Language as Mechanisms for Interaction*

Ruth Kempson†*, Ronnie Cann**, Eleni Gregoromichelaki****, and Stergios Chatzikyriakidis***

*King's College, London
**University of Edinburgh
***University of Gothenburg; Open University of Cyprus
****University of Osnabrück and King's College London

April 21, 2016

*There are many who deserve our thanks for ongoing work on the *Dynamic Syntax* framework; but our particular, fervent thanks go to those involved in work done on *The Dynamics of Conversational Dialogue ESRC-RES-062-23-0962*, on which the results reported here are grounded: Arash Eshghi, Patrick Healey, Christine Howes, Julian Hough, Wilfried Meyer-Viol, Gregory Mills, Matthew Purver, and Graham White. The last author gratefully acknowledges support from the Vetenskapsrådet, project 2014-39, as well as the ANR ContInt Polymnie project in France (ANR-12-CORD-0004).

†Corresponding author. Mailing address: correspondence and transmission of proofs: 67 Hendon Wood Lane, London NW7 4HT, UK

E-mail: ruth.kempson@kcl.ac.uk; r.cann@ed.ac.uk; elenigregor@gmail.com; stergios.chatzikyriakidis@gu.se

Abstract

Language use is full of subsentential shifts of context, a phenomenon dramatically illustrated in conversation where non-sentential utterances displaying seamless shifts between speaker/hearer roles appear regularly. The hurdle this poses for standard assumptions is that every local linguistic dependency can be distributed across speakers, with the content of what they are saying and the significance of each conversational move emerging incrementally. Accordingly, we argue that the modelling of a psychologically-realistic grammar necessitates recasting the notion of natural language in terms of our ability for interaction with others and the environment, abandoning the competence-performance dichotomy as standardly envisaged. We sketch *Dynamic Syntax*, a model in which underspecification and incremental time-relative update is central, showing how interactive effects of conversation follow directly. Finally, we note the changing cognitive-science horizons to be explored once a language-as-action view is adopted.

1 Introduction

In this paper, we start from some simple observations about how language is used in every day conversation to facilitate human interaction. We argue from these that, contrary to standard foundational assumptions of theoretical linguistics, the most plausible accounts of linguistic use show that the concept of language itself needs to reflect the processes of action-coordination that make human interaction possible.

Firstly, we show the challenges that the incrementality of producing/comprehending language poses for current assumptions about formal language modelling, especially in view of the fact that such incrementality crucially underpins the interactivity displayed in conversation. On this basis, we argue against the standard methodology of isolating classes of phenomena as independent of language use and analysable through declarative propositional knowledge ('competence') invoked during online processing. Such a methodology, although a legitimate initial methodological simplification, eventually appears inadequate especially when faced with data that, even though arguably constituting an integral part of any such theory's empirical basis, nevertheless cannot be subsumed under the idealisations that have been imposed. We then argue that the only way to get to grips with these challenges is to recast our concept of language in terms of ability for action within the social and physical surroundings of language use, making the temporal flow of utterance processing central to the explanation of structural properties of language. We set out a formal model which adopts this dynamic stance, recasting the concepts of syntactic and morphosyntactic constraints in terms of goal-directed actions that enable predictive and incremental linguistic processing.

The explanatory basis of the account to be provided is based on two basic ingredients for sustaining an incremental account in the face of various grammatical dependencies and the vagaries of context-shifting: integrating a notion of *underspecification* and a correlated notion of subsequent *update* across all grammatical mechanisms, underpinning, for example, local/distant discontinuities, as well as anaphora and ellipsis. From such a perspective, instead of underspecification being seen as a defect to be remedied through pragmatic inference, it is taken here as a basic ingredient of cognitive processing and a defining characteristic of natural language grammars which enables them to serve as coordinative devices in dialogue.

2 Split utterances in dialogue: the challenge of incrementality

The first thing that strikes one on looking at informal conversations is how crucially we depend on context for what we say and understand. Despite the huge amount of work on modelling semantic/pragmatic concepts of context (Stalnaker 2014, Kamp 1981, Kamp and Reyle 1993, Bittner 2001, 2007, Asher and Lascarides 2003 among many others), there is reason to think that the context-dependence of natural languages is systemic in ways that have not yet received due recognition, affecting not merely semantic dependencies, but syntactic and morphosyntactic ones also (see e.g. Kempson et al 2001, 2007, Bittner 2014ab). The phenomenon that we focus on here is that, in conversation, commonly, we do not speak in complete sentences. Instead, we rely on the surrounding context, whether linguistic or non-linguistic, provided by ourselves or others, to coordinate our verbal and non-verbal actions. As a result, what can be considered as "complete sentences" may emerge through a sequence of non-sentential contributions, with each interlocutor adding some fragment to a partial structure. However, switching of roles between speaking and listening is unrelated to standard syntactic boundaries – at any time, people can join in on what the other person is setting out to say, often helping the other to find the appropriate add-on (1), but sometimes simply adding to it as in (2):

(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]

Each such contribution can add unproblematically to whatever partial structure has been set out so far, irrespective of whether or not what precedes it is a full constituent or sentence.

The freedom is essentially unrestricted. In informal dialogue, people can take over from each other at any point in a clausal sequence: (1) involves a split between determiner and noun; (2) a split between preposition and its complement, and (3) a split between a post-auxiliary verb subject and main verb.

(3) A: Did Jo...
B: stumble? I hope not.

(4) splits apart a long-distance dependency between the left-peripheral *wh*-form and its associate "gap" position as object of *axe*, and also between an auxiliary and its dependent bare stem verb-form:

(4) A: Which unit are we thinking we should . . .
B: axe? None.

There is also no natural limit to the complexity of such interruptable structure. Split dependencies can be arbitrarily complex, displaying multiple, interleaving semantic dependencies, with all the characteristics of locality. For example, in (5), the negative polarity *any of the saliva kits* displays a dependency on A's initiating utterance across B's intervention, while, nevertheless, the containing fragment in which it occurs is construed as an extension of B's intervention. Additionally, the quantifier introduced by the A's utterance binds a pronoun in B's utterance, and the function of B's utterance as a request for confirmation cannot be achieved unless local binding is assumed:

(5) A: Has every female gymnast handed in
B: her blood sample?
A: or even any of the saliva kits?

These data pose very considerable challenges for standard linguistic models. Under standard assumptions, the grammar presumed to underpin natural language is taken to constitute a model of the individual speaker's linguistic capacity, with the concept of 'sentencehood' providing the remit within which grammatical explanations are formulated. Accordingly, syntax is defined as licensing well-formed sentence-sized units; semantics is based on evaluating content for such syntactic structures to yield full propositions relative to invariable contextual parameters which determine content globally (e.g. speaker/hearer roles, salient discourse referents etc.); and these semantic characterisations then feed into some extragrammatical account of pragmatic inference operating under the assumption that propositional intention recognition will resolve whatever underspecification remains. However, the split utterance data contradict all such assumptions.

Firstly, such data demonstrate that participants can manipulate structures effectively and strategically at a level considerably lower than that of sentences or propositions. Secondly, as (6) shows, the significance of the uttered sentence, instead of itself determining "enrichment" processes, may only become apparent as the conversation progresses, with no one contributor having any idea of the overall structure from the outset. For example, (6) below only turns out to be a conditional very late on in its development, and, even then, the conditional is over a non-sentential unit (*with the dogs*), not over the full sentence (going to see Hugh's aunt is not conditional on Eliot controlling the dogs):

(6) Alex: We're going to . . .
Hugh: Burbage, where Auntie Ann lives
Eliot: with the dogs?
Hugh: if you look after them.

To add to this problem, continuations or interruptions are by no means necessarily intended as "cooperative" contributions to the expression of the propositional intention of the other interlocutor:

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

Therefore the processing of such partial strings is not necessarily underpinned by the prior recognition of a propositional intention providing the missing form/content, which is then deflected by the continuation. Moreover, such interruptions, even though displaying strong syntactic/semantic "connectivity" effects, do not reflect contextual continuity in terms of the parameters that resolve the "characters" (Kaplan 1989) of the context-dependent expressions employed, whether these resolve participant roles or the worlds at which the content of the utterance needs to be evaluated:

(8) A: Every female athlete approached me and said that
B: you deceived her.

Hence evaluation and contextual resetting needs to be much more fine-grained than its usual spanning of whole clauses (Gregoromichelaki, to appear, cf Bittner 2014a).

The emergent strings may, even, never add up to a sentence at all, even though the dialogue itself is perfectly coherent and successful (as noted by Stainton 2006):

(9) A: Covent Garden?
B: Right at the lights. Then straight on up.

