2010. *Language, Usage and Cognition.* Cambridge: Cambridge University Press.

Cann, Ronnie. 2011. Towards an account of the English auxiliary system. In Kempson, R. Gregoromichelaki, E. & Howes, C. (eds.) *The Dynamics of Lexical Interfaces.* CSLI Publications: 279-317.

Cann, Ronnie & Ruth Kempson. forthcoming. What do words do for us? *Diachronica.*

Cann, Ronnie, Ruth Kempson & Lutz Marten. 2005. *The Dynamics of Language.* Oxford, Elsevier.

Chatzikyriakidis, Stergios & Ruth Kempson. 2011 Standard Modern and Pontic Greek Person Restrictions: a feature-free dynamic account. *Journal of Greek Linguistics* 11(2): 127-166.

Chomsky, Noam. 1965. *Aspects of the Theory of Syntax* Cambridge, MA, MIT Press.

Chomsky, Noam. 1986. *Knowledge of Language: Its Nature, Origin and Use.* Westport, CT, Praeger.

Christiansen, Morton & Chater, Nick. 2016 *Creating Language: Integrating Evolution, Acquisition and Processing* Cambridge, Mass: MIT Press.

Clark, Andy. 2013. Predictive brains, situated agents, and the future of cognitive science. *Brain and Behavioral Science* 36, pp. 181-204.

Clark, Andy. 2016. *Surfing Uncertainty: Prediction, Action and the Embodied Mind.* . Oxford: Oxford University Press.

Clark, Eve. 2009. *Language Acquisition* Cambridge: Cambridge University Press.

Clark, Eve & Marisa Casillas. 2016 First language acquisition. In Keith Allan (ed.) *The Routledge Handbook of Linguistics* pp. 311-329. New York: Routledge.

Clark, Eve & Estigarribia, Bruno. 2011 Using speech and gesture to introduce new objects to young children. *Gesture*, 11, 1–23.

Clifton, Charles, Antje Meyer, Lee Wurm & Rebecca Treiman. 2012. Language comprehension and production. In A. F. Healy & R. W. Proctor (Eds.), Vol.4 *Experimental Psychology. Volume 4* in I.B. Weiner (Editor-in-Chief), *Handbook of Psychology*, New York, Wiley.

Cross, I. 2012. Music as a social and cognitive process. In P. Rebuschat, M. Rorhrmeier, J. A. Hawkins & I. Cross (eds.), Language and music as cognitive systems (pp. 313-328). Oxford: Oxford University Press.

de Villiers, J.& H. Tager Flusberg. 1975. Some facts one can't deny. *Journal of Child Language* 2. 279–286.

Eshghi, Arash, Matthew Purver & Julian Hough. 2011. DyLan: parser for Dynamic Syntax.

Technical Report, Queen Mary, University of London.

Eshghi, Arash, Matthew Purver, Julian Hough. 2013. Probabilistic Induction for an Incremental Semantic Grammar. In *Proceedings of the 10th International Conference on Computational Semantics (IWCS)*, Potsdam, March 2013.

Eshghi, Arash, Christine Howes, Eleni Gregoromichelaki, Julian Hough & Matthew Purver. 2015. Feedback in conversation as incremental semantic update. *Proceedings of the 11th International Conference on Computational Semantics (IWCS)*, London.

Fitch, W.Tecumseh, Marc Hauser, & Noam Chomsky. 2005. The evolution of the language faculty: clarifications and implication. *Cognition* 97. 197–210.

Fodor, Jerry A. 1983. *The Modularity of the Mind*. Cambridge, MA, MIT Press,

Friston, Karl & Frith, Christopher. 2015 A duet for one *Consciousness and Cognition* 36. 390–405.

Frith, Christopher. 2007. *Making up the Mind: How the Brain Creates our Mental World*. Oxford: Wiley-Blackwell.

Gargett, A., Gregoromichelaki, E., Howes, C. & Sato, Y. 2008. Dialogue-grammar correspondence in Dynamic Syntax. In *Proceedings of the 12th Workshop on the Semantics and Pragmatics of Dialogue. (LONDIAL'08).*

Gregoromichelaki, Eleni. 2006. Conditionals in Dynamic Syntax. Ph.D King's College London.