(10) Fire!/Scalpel/Sutures/... [Yanofsky 1978]

Such non-sentential fragments clearly show that people are able to perform speech acts just by employing words and phrases (Barton 1990, Elugardo & Stainton (2005)). It is also obvious that people are able to produce and understand words and complex phrases in isolation, for example, in titles, lists etc. without either any proposition needing to be generated or any speech act recognised to be being performed:

(11) "Frankie Goes To Hollywood" (band name) [from Barber 2005]

(12) "Sainsbury's"/"Jimmy's" (supermarket/restaurant names)

Such cases have been addressed previously under standard (generative grammar) assumptions in terms of the grammar[1] with appeal to pragmatics to explain further cases like (9), where supposedly whole propositions need to be derived in order to justify the performance of speech acts (Stainton 2006). However, such accounts, even though dispensing with the sentential restriction for grammatical licensing, still insist on some kind of standard constituency requirement for the licensing of such non-sententials. This, though, immediately prevents various types of data from being licensed. There are many types of non-constituent "fragmentary" utterances occurring in discourse:

(13) A (showing a jug): Purchased in Germany.

(14) A: seen Tom? [Napoli, 1982]

(15) A: Are you reading? B: No, writing a linguistics paper. [Thrasher, 1974, cited in Napoli 1982]

[1] See Morgan, 1973, 1989; Barton 1990, Barton and Progovac 2005.

It might then seem that these phenomena, along with the split-utterance data in (1)-(8), can be resolved by subsuming them into the category of non-standard constituency effects, of which more familiar phenomena within the literature are coordination or parenthetical insertion. Nonconstituent coordination/parentheticalisation data have been partially addressed within current grammar formalisms, specifically by categorial grammar, e.g. Combinatory Categorial Grammar (CCG: Steedman, 1996, 2000):

(16) Mary, if she still wants to, will come to Burbage, too.

(17) Mary will buy, if she has enough money, that nice painting at the exhibition.

However, the requisite conception of non-constituency licensing is far more general than the CCG account allows for, as this arises both in monologue and dialogue. For example, the kinds of disruption displayed in (2)-(8) can also appear within a single sentence uttered by just one person, for example, parentheticals disrupting the incremental flow of the primary structure but nevertheless contributing to the compositional truth conditions of the main clause despite the violation of traditional constituency (Gregoromichelaki, 2006, to appear):

(18) Mary will come to Burbage for sure, and with the dogs, if you look after them.

(19) Mary and, I think, John went to the exhibition.

(20) Well, they dropped cards in *I suppose* the doors (ICE-GB[2]: s1a-020 #177; from Dehé 2014)

(21) I mean in in that piece we've just heard from The Revenger's Tragedy it's a mixture *isn't it* of original instruments and kind of what sound to me like modern trumpets (ICE-GB: s1b-023 #140, from Dehé 2014)

Such phenomena, along with the split-utterance data in (1)-(8), cannot be tackled by the usual CCG mechanisms unless those are generalised to a degree that undermines both the formalism's philosophical and psychological 'competence' grounding and declared desiderata in terms of complexity results.

In general, for any standard grammar, the problem is that natural linguistic behaviour undermines basic theoretical notions like the abstract and folk-linguistic concepts of "sentence" and any assumptions about string-level constituency. Given that the distributed effects of these split utterances are uncontentiously wellformed, they need to be licensed by the grammar. So standard grammars risk failing a minimal criterion of adequacy: that of licensing all and only the well-formed occurrences of the phenomenon displayed across the split. Furthermore, as regards pragmatic models, Gregoromichelaki (2013) and Gregoromichelaki and Kempson (2015) argue on the basis of such data that there is no restriction that whole propositions are ever required for the performance of speech acts (contra Stainton (2006) among many others): various distinct conversational acts can be performed incrementally, with subsentential/subpropositional constructs (see e.g. (6), (9)); it is even the case that exactly the incompleteness of initiating some grammatical dependency can be recruited for speech act performance (see e.g. (2) earlier and (32)-(36) below). In such cases it is transparent that analysing such shared strings as independent utterances misses not only the rhetorical significance of such uses but also the role of the grammar, once appropriately defined, as one of the means of achieving, not only the formulation of individual thoughts, but also coordinated action among people.

3 The competence-performance perspective: invoking ellipsis

Given that current formal linguistic theories take competence-performance (Chomsky, 1995), modularity (Fodor 1983), and Marrian computational vs algorithmic level distinctions (Marr 1982, Steedman, 2000,

[2]ICE-GB: International Corpus of English, British Component

Kobele, 2012b) as fundamental,[3] it is worth examining whether standard models can be enhanced with tools that handle the dialogue data while maintaining standard foundational assumptions. In such accounts, the grammar deals only with the licensing of 'expressions' (structured string - meaning pairs) in order to satisfy the required bottom-up compositionality of content for sentence-strings, with this level being unrelated to the order of words.[4] Given such assumptions, dialogue data need to be subsumed under existing accounts of ellipsis, whether syntactic, semantic, or pragmatic. We examine these in turn below.

3.1 Syntax and the structure of silence

In accounts of ellipsis in the generative tradition, elliptical strings are taken to be sentences with full clausal structure but with all phonological information except that of the fragment having to be deleted (Sag 1976) or, alternatively, by involving copying of syntactic structure (Fiengo and May 1994, and many others following). This is the strategy followed by, for example, Merchant (2005) who analyses NP fragments that occur as answers to questions as involving an underlying full sentential structure which is deleted after "movement" of the pronounced fragment *him* to the left periphery:

(22) A: Who did you see in Burbage?
B: Him (, in the afternoon)

Additionally, "fragments" like (9)-(10) are also similarly analysed as elliptical under movement and deletion following some type of discourse "accommodation" by Merchant (2005), van Craenenbroeck (2012), Weir (2014) among others.

From an empirical point of view, in such analyses, the appeal to 'ellipsis' is supported by syntactic 'connectivity effects', i.e., the licensing of dependencies between the fragment and its assumed elided structural linguistic context (e.g. in the licensing of accusative in B's reply).[5] However, because the only architectural arrangement allowed by the theory is the mapping [sentence $\rightarrow$ proposition $\rightarrow$ speech act], the only assumption that can be made to account for the functions and syntactic restrictions of such "fragments" is that, despite appearances, they are whole sentences underlyingly. But such an account does not yield empirically accurate results, for example, the perfectly natural addition of the phrase *in the afternoon* in (22), which will yield a non-constituent unable to be accounted for as "moved" before deletion.[6] More importantly, the standard assumption that only constituents can be deleted makes such an approach completely non viable for the split-utterance data in (1)-(2), (4), (5).

Moreover, from a theoretical point of view, such analyses (e.g. Merchant 2005, Kobele 2015) do not characterise the general phenomenon of *ellipsis* in an adequately generalisable manner, namely as contextual dependency (*underspecification*), but, instead, as *ambiguity* that needs to be resolved on a case-by-case basis. As a result, such accounts, besides the invocation of 'movement', also involve the postulation of multiple underlying structures/derivations/types of strings in any case of ellipsis. This then even extends to cases where ambiguity needs to be invoked not only at the ellipsis site but also for the antecedent, even though that antecedent is not ambiguous. For example, (23) allows interpretations involving Bill's washing his own socks (the "sloppy" interpretation) as well as Bill washing John's sock (the "strict" interpetation), even though *John washed his socks* is not ambiguous in this particular context with *his* interpreted as 'John's':

(23) A: John washed his socks
B: (And) so did Bill / Bill too.

[3] Various aspects of these distinctions are currently being disputed see e.g. Stokhof and van Lambalgen (2011), Piccinini & Craver (2011), Milkowski (2013), Chater et al (2015), Christiansen and Chater (2015), Lewis and Phillips (2015), Stabler (2013).

[4] However, this has not been universally the case , see Hausser (1989), Milward (1994), Hawkins (1994), Phillips (1996), Poesio & Rieser (2010), Bittner (2014)a,b.

[5] But see Barton and Progovac (2005) for relevant counterevidence regarding case in English.

[6] For further empirical counterarguments, see e.g. Stainton (2006), Ginzburg & Sag (2000) Ginzburg (2012).

Attempting to account for the phenomenon of split utterances in these terms would compound this problem since the split can occur at any point in a string. So analysing all such "fragments" as incomplete sentences in their own right would make ALL (even non-constituent) sequences by definition sentence-sized objects, jeopardising the possibility of any viable, internally-coherent concept of 'convergent' derivation. The usual ("non-pedestrian", Stabler 1991) alternative in such cases is to invoke the operation of grammar-external mechanisms to explain the phenomena, e.g. to provide the missing grammatically-licensed antecedents. In current accounts of this kind (e.g. Kobele 2012a,b, Kobele 2015), an incremental parser can provide already derived parsing steps as antecedents for ellipsis sites that are syntactically constrained to be of various particular syntactic types. However, despite its proliferation of syntactic constructions, and the appeal to another level of "constituency" analysis, even such a concession would not provide an adequate account in terms of covering all the pretheoretically acceptable data. This is because, in many cases, there is no basis for analysing such partial strings as surface-incomplete sentences with content that can be provided through the previous linguistic context. In (24), both speaker and hearer are attempting to fully integrate semantically and pragmatically the first subject NP *the doctor*, efficiently interacting with each other in order to establish its interpretation by utilising another non-sentential element (Kempson et al, 2007; Gargett et al 2008, Gregoromichelaki et al 2011), and only then proceeding with the rest of the proposition:

(24) A: Er, the doctor
B: Chorlton?
A: Chorlton, mhm, he examined me, erm, he, he said now they were on about a sli[ght] [unclear] on my heart. [BNC]

This shows that people are able to manipulate sub-sentential elements, and strive to integrate their content without waiting for a full sentential/propositional licensing to become apparent. Appeal to unorthodox "constituents" ('contexts', Kobele 2012a,b) generated by an independent parser/generator does not solve the problem because the general problem of ellipsis is a problem of grammar-internal context-dependency (i.e. underspecification) not syntactic ambiguity. This extends to categorial grammar accounts which are in principle able to license some word-by-word incremental accumulation of content within the grammar formalism, but only by defining mechanisms such as type-lifting, which multiplies syntactic and semantic types. Such multiplication is supposed to be tightly restricted and predefined within the formalism so that no undesirable overgeneration and complexity can occur. But this leads to a further empirical problem. In many split utterance cases, as well as many of the parenthetical cases in (16)-(21), the break does not occur within those well-defined limits. Moreover, there is no indication that, either in monologue or dialogue, the continuation to some initial sequence must have been planned well in advance at the sentential/propositional level so that appropriate globally licensing derivations can be invoked by the parser/generator.

3.2 Incomplete sentences - complete thoughts: the semantic account

It might seem that an account of split utterances exclusively at the level of semantics has a better chance of success, using tools developed over decades for modelling ellipsis at the propositional level. Indeed, one such semantic account of ellipsis, the Dalrymple et al (1991) higher-order unification account, was explicitly formulated to avoid the stipulatory ambiguity of syntactic accounts of VP-ellipsis indicated earlier in section 3.1. For this purpose, a form of higher-order abstraction was defined over some preceding propositional content to yield an appropriate predicate to apply to the content of the fragment. On this style of analysis, the non-sentential expressions which constitute the fragment are taken to be propositional, with the mismatch between syntactic and semantic categorisation driving the interpretation process. However, straightforward application of such a procedure would at best not cover all instances of split-utterance construal, as witness again (24). Applying an abstraction mechanism there would inappropriately provide an abstract defined over the previous proposition, when what is required is a clarification of the immediate context given solely by the expression *the doctor*.

There is however a more general problem facing all accounts of ellipsis based on different variants of abstract-construction. These accounts are invariably defined relative to a sentence $\rightarrow$ proposition mapping concept of grammar. Accordingly, the input to any context-relative evaluation is that of a sentence-string paired with an articulated propositional structure, often with co-indexing determined by mechanisms underpinning anaphoric dependence within the syntax and so articulated independently of any semantic account (as noted by Bittner 2007). Whatever notion of context-relativity the semantic account might then be able to express will inevitably involve incrementality of update only on a sentence-by-sentence basis, because these are the inputs provided to the semantics for each such context-relative evaluation. To invoke the parser/generator to deliver incremental semantic derivations at this juncture (Stabler 1991) violates the 'strict competence hypothesis' (Steedman 1992) which, in our view, removes any motivation for maintaining the competence-performance distinction. It is notable in this connection that Discourse Representation Theory (DRT), despite explicit commitment to characterising how gradual construction of Discourse Representation Structures provides an interface for defining progressive semantic update, involves sentence-by-sentence update. All updates are truth-conditionally evaluable after all embeddings have been processed (Kamp and Reyle 1993; for cogent radically-contextualist moves in a different direction, see Jaszczolt et al (2015) and, especially, Bittner 2014b).

In any case, there remains reason to doubt the viability of any purely semantic characterisation of non-sentential utterances in conversation. This is because, as pointed out by syntactic accounts of ellipsis, various connectivity effects occur so that exorcising such data by appealing to semantics, pragmatics or the parser would not provide an adequate, unified account. For example, in morphologically-rich languages, non-sentential expressions have to adopt a form appropriate for what would be expected in an overt clausal sequence. Thus in (25), the form of the pronoun has to be the nominative form *ego*, not the accusative *emena*:

(25)	*A:*	*I*	*Maria*	*to*	*egrapse*	*to*	*grama?*
		the-NOM	Maria-NOM	it-ACC	write-PAST.3SG	the	letter-ACC

A: 'Did Maria write the letter?'

B:	*Oxi,*	*ego*	/	*#emena.*
	No	I-NOM	/	#me-ACC

B: 'No, I did / #Me did'

The challenge of defining a model that is appropriate for accounting for such phenomena is that the model must crucially define procedures for such syntax-context interaction within the grammar.

3.3 Combining modular resources to face the dialogue challenge

Such interactions provide the motivation for the dialogue model proposed by Jonathan Ginzburg and colleagues (Ginzburg & Cooper 2004; Fernandez et al 2007; Purver & Ginzburg 2004; Ginzburg 2012). This model is currently based on a constructional version of HPSG expressed in the TTR representational framework (HPSG$_{TTR}$, Ginzburg 2012). Consequently, it is able to capture syntactic, semantic and pragmatic aspects of conversational utterances including speech-act functionality, all within a single formalism. Non-sentential utterances are modelled by defining a mapping from a sui-generis syntactic derivation for some sub-sentential constituent (e.g. mapping directly from NP to S) onto the range of interpretations open to it, including fine-grained speech-act identifications. So this account acknowledges the problem of ambiguity that confronts syntactic approaches to ellipsis and attempts to restrict it by specifying *constructions*, e.g.

various types of non-sentential *clarification request*s, encoding idiosyncratic syntactic-semantic-pragmatic interactions.

It is an empirical issue whether structures and their attendant contextual interpretations can be limited in such respects. However, due exactly to the fact that the HPSG$_{TTR}$ licensing rules are defined across such syntactic-semantic-pragmatic matchings, the account loses the flexibility and generality required to deal with open-ended phenomena like the incremental licensing of non-standard constituency and speech act effects seen in (1)-(8). Relatedly, the same problem arises for fragments of the types seen in (9), (13)-(15), because the morphosyntactic constraints that license the already defined HPSG$_{TTR}$ 'constructions' obtain identically also in cases where there is no overt linguistic antecedent. For example, in morphologically-rich languages, non-sentential expressions are required to take a definitive morphological form whether or not there is some overt antecedent in a previous clause (contra Stainton 2006):

(26) [Context: A and B enter a room and see a woman lying on the floor:]
A to B: Schnell, den Arzt /#der Arzt
Quick, the doctor-ACC /#the doctor-NOM
[German, Gregoromichelaki 2012]

(27) [Context: A is contemplating the space under the mirror while re-arranging the furniture and B brings her a chair]
A to B: tin karekla tis mamas? / #i karekla tis mamas?
the-ACC chair-ACC of mum's? / #the-NOM chair-NOM of mum's?
(Ise treli?) (Are you crazy?)
[clarification, Modern Greek, Gregoromichelaki 2012]

Such data will not be covered by the already defined HPSG$_{TTR}$ constructions because the specifications of such constructions in the grammar rely on linguistic expressions being present. So even an account which brings together syntactic and semantic considerations to devise a fully comprehensive account of dialogue does not seem able to address the full range of licensing non-sentential elements in language use. This is because of the competence-performance distinction which confines 'syntax' to the licensing of linguistic strings and disallows parsing/generation features within the grammar (however, see Ginzburg et al (2012) where the first steps towards incrementality are taken within this model).

The only alternative would then appear to be to posit a purely pragmatic account (e.g. Stainton 2006) at least for such non-sentential fragments. However, the obligatory morphosyntactic features required to license the presence of fragments such as (25)-(27), as well as the fact that such an account relies on necessarily deriving propositional structures via inferential mechanisms, precludes such an account from encompassing the full range of data (see also Ginzburg 2003, Gregoromichelaki et al 2011, Asher & Lascarides 2013 for arguments against Gricean inference necessarily underpinning linguistic processing).