Gregoromichelaki, Eleni. 2012. Review of J. Ginzburg (2012) *The Interactive Stance*. *Folia Linguistica* 47:293-316.

Gregoromichelaki, Eleni. 2013. Clitic left dislocation and clitic doubling: A dynamic perspective on left-right asymmetries in Greek". In Gert Webelhuth, Manfred Sailer, Heike Walker (eds.) *Rightward Movement in a Comparative Perspective*. Berlin: John Benjamins.

Gregoromichelaki, Eleni (2016). Quotation in dialogue. In Paul Saka & Michael Johnson, M. (eds.) *The Pragmatics of Quotation*. Springer.

Gregoromichelaki, Eleni & Ruth Kempson. 2015. Joint utterances and the (split-)turn taking puzzle. In Alessandro Capone & Jakob L. Mey. *Interdisciplinary Studies in Pragmatics, Language and Culture*. pp. 703–744. Dordrecht: Springer.

Gregoromichelaki, Eleni, Ruth Kempson, Matthew Purver, Gregory J. Mills, Ronnie Cann, Wilfried Meyer-Viol & Patrick G. T. Healey. 2011. Incrementality and intention-recognition in utterance processing. *Dialogue and Discourse*, 2(1): 199–233.

Gregoromichelaki, Eleni, Ruth Kempson, Christine Howes & Arash Eshghi. 2013 On making syntax dynamic: the challenge of compound utterances and the architecture of the grammar. In: I. Wachsmuth, J. de Ruiter, P. Jaecks, and S. Kopp (eds.) *Alignment in Communication: Towards a New Theory of Communication*. John Benjamins.

Grice, Paul. 1975. Logic and Conversation. *Syntax and Semantics*, 3(S 41):58.

Hawkins, John. 2004. *Efficiency and Complexity in Grammars*. Oxford: Oxford University Press.

Hendriks, Petra. 2014. *Asymmetries between Language Production and Comprehension*. Dordrecht: Springer.

Hilbrink, Elma, Herideth Gattis & Stephen Levinson. 2015. Early developmental changes in the timing of turn-taking: a longitudinal study of mother-infant interaction. *Frontiers in Psychology* 6, 1492.

Hough, Julian. 2015. Modelling Incremental Self-Repair Processing in Dialogue. Ph.D Queen Mary University of London.

Hough, Julian & Matthew Purver. 2014. Probabilistic type theory for incremental dialogue processing. in *Proceedings of the EACL 2014 workshop on Type Theory and Natural Language Semantics*, pp. 80-88. ACL 2016, Gothenburg Sweden

Jespersen, Otto. 1949. *A Modern English Grammar on Historical Principles, Part III Syntax.* New York, London: Routledge.

Kamp, Hans & Uwe Reyle. 1993. *From Discourse To Logic.* Dordrecht, Kluwer Academic Publishers.

Kempson, Ruth, Wilfried Meyer-Viol & Dov Gabbay. 2001. *Dynamic Syntax: The Flow of Language Understanding.* Oxford, Blackwell.

Kempson, R, Gargett, A. & Gregoromichelaki, E. 2007. Clarification requests: an incremental account. In Proceedings of the 11th Workshop on the Semantics and Pragmatics of Dialogue (Decalog 2007).

Kempson, Ruth, Ronnie Cann, Arash Eshghi, Eleni Gregoromichelaki & Matthew Purver. 2015. Ellipsis. In Lappin, S., Fox, C (eds.) *Handbook of Contemporary Semantics 2nd editn.* 114–140. New York, London: Wiley-Blackwell.

Kempson, Ruth, Cann Ronnie, Gregoromichelaki, E. Chatzikyriakidis, S. 2016. Language as mechanisms for interaction *Theoretical Linguistics* 42(3-4), 203-276

Kempson, Ruth, Ronnie Cann, Eleni Gregoromichelaki & Stergios Chatzikyriakidis. 2017a. Grammar-based action. *Theoretical Linguistics* 43(1–2).

Kempson, Ruth, Ronnie Cann, Christine Howe, & Stergios Chatzikyriakidis. 2017b Cognitive science, language as a tool for interaction, and a new look at language evolution. Proceedings for *Formal approaches to the dynamics of linguistic interaction* ESSLLI 29, Toulouse.