3.4 Incremental sharing of sentences and intentions

A welcome alternative view that seems to undermine the competence-performance assumption by modelling the incrementality of processing and the full interaction of the grammar with pragmatic notions is that of Poesio & Rieser (2010). This account indeed extends naturally to (part of) the split-utterance data. Poesio & Rieser offer an in-depth account of one type of split utterance, so-called *collaborative completions* as in (1), in which the second speaker is trying to contribute to some task in collaboration with the interlocutor. In such cases, the second speaker can be modelled as being able to intervene with a follow-up fragment completing the first speaker's utterance, because such a completion can be derived through inference from mutual knowledge/common ground. Poesio & Rieser's account of such phenomena combines an incremental grammar formalism based on LTAG (Lexicalized Tree Adjoining Grammar) with a model of coordination in dialogue relying on Gricean-style rational reasoning to achieve joint-intention establishment and recognition. Such reasoning is necessarily proposition-based at the pragmatic level; and, at the level of the grammar, the standard view of a syntactic string-based level of analysis is maintained for it is this which provides the top-down predictive element allowing the incremental integration of continuations. However, exactly this assumption impedes a more general account of split utterances, since there are cases where the contribution made by the second speaker cannot be seen as an extension to the string of words/sentence offered by the first speaker:

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

(29) Eleni: Is this yours or
Yo: Yours. [natural data, from Gregoromichelaki et al, 2011]

In (29), the string of words (sentence) that the completion yields is not at all what either participant takes themselves to have constructed, collaboratively or otherwise. Similarly, in (28), even though the grammar is responsible for the dependency that licenses the reflexive anaphor *myself*, the explanation for A's continuation in the third turn of (28) cannot be string-based as then *myself* would not be locally bound (its antecedent is *you*). Any account that relies on sentence-proposition mappings as the basis of syntax (e.g. Kobele 2015) cannot deal with such structures. Moreover, in LTAG, parsing relies on the presence of a head that provides the skeleton of the structure. Yet, as (4), (24) indicate, utterance take-over can take place without the appearance yet of the head that determines argument dependencies (see also Gregoromichelaki et al 2011; Purver et al 2010; Howes et al 2011).

There are further considerations threatening the explanatory generality of intention-based accounts like Poesio & Rieser's. Despite its incremental syntax, this account (like Stainton, 2006; Barton, 1990) relies on the generation and recognition of the speaker's propositional intentions as the basis for licensing the continuation. Yet though this stance is well suited to the specifically collaborative tasks modelled by Poesio and Rieser, in informal conversation, as we saw earlier in (24), such fragments can occur well before the informative intention, which is standardly defined as requiring a propositional object, has been made manifest. Moreover, unlike what happens in Poesio and Rieser's task-oriented dialogues, in everyday conversation, many fragments do not involve straightforward participant cooperation or inference as to the speaker's intended utterance, hence Gricean notions of reasoning are not applicable (see Gregoromichelaki et al, 2011). For example, in (30), the son not only isn't being cooperative in the standard Neo-Gricean sense but also does not feel the need to wait for the formulation of the mother's intentions despite her second explicitly introductory *then*:

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

The fact that recovery of propositional intentions is not integral to communicative success is also shown in cases such as (31) where various speech acts are accomplished within the unfolding of a shared single proposition (Gregoromichelaki et al 2013):

(31) Hester Collyer: It's for me.
Mrs Elton the landlady: And Mr. Page?
Hester Collyer: is not my husband. But I would rather you continue to think of me as Mrs. Page.
[from The Deep Blue Sea (film)][7]

[7] Along with natural data, constructed data from literature, film scripts etc. are particularly relevant in this context as they show that such constructions are not "speech errors" that can be easily dispensed with.

The ability of such shared constructions to sustain multiple subsententially-performed speech-acts indicates that imposing the explicit derivation and representation of propositional speech-act intentions is too strong a requirement on language processing. In fact, coordination relies on the flexibility/underspecification of the grammatical devices employed: necessarily (meta)representing descriptions of the speech acts performed would impede joint-action, rather than facilitate it. For example, in psychotherapy sessions, invited completions have been argued to (consciously or unconsciously) exploit the indeterminacy of the speech act involved to avoid overt/intrusive elicitation of information (Ferrara 1992):

(32) Ralph (therapist): Your sponsor before ...
Lana (client): was a woman.
Ralph: Yeah.
Lana: But I only called her every three months.
Ralph: And your so your sobriety now, in AA ::[(is)]
Lana: [is] at a year. [Ferrara 1992]

The fact that explicit metapragmatic representations of the acts performed can be invoked in cases of resolving trouble in conversation does not imply that such explicit descriptions of the speech acts performed are always present and guiding dialogue. Instead it seems that perfectly intelligible moves in conversation can be achieved simply by initiating a grammatical dependency which prompts either interlocutor to fulfill it without specific determination or identifiability of a given speech-act.

In sum, the seamless fluency with which individuals take on or hand over utterance responsibility subpropositionally presents a formidable challenge to both grammar formalisms and pragmatic models.

4 Rejecting the competence-performance dichotomy: an alternative grammar architecture

To tackle the dilemmas faced by static, modular accounts of linguistic knowledge when applied to dialogue data we need a radical shift of perspective. We need to turn to the alternative challenge of articulating a psycholinguistically-motivated incremental grammar, conceptualised as knowledge-how that people apply to various domains of processing. Under such a view, all grammatical dependencies, which, as we saw, are able to function as coordinating devices (2), (31), (32), could be characterised as generating predictive goals to be fulfilled by either interlocutor in the very next steps which they will be taking. This explains the interlocutors' ability to either fulfill such predictions through non-linguistic contextual provisions or by offering their chosen linguistic contribution. However, this presupposes that not only does the grammar incorporate processing features like incrementality and predictivity but also that it provides a shared "workspace" (Kempen et al 2012, Kempen 2014) for both production and comprehension to operate and interact (Gregoromichelaki et al 2013a,b). Minimal acts of coordination (i.e. "speech acts") might then be achieved incrementally at each step without requiring the employment of full sentences/propositions or inferential reasoning, just by employing the potential of grammatical lexical and computational resources. For example, completions might be explicitly invited by the speaker to form what externally can be described as a question-answer pair as in (33)-(35):

(33) A: And you're leaving at . . .
B: 3.00 .

(34) A: And they ignored the conspirators who were . . .
B: Geoff Hoon and Patricia Hewitt. [BBC Radio 4, Today programme, 06/01/10]

(35) George: Cos they [unclear] they used to come in here for water and bunkers you see.
Anon 1: Water and ...
George: Bunkers, coal, they all coal furnace you see, [BNC]

The functional role of each such contribution does not have to be modelled as an explicit descriptive metarepresentation of what a participant does with their utterance. A psycholinguistically-motivated model includes in its definition of ‘context’ not only the current grammatical goals, incrementally shifting contextual parameters, and discourse referents established, but also the various previously projected processing paths that unfold after each word-utterance (Eshghi et al 2015, Gregoromichelaki to appear). Within such a context, even elaborate dialogue actions can be modelled implicitly and mechanistically: various phenomena externally described as e.g. clarifications, disagreements or (self-)corrections simply involve retracing processing steps and opening up again various paths that had been projected up to completion due to the normal structural/semantic/pragmatic properties of the usual processing actions (Kempson et al 2007, Gargett et al 2008, Hough 2015, Eshghi and Lemon, 2014, Eshghi et al 2015, Gregoromichelaki to appear). Additionally, with this context being token-identical for both parsing/generation, split-utterances are predicted to be feasible at any point to subserve the traversal of such paths in functionally-relevant ways.

The ability to mechanistically manipulate such a context thus does not presuppose any ‘rational’ high-order inference, standardly taken to be the basis of all successful human communication (Sperber & Wilson 1986, Clark 1996, Carston 2002 and many others). Accordingly, it is predicted that such split-utterances are also employed in interactive exchanges with very young children, and are indeed one means for detecting complex conceptual abilities like negation, suggesting, on the one hand, that acquisition need make no reference to higher order reasoning, and, on the other, that recognition of the content of other people’s intentions is not a necessary condition for acts of communication to be successful (Gregoromichelaki et al 2011):

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

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

In such cases both adult and child alike rely on the predictive potential afforded by subpersonal grammatical mechanisms, without the mediation of complex metarepresentational reasoning about intentions. These mechanisms, e.g. the fulfillment of a syntactic dependency, prompt the child to retrieve and verbalise appropriate continuations, e.g. say their name, so that a joint communicative event is achieved (Christiansen and Chater 2015).

The alternative architecture that we believe resolves the difficulties presented by the dialogue data is embodied in the theory of Dynamic Syntax to which we turn next.

5 Dynamic Syntax

Dynamic Syntax (DS, Kempson et al 2001; Cann et al 2005a; Kempson et al 2015, Eshghi et al 2011) is a grammar formalism of which the core notion is incremental interpretation/linearisation of word-sequences/contents relative to context. Production and comprehension are modelled symmetrically as operating on the same representations via the mirroring of parser/generator actions; and, consequently, the split utterance data emerge as core data directly predicted from the system itself.

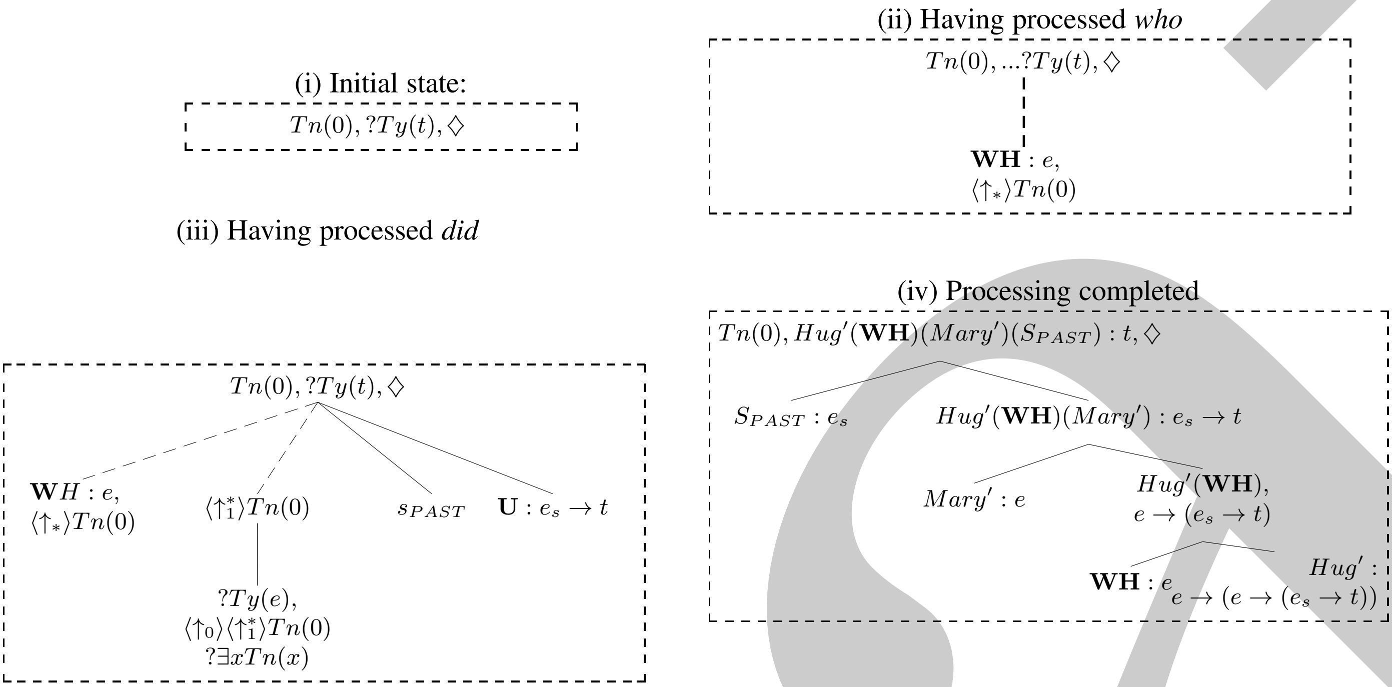
What has gone in this approach to language modelling is the traditional concept of *syntax*, defined as the structural properties of sentences of the language independent of either context or the dynamics of their usability. In its place are the twin concepts of underspecification and update of conceptual representations and strings; and it is the burden of this paper to argue that not only does this match predictive successes of traditional syntax but also it shows up how the complexity of much of current attempts to model more recalcitrant data is an artifact of precluding any such dynamics within the grammar. The DS formal devices are defined to model how to pair emergent conceptualisations of eventualities with words being uttered in sequence. So the system is articulated in terms of goal-driven *actions* that give rise to expectations of further actions, all involving incremental updates towards constantly generated new goals. On this view, the setting out of such dynamics constitutes the grammar: words, syntax, and morphology are all triggers of pre-specified sets of actions (*macros*) that induce such conceptualisations or associate word sequences with them. *Context* becomes an integral part of grammar, equally dynamic and evolving. It is not only a record of the emergent (partial) structures and cross-modal contributions, but also a record of the sequences of actions used to construct representations of content as performed so far.

To model this action-directed perspective, DS is founded on a dynamic modal logic that defines the transitions among states taken to constitute the current context of processing at each point (see Kempson et al 2001 for formal details). The accessibility relations among these states are defined through *actions* which license goal-driven, incremental transitions from state to state. Such states can be taken to model the total context of each processing step, linguistic and non-linguistic, so that the whole system licenses mappings from context to context. The DS system is defined as a set of constraints expressed in a formal language whose models are such context states (Kempson et al 2001, chapter 9). As such, DS's approach to grammar is *model-theoretic* (unlike categorial grammars and minimalism, see Pullum & Scholz 2001). This approach does not recursively enumerate well-formed expressions, instead it allows the licensing of contexts as models that satisfy the constraints imposed by the system. Among the various advantages conferred by such an approach is the potential to express notions of gradient grammaticality, the licensing of "partial" (i.e., in DS terms, extendible) constructs, and the modelling of "lexical flux" (Pullum & Scholz 2001, Gregoromichelaki to appear). For simplicity of illustration purposes, we display partially below the goal-driven development of models, i.e., conceptual representations of content illustrated as partial trees.

5.1 Sketching the dynamics

The general pattern is sketched in (38) for an idealised null-context parse of the string *Who did Mary hug?*.

(38) Processing *Who did Mary hug?*



As (38) displays in (iv), a complete tree resulting from some successful parse is a binary-branching structure consisting of nodes inhabited by representations of content expressed in the lambda calculus augmented with the epsilon calculus of Hilbert & Bernays (1939).[8] The progression onto that content in association with the linear production of a string is defined across sequences of partial trees, (i)-(iii), which are partial across parameters both of content and structure, such partial trees being annotated not only with typed conceptual representations as they become available, but also with other annotations that drive or constrain the process of construction. The input tree of the idealised case (38) is (i), a single-node tree whose only annotation gives a skeletal indication of the goal to be achieved: a requirement for a tree of propositional type, represented as $?Ty(t)$.[9] (ii) is the update achieved in having processed *who*, (iii) that achieved by processing *did*, and, collapsing a number of steps, (iv) is the final state, having processed *hug*. Every tree contains a *pointer*, $\Diamond$, which at each interim stage indicates which node is to be the next one under development: in the sequence in (38), this shows how development can start from some initiated rootnode and then return to it at various subsequent stages, as in the second and third trees in (38), and finally return to that node when all subgoals that have been introduced get subsequently satisfied.

Trees are by definition binary, being representations of functor-argument structure, with the convention that functor nodes appear on the right and argument ones on the left.[10] As such, trees do not represent word order or syntactic constituency, the former only being recoverable from the trace of transitional states that lead to the current tree, the latter being considered epiphenomenal. On each node, if there is a content formula, it is written to the left of the colon ":" with a type specification to its right, e.g. $Sue' : e$ is the annotation for a formula Sue' of entity type e. DS uses only a restricted set of types: e for individual entity, e_s for event or situation entity, $e_{s \rightarrow t}$ for predicates over events, $e \rightarrow (e_s \rightarrow t)$ for (one-place) predicates over individuals, and so on. Individual terms, whether quantified or not, are invariably terms of type e as defined

[8]The epsilon calculus constitutes the formal account of arbitrary names as used in Predicate Logic natural-deduction proof systems. There has been considerable work on the epsilon calculus since that time (see e.g. Meyer-Viol 1995), but in not addressing quantification in this paper, we will leave all related issues aside.

[9]In fact, any type-requirement can serve as the initiating context, and, since this is an incremental framework, the initial input which the processor develops will commonly be some tree already partially developed in the immediate context.

[10]The one-dimensional page display means that (iii) is misleading in suggesting that trees may have more or indeed less than a pair of sister nodes. This is because, on the one hand, lexical actions can induce any subpart of such binary trees. On the other hand, unfixed nodes are relations in search of a host, and so are packages of annotations seeking a site for unification.

in the *epsilon calculus* with particularised type e_s terms for event/situation terms (Gregoromichelaki 2006), encoding tense, modality, aspect and other information (Cann 2011) (although only tense is represented here and, even that, only schematically). Scope relations are not represented architecturally in the tree, but indicated through incrementally collected scope constraints indicating the various dependency choices. These feed special interpretive rules dictating the interpretation of all terms as witnesses for their containing proposition (Kempson et al, 2001:ch. 7, Cann & Kempson 2016). Pairs of trees, one as adjunct to the other, can also be incrementally built up across a defined transition from some node in a current tree to the initiation of a second LINK*ed* tree (see section 5.5). The dynamic process of tree construction is partly top-down, as (i)-(iii) display, as macros of actions given by general strategies and individual words feed from one to another to successively update the previous transition, and then in turn become part of the context. But it is also partly bottom-up as kaleidoscoped in (iv) as actions induced by the words feed into the bottom-up process that operates from a mother node applying a modalised form of functional application to its two daughter nodes to yield the effect of compositionality defined on the resultant tree.

The trees in (i)-(iii) span the different types of underspecification available at intermediate steps in the tree-building process. Trees may simply lack a full set of the requisite nodes for expressing the resultant content, even to the limit of only containing a single node as in (i). Tree nodes may be assigned a type specification and yet lack fully determined specifications of content, as the place-holding terms of (ii)-(iii) show: $\mathbf{WH} : e$, the place-holder projected by *wh* terms, $S_{PAST} : e_s$ for the skeletal indication of the event term, $\mathbf{U} : e_s \rightarrow t$, the placeholder of the open propositional structure yet to be developed. And the presumption of underspecification extends even to the construction of tree relations, as in (ii) and (iii), the dashed lines indicating a node introduced into a tree without as yet a fixed relation to any node other than the one from which it was constructed, $\langle\uparrow_*\rangle$ being the more general relation, $\langle\uparrow^1_*\rangle$ a localised variant (more details later).

5.1.1 Building up or linearising content

A distinctive attribute of this system is that nothing in the construction of trees from such underspecifying input, or the linearisation of content via underspecifying devices, precludes overlapping processes of tree construction, subject to the actions triggered being commensurate. Both general and word-triggered strategies provide options which, subject to whatever prerequisite is defined for that option, are freely licensed: given the definition of the formal system as model theoretic, constraints can apply cumulatively through various compatible sources.

One primary illustration of such overlap is the feeding relation between *auxiliaries* and main verbs (but also other content elaborations on a single node, see e.g. (24) earlier). All verbs are defined as triggering actions inducing a skeletal propositional template; auxiliary verbs are defined similarly, but without the inducement of full predicate-argument structure. As such, auxiliary and main verbs will overlap in the actions they project. Taking a parsing perspective, the actions defined for the auxiliary *did*, for example, expand the partial tree of (ii) to yield (iii) with its (partially specified) event term, an event predicate node underspecified for content (shown by the *metavariable* $\mathbf{U}$), and one positionally underspecified individual argument node. The actions triggered by the bare verb-stem *hug* then overlay this, and go on to extend the emergent structure by determining the node annotated by the formula $Mary'$ as the logical subject, building and annotating the main predicate node with the formula Hug' and finally constructing an open individual term node for the internal argument of that predicate. Unifying the initially unfixed **WH**-annotated node and the open argument node then allows functional application to apply three times to yield the final tree (iv).

Such overlap in the projections of strings/structures distinguishes DS characterisations from all other grammars for whom it is words which inhabit structure; and it has the immediate advantage of making it possible to express local *discontinuity effects*, as nothing precludes the occurrence of an expression intervening between two co-dependent items as long as the structural condition for the update associated with the second item isn't jeopardised by the actions induced by that intervening expression. Again, (38) is illustrative, as the applicability of the update actions triggered by the verb *hug* is unaffected by the annotation of $Mary'$ having been added to the emergent tree in virtue of processing the word *Mary*. This is because the development the verb *hug* induces simply updates the structure provided by *did*, irrespective of whether the sequence of words is *did Mary hug* or *Mary did hug*.

There are further advantages to this shift into an action-based perspective in which underspecified representations are progressively induced. First, associating place-holding metavariables with lexical expressions such as pronouns, indexicals, and (English) auxiliaries provides a vehicle for expressing the way in which some types of natural language expression encode an under-determination of attributable content,[11] in this way giving overt status to the gap separating words from their assigned contents in context. It also provides a basis for seeing words as projecting very much more than just access to an address in memory where various associations and conceptualisations involving the contextually-assigned referent are stored. Taking the example of an auxiliary-main verb combination, we see that both auxiliary and full verb project the over-arching propositional structure, with the auxiliary projecting the finer levels of detail with respect to the event term, while the verb projects the greater granularity of argument specification and determines the relatively rigid word order via the generation of predictions as to how and when the verbalisation of the arguments will occur. This intrinsic procedurality of lexical "content" provides a natural basis for expressing variability across languages in terms of constraints on the contexts in which words such as verbs can be parsed. English is again illustrative: the actions of the English finite non-passive verb form turn on there being in the immediate context a locally unfixed type e node, which they then fix as logical subject, and end with a prediction for an object following, in DS terms, with pointer placement at the verb's object node, thereby ensuring the default SVO construal of DP V DP sequences in English.[12] The various feeding relations between these sequences of actions for update thus yield the ordering effects familiarly thought of as warranting a concept of autonomous syntactic structure to be defined over sentence-strings. Indeed, both local and long-distance dependency effects can be expressed through the articulation of the construction process itself; and even though bottom up compositionality of content as defined over the string is relinquished, compositional accumulation of contextually-derived content (see also Recanati 2010, Jaszczolt 2005) at each incremental step is achievable (Hough 2015; Purver et al 2011; Purver et al 2014).

In DS, this general mixed top-down bottom-up dynamic forms the central underpinning of the grammar, adopted equally in *parsing* and *generation*. Parsing includes aspects of generation as, in order to build conceptual structure, the parser does not wait passively for input to process, but generates predictions as to what steps and inputs are going to follow next. Such potential next steps, along with the inputs and actions that have already been utilised, are stored in a directed acyclic graph (context DAG: Sato 2011, Hough, 2015) that models the transitions from state to state. Generation, on the other hand, builds upon the same context state as the parser with the only difference that there is an additional control, namely, some more particular conceptual goal to be achieved through the transition to the next context state (simplistically expressing this development here only through tree-transitions, this control state can be termed as the *goal tree*). The goal-tree is a partial tree, exactly like a parse tree in form, that is at least a one-step extension of the current parse tree, serving as a restriction for future generation steps, guiding lexical access and production of words. As an example, for an already planned whole interrogative structure to be expressed, the tree in (38)(iv) would serve as the goal tree and the rest of the trees in the sequence in (38) will be licensed after being checked for their 'subsumption' of this tree (see Purver et al 2006, Purver et al 2014). Thus

[11]The intrinsic anaphoricity of auxiliaries is idiosyncratic to English, as witness their hosting predicate ellipsis unlike many other languages.

[12]Adjuncts as in (i)-(ii) would not undermine this result as they would invariably be taken to annotate a node in an adjunct LINKed tree:
(i) The dogs, yesterday, disappeared for two hours.
(ii) John, the idiot, disappeared for two hours.
Such an adjunct transition is characteristically indicated by an intonation break.

lexical access for the generator is followed by implicit testing of the syntactic mechanisms and contextual resources to determine that the update associated with a potentially produced word is extendable towards that goal tree so that the word will be licensed to be produced.[13] Thus speaker and hearer always perform mirrored actions induced through incremental complementary initiatives. As both speaker and hearer are modelled as assuming concurrently the roles of the parser (in integrating word-macros' outcomes in their context) or the generator (by inducing predictions/goals for upcoming structure), each next step initiative for performing an action is equally available to both; but once such an initiating action has been made manifest by the interlocutor both participants' contexts need to be updated ('grounding' Clark 1996). Any mismatches detected can then induce backtracking or further elaboration of the troublesome update (Eshghi et al 2015, Hough 2015). Accordingly, well-formedness is context-dependent and characterisable without making reference to any individual agent, hence applicable to both split and non-split dependencies. Since the grammar deals with contextual updates (and not just strings or string-meaning pairs), a contextual update is wellformed if and only if there is at least one possible extension of the partial structure induced by the rules and relative to the context within which the string is interpretable. So B's answer in (39) below is well-formed in the context provided by the question but not in many other contexts:

(39) A: What does Mary like? B: John thinks cake.

Inability to induce an update can induce contextual search backwards, the generation of clarification requests (Eshghi et al 2015) or coercion procedures in an attempt to accommodate the input by extending the grammatical resources (Gregoromichelaki, to appear).

5.2 Defining the context-update process

5.2.1 Formal properties of trees

To flesh out this dynamic perspective and see the substance of the constraints it imposes, we need first a vocabulary for defining the trees that, as a simplification here, we take to model conceptual content. This vocabulary incorporates means for defining how the nodes of such trees can be referred to from the perspective of any other node on a tree so that we can formulate predictions of upcoming structure and constrain future developments according to content having been processed or being expected. So, embedded within the incremental dynamic logic of context-state transitions, DS articulates a vocabulary for describing tree structure in the form of a modal logic language for defining, perhaps partially, tree properties at each state transition (Logic of Finite Trees (LOFT), Blackburn & Meyer-Viol 1994). In LOFT, trees are defined by the set of nodes that they include, and each node is uniquely identified by its relation to all other nodes in the tree: a tree node identifier accompanying each node in effect encodes this. So the vocabulary for describing trees eventually labels the root node as $Tn(0)$ (*Tn* being a predicate taking tree-node identifiers as its value), $Tn(00)$ for its argument daughter, $Tn(01)$ for its functor daughter and so on. At intermediate stages a top-node will be labelled as $Tn(a)$ and other nodes can be invoked by reference to this node. In LOFT, tree-nodes are seen as states and modalities are defined governing accessibility relations from one node to another. There are two basic modalities that express traditional dominance relations, $\langle\downarrow\rangle$: '$\langle\downarrow\rangle X$ holds at a node if X 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 there is the inverse $\langle\uparrow\rangle X$: '$\langle\uparrow\rangle X$ holds at a node if X holds at its mother (one node up)', equally with argument and functor variants indicative of the status of the daughter-mother relation so identified. Kleene star operators are also defined yielding concepts of (none or) multiple steps of *dominate* and *be dominated by*: $\langle\downarrow_*\rangle Tn(n)$ holds at a node when a node $Tn(n)$

[13]Part of the challenge to be addressed in such a task is how it is that the concept of prosodic word so commonly mismatches the units over which linguists standardly define their object of analysis, a puzzle standardly ignored (though see Lahiri and Plank 2010).

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. There are then the analogous $\langle\downarrow^1_*\rangle$ and $\langle\uparrow^1_*\rangle$ operators defining a node somewhere down/up along an arbitrary sequence of functor relations (a functor spine) (see (38) where the sequence of functor nodes decorated forms the sequence of right-edge nodes through the local tree); and these will give us concepts of *locally dominate* and *be locally dominated by*, which are used to refer to nodes within a single propositional domain. Tree-node identification can thus either be in terms of the direct algorithmic labelling of the nodes, $Tn(0), Tn(00), Tn(01), Tn(011)$ etc, or in terms of their modally specified relation to others in the tree under construction, as LOFT makes available. In order to keep the formal vocabulary to a minimum, our characterisations of tree-node identifiers will largely be expressed in terms of these modal relations, making use of just one itemised tree-node identifier, that of the root as $Tn(0)$. Hence the logical subject of a subject-predicate structure derived from an unembedded clause, being a node with the treenode address $Tn(010)$, is identified by the modality: $\langle\uparrow_0\rangle\langle\uparrow_1\rangle Tn(0)$. In either notation, the identification picks out each node in the tree uniquely. As we shall see, this unique identifiability of any node allows flexibility in the tree and string development process, as it allows these processes to be defined without having to explicitly exclude any overlap in the construction of nodes. The uniqueness of any node identification will ensure that if any node is constructed twice, such duplication will never lead to distinct nodes: any follow-up construction of a given relation will simply conflate with the first legitimate occurrence of that relation, overlaying it to yield a wellformed outcome as long as the annotations accumulated by such composite actions are commensurate. As a universal constraint on update, this will have an important role in what follows.

5.2.2 Defining underspecification

The core dynamic of the update process turns on the concept of underspecification, with a twinned notion of *requirement*, $?X$, for some annotation X. Such requirements are imposed goals guiding the parsing/generation process and allowing predictive unfolding of either tree structures, word strings, or contextual updates. Some requirements may be met immediately, but others only substantially later, as in (38), where the initial requirement $?Ty(t)$ isn't satisfied until the final step of the derivation. Such requirements apply to all types of annotation and may be modal as well as non-modal. For example $?\langle\uparrow_0\rangle Ty(e_s \rightarrow t)$ is a requirement expressing the constraint that the current node must eventually be immediately dominated by a completed event predicate node, a functor typed as mapping an event term node onto a closed propositional type. Context evolves along with such emergent representations of content and strings. Every step of update, once achieved, becomes part of the context relative to which the next update takes place.

The primary types of underspecification which have such attendant requirements can now be articulated in terms of the tree-description vocabulary. Firstly, trees may be partially specified in virtue of some formulae remaining yet to be developed. Any such node will have some requisite type requirement, $?Ty(X)$ for some possible type value X. Secondly, we have the relatively familiar underspecification of content formulae. Such specifications include a type but their content value awaits replacement. This is the way pronouns and predicate ellipsis are modelled. They are both taken to induce some place-holding metavariable of given type e.g. $\mathbf{U} : e$ for pronouns and definite DPs, $\mathbf{U} : e_s \rightarrow t$ for open propositional placeholders, as introduced by auxiliaries in English. These have the attendant requirement $?\exists\mathbf{x}.Fo(\mathbf{x})$, interpreted as a requirement for a fully specified formula value: metavariables, being mere procedural place-holders, do not qualify as appropriate permanent values as they are not interpretable objects in the formulae object language. Such *metavariables* are indeed a standard way of encoding how contribution made by some words achieves denotational status only in virtue of triggering selection of their construal from context. In English this attribute is intrinsic not only to pronouns and some determiners (*the*, and the demonstratives) but also to the auxiliary and modal verbs *do*, *be*, *have* etc. We also assume a particularised metavariable $\mathbf{WH}$ for *wh* terms of WH-questions.[14]

Thirdly, and more radically, tree-node relations may be incompletely specified. In (38)(ii)-(iii), we have two sorts of such underspecification involving *unfixed nodes*. One involves a simple dominance relation, defined from the perspective of some newly constructed node from some tree node $Tn(a)$ as $\langle\uparrow_*\rangle Tn(a)$ (38)(ii), a modality that is satisfied just in case the node with address $Tn(a)$ ultimately comes to dominate the relevant node at some fixed position. This provides the means of modelling long-distance dependencies. The more restricted modality $\langle\uparrow_0\rangle\langle\uparrow_*^1\rangle Tn(a)$ (introduced in (38)(iii)), induces an argument node constructed from some node which itself stands in an unspecified functor relation to the node from which this update was induced; and this requires resolution along an unbroken chain of functor-relations – 'a functor spine' – hence within a local predicate-argument structure. This strategy underpins locally discontinuous dependencies, an effect in other frameworks commonly requiring some externally imposed process of permutation. All unfixed nodes, whatever their modality, remain in need of subsequent updating by being assigned a fixed position within the tree that satisfies the modal relation between the relevant node and its dominator. This is determined by the requirement $?\exists\mathbf{x}.Tn(\mathbf{x})$ which can only be satisfied by some fully specified value for the treenode attribute Tn.[15] As noted above and discussed in more detail below, the modal logic underpinning the tree displays has the necessary consequence that there can never be more than one instance of an unfixed node with the same modality within some partial tree. Consider the attempted analysis of **What did Mary John give?* (with the intended interpretation 'What did John give Mary?'). The first three steps in the analysis would be as in (38)(i)-(iii) with the parse of *Mary* giving an appropriate annotation for the locally unfixed node, giving us a node with the specification $\{\langle\uparrow_0\rangle\langle\uparrow_*^1\rangle Tn(0), Fo(Mary'), Ty(e)\}$. A further attempt to add another locally unfixed node, this time annotated by a parse of *John* will induce a node $\{\langle\uparrow_0\rangle\langle\uparrow_*^1\rangle Tn(0), Fo(John'), Ty(e)\}$ but as this has exactly the same modal relation to the rootnode as the previous locally unfixed node, the two nodes inevitably collapse to give $\{\langle\uparrow_0\rangle\langle\uparrow_*^1\rangle Tn(0), Fo(Mary'), Fo(John'), Ty(e)\}$ which is incoherent by virtue of the distinct nature of the concepts named by the formula values.

5.2.3 Defining context and actions for context-state extension

The concepts of underspecification given so far are a heterogeneous set, but nonetheless are all static notions, properties of a given content-string configuration available for update. The next step is thus to define the concept of context, in order to see what concept of update is definable.

Context With actions as the central player in the process defined, the concept of context itself has to reflect the integral dynamics.[16] Accordingly, context is not just a record of whatever complete and partial trees emerge across the construction process: it also records the action sequences that have yielded these updates through the successive transitions, evolving in tandem with the ongoing construction of content, making it substantially richer than is expressible in either model-theoretic accounts or semantically-blind syntactic accounts. *Context*, thus, is a dynamic, multi-modally induced record of (a) words; (b) conceptual content notated as tree structures; and (c) the sequence of steps in building the emergent trees – a sequence of partial trees, and the actions that effect the transitions between them. In Sato (2011), Purver et al (2011) and Eshghi et al (2012) this is extended and expressed as a Directed Acyclic Graph (context DAG) where each node represents the current (partial) tree and each edge in the graph records a potential action to be taken, to allow for full characterisation of the availability of choices at various stages of the derivation, and

[14] The precise specification of the **WH** metavariable is not provided here, given that quantification is also not taken up in this paper.

[15] Low case variables shown in bold are rule-level variables (not metavariables) that will unify with whatever specified value occurs on the node label (Kempson et al. 2001).

[16] An initial formal definition of the contents of context states was given in Cann et al (2007).

the possibility of clarification, acknowledgement, correction and so on (see also Hough 2015, Kempson et al 2015).

A consequence of this richer concept of context that the grammar manipulates is that processes which trigger recovery of information from context, are licensed to recover various types of information: re-use of content (semantic formulae) copied from a node in some (partial) tree; re-use of sequences of actions used to construct the current partial tree; re-use of phonological strings in cases of clarification or repetition in general; and re-use of structure, i.e. extension of some partial tree in context. To these we add the potential recoverability of information directly from the utterance scenario as represented by the processor, yielding indexical construals (see Cann 2015 for discussion of these).

The process of developing this richer notion of context is modelled through the dynamic logic underpinning DS which includes a set of actions for formulating context updates. Here, confining ourselves to actions for tree/string extensions, there are four primitive actions stipulated: $\texttt{make}(X)$, $\texttt{go}(X)$ and $\texttt{put}(Y)$ operations, where X and Y are tree relations and node-annotations respectively, plus an $\texttt{Abort}$ action, which terminates an action sequence. These are used to define a range of stored action sequences ('macros'). It is these macros of sequenced actions which come to form part of the evolving context as that develops in tandem with each step of content construction. The action macros fall into two different types: general *computational actions* and *lexical actions* though these differ solely in whether they are licensed only by some encoded trigger. All macros of actions, however induced, are given in a standard conditional $\langle$ IF-THEN-ELSE $\rangle$ format, enabling whatever conditions need to be imposed on the context update to be explicitly formulated as a pre-condition of the macro's application.

Computational Actions constitute generally available strategies, either inducing the unfolding of an emergent tree on a top-down basis, reference to context for retrieval of elements already processed, or inducing bottom-up processes which, once appropriate terminal nodes are annotated, lead to annotations for all non-terminal nodes. The licence to construct a node characterised solely as being *unfixed* within a given tree domain is, as we've already seen, central to the system. This is one of a very restricted number of initial steps in a derivation for some novel tree: a first step licensing the building of just one as yet not fixed relation, expressed by the $\langle\uparrow_*\rangle$ operator, with its attendant requirement for a fixed treenode position, $?\exists\mathbf{x}.Tn(\mathbf{x})$. The rule inducing this step is constrained to operate only if no dominated other node already exists within this newly emergent tree (the second preparatory IF condition in (40)):[17]

(40) *ADJUNCTION

Actions

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

Output Tree

$?Ty(t), Tn(a)$

$?Ty(e),$
$?\exists\mathbf{x}.Tn(\mathbf{x}),$
$\langle\uparrow_*\rangle Tn(a)$
$\diamondsuit$

The alternative variant of an unfixed node which licenses the building of a locally unfixed node, without any pre-condition for the absence of other already constructed nodes is:

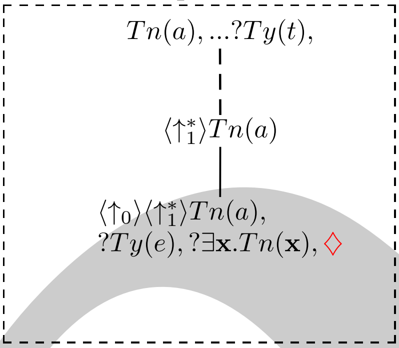
[17]The modality $\langle\downarrow\rangle\langle\downarrow_*\rangle$ refers to some node that is constructed under the current node while $\top$ requires that node to be annotated in some way.

(41) LOCAL*ADJUNCTION

Actions

IF	$?Ty(t), Tn(a)$
THEN	$\texttt{make}(\langle\downarrow_1^*\rangle); \texttt{go}(\langle\downarrow_1^*\rangle);$
	$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle)$
	$\texttt{put}(\langle\uparrow_0\rangle\langle\uparrow_1^*\rangle Tn(a), ?Ty(e), ?\exists\mathbf{x}.Tn(\mathbf{x}))$
ELSE	Abort

Output tree



The difference between these variants lies primarily in the distinct modality, thereby licensing their co-presence in a partial tree as in (38)(iii). The action of the first is restricted in that the position of the node has to be a position within the emergent hierarchical configuration within which it is to be constructed having no other node as yet constructed (this is a hard-coding of the restriction which arguably follows from general principles). The restriction applies also in embedded domains, so it is a constraint on semantic tree developments which yields a left-periphery effect within a clause, rather than a constraint ensuring initial position in some overall string. The action of the localised variant is less restricted in so far as it can operate freely within a given domain yielding "scrambling" effects (see section 5.4.1).[18] However, the domain within which it operates is more restricted: the argument whose construction it licenses must have its role within the containing structure resolved along an unbroken functor spine; and this restricts it to being resolved within a local predicate-argument structure. Both of these transitions will be used in our sequence of partial trees depicting a process of semantic tree development for (38).

Lexical actions It is now straightforward in principle to articulate what words and morphemes contribute to the ongoing development of partial trees, even though here is where the detailed language-specific granularity needs to be expressible.[19] The same action vocabulary is used for both general computational actions and lexically triggered actions, the latter being listed in the lexicon, together with the word that calls up the macro it induces. For each such lexical entry, given a certain (sometimes complex) set of conditions, a macro of action sequences induces some update yielding a distinct partial tree, string, or context state. The simplest macro of lexical actions is that which is triggered by the parse of a proper name like *Mary*. Given a context that requires the construction of a term (the *trigger*), the parse of the name merely annotates the node with the formula value representing the concept specified by the name and the information that a formula of the appropriate type has been identified:[20]

(42) *Mary*:

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

[18] Cann 2011 argues that the applicability of Local *Adjunction in English carries the same restriction as *Adjunction in not being re-iterable.

[19] DS makes no essential difference between lexically encoded elements according to whether they may be considered to be independent words or bound morphemes, provided that they are associated with specific sets of actions yielding some monotonic/compositional output. A theory of morphology within the framework has yet to be articulated.

[20] This simplicity is partly spurious, with no account given here of any putative internal logical structure of names. Additionally, such lexical entries are a simplified presentation of allowing the grammar to access memory locations storing various forms of encyclopedic and inferential information feeding mechanisms of concept construction (see Cann & Kempson 2016, Gregoromichelaki, to appear).

As already graphically displayed, *verbs* induce more structure than other words. They are defined as a conditional action inducing all the nodes that are needed to project onto a propositional structure using the predicate which the verb corresponds to, hence a node for their own predicate (of whatever adicity) and nodes for its attendant argument formulae. However, the building of these may be subject to a range of conditions. On the one hand, languages vary as to what options for the argument node annotations they license (see e.g. Bittner (2014b) for a range of variations). Like other Germanic languages, the actions defining English verbs specify argument nodes as merely having a requirement of the form $?Ty(e)$ (or $?Ty(t)$ for propositional attitude verbs). This ensures that in each case, there has to be some further step of processing in order to satisfy the type requirement with some appropriate formula value. Less rigidly, verbs in pro-drop languages may license their argument nodes with a place-holding metavariable, allowing their value to be identified directly from context. On the other hand, languages vary as to where in the sequence of expressions inducing such structure the verb may occur; and, English being highly restricted, this is determined by the particular form of the verb. Finite verb forms of English determine two such conditions. First they cannot occur initially:[21] this is the secondary preparatory condition of the verb *hugged* in (43), which determines that, for this update to take place, there must already be some term decorating a locally unfixed node within that structure, expressions like **Hugged Mary Jo?* being automatically excluded. Main verbs in English also determine the fixed position of this unfixed node as logical subject for active forms (via the requirement $?\langle\uparrow_0\rangle\langle\uparrow_1\rangle Tn(0)$) and some logical object for passive ones (Cann 2011). The post-verbal position of the syntactic object in active forms is determined by the construction of a logical object node that carries the pointer $\diamondsuit$, being the final subsequence of actions, thus ensuring that this is the next goal to be achieved.

hugged:

(43)
$$\left|\begin{array}{lll}
\text{IF} & Tn(a)?Ty(t) & \\
\text{THEN} & \text{IF} & \langle\downarrow^1_*\rangle\langle\downarrow_0\rangle Ty(e) \\
 & \text{THEN} & \texttt{go}(\langle\downarrow^1_*\rangle\langle\downarrow_0\rangle);\texttt{put}(?\langle\uparrow_0\rangle\langle\uparrow_1\rangle Tn(0)) \\
 & & \texttt{go}(\langle\uparrow_0\rangle\langle\uparrow^1_*\rangle Tn(a)) \\
 & & \texttt{make}(\langle\downarrow_0\rangle);\texttt{go}(\langle\downarrow_0\rangle); \\
 & & \texttt{put}(Ty(e_s), Fo(\mathbf{U}_{PAST}), ?\exists\mathbf{x}Fo(\mathbf{x}));\texttt{go}(\langle\uparrow_0\rangle) \\
 & & \texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);\texttt{put}(?Ty(e_s\rightarrow t)); \\
 & & \texttt{make}(\langle\downarrow_0\rangle);\texttt{go}(\langle\downarrow_0\rangle);\texttt{put}(?Ty(e));\texttt{go}(\langle\uparrow_0\rangle); \\
 & & \texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);\texttt{put}(?Ty(e\rightarrow(e_s\rightarrow t))); \\
 & & \texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle); \\
 & & \texttt{put}(Fo(Hug'), Ty(e\rightarrow(e\rightarrow(e_s\rightarrow t))));\texttt{go}(\langle\uparrow_1\rangle); \\
 & & \texttt{make}(\langle\downarrow_0\rangle);\texttt{go}(\langle\downarrow_0\rangle);\texttt{put}(?Ty(e)) \\
 & \text{ELSE} & \text{Abort} \\
\text{ELSE} & \text{Abort} &
\end{array}\right.$$

As noted above, the burden of structure projection is often shared between like-category expressions, and in English, this is a role carried by *auxiliary verbs* supporting the verbs they co-occur with. First, they project the topmost structure of this propositional skeleton, including a node for an event term, to which they project a typed place-holder as a proxy event term to carry whatever tense/aspect restrictor is projected by other expressions yet to be processed. But secondly, they also play a crucial role in determining how the structure emerges across the sequence of words. Without the constraint encoded in main verbs that, in English, prevents their appearance string-initially, the auxiliaries are free to occur in that position and moreover function to identify such an order as within a particular speech