Kirby, Simon, Kenny Smith & Hannah Cornish. 2008. Language, learning and cultural evolution: how linguistic transmission leads to cumulative adaptation. In Robin Cooper & Ruth Kempson (eds.)*Language in Flux; dialogue coordination, language variation, change and evolution.* 81–108. London: College Publications.

Lerner, Gene 2004. On the place of linguistic resources in the organization of talk-in-interaction: Grammar as action in prompting a speaker to elaborate. *Research on Language and Social Interaction*, 37(2) 151-184.

Lightfoot, David. 1999. *The Development of Language: Acquisition, Change and Evolution.* Oxford: Blackwell.

Marr, David. 1982. *Vision.* MIT Press.

Miller, Jim & Regina Weinert. 1998. *Spontaneous Spoken Language.* Oxford, Clarendon Press.

Newmeyer, Fritz. 1998. *Language Form and Language Function.* Oxford: Oxford University Press.

Pezzulo, Giovanni. 2011. Shared representations as coordination tools for interaction. *Review of Philosophy and Psychology*, 2(2):303-333.

Pickering, Martin & Simon Garrod. 2004. Toward a mechanistic psychology of dialogue. *Behavioral and Brain Sciences* 27(2):169-226.

Pickering, Martin & Simon Garrod. 2013. An integrated theory of language production and comprehension *Behavioral and Brain Sciences*, 36(4):329–347.

Pinker, Steven & Paul Bloom. 1990. Natural language and natural selection. *Brain and Behavioral Sciences* 13. 707–784.

Poesio, Massimo & Hannes Rieser. 2010. Completions, co-ordination and alignment in dialogue. *Dialogue and Discourse* 1: 1–89.

Purver, Matthew, Ronnie Cann & Ruth Kempson. 2006. Grammars as parsers: meeting the dialogue challenge. *Research on Language and Computation* 4, 289-326.

Purver, Matthew, Eleni Gregoromichelaki, Wilfried Meyer-Viol & Ronnie Cann. 2010. Splitting the I's and Crossing the You's: context, speech acts and grammar. In *Proceedings of the Semantics and Pragmatics of Dialogue(PozDial)*. Poznan, Poland. June 2010.

Putnam, Hilary. 1967. The nature of mental states (originally published under the title 'Psychological Predicates') In *Mind, Language and Reality: Philosophical Papers, Volume 2*. Cambridge: Cambridge University Press.

Reinhart, Tanya. 1983. Coreference and bound anaphora. *Linguistics and Philosophy* 6. 47–88.

Rose, Nikolas & Abi-Rached, Joelle. 2013. *Neuro: The New Brain Sciences and the Management of the Mind.* Princeton: Harvard University Press.

Sato, Yo. 2011. Local ambiguity, search strategies and parsing in DS. In R. Kempson, E. Gregoromichelaki & C. Howes (eds.) *Dynamics of Lexical Interfaces*, Chicago: CSLI Press, 201-228.

Saxton, Matthew. 2010. *Child Language: Acquisition and Development*. Sage Publications.

Scott-Phillips, Thomas, Kirby Simon, & Ritchie Graham. 2009. Signalling signalhood and the emergence of comunication. *Cognition* 113(2):226-233.

Simon, Herbert. 1956. Rational choice and the structure of the environment. *Psychological Review* 63(2), 129–138.

Sober, Elliott, & Wilson, D.S. 1998. *Unto Others. The Evolution and Psychology of Unselfish Behavior*. Boston Mass: Harvard University Press.

Sperber, Dan. 1996. *Explaining Culture: A Naturalistic Approach*. Oxford: Blackwell.

Sperber, Dan. 2002. In defense of massive modularity. In E. Dupoux (ed.) *Language, Brain and Cognitive Development: Essays in Honor of Jacques Mehler*. 47-57. Cambridge MA: MIT Press.

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

Stainton, Robert. 2006. *Words and Thoughts* Oxford, Oxford University Press.

Tomasello, M. 2008. *The Origins of Human Communication*. Cambridge Mass: MIT Press.

Wilson, David Sloan. 2002 *Darwin's Cathedral: Evolution, Religion and the Nature of Society*. Chicago: University of Chicago Press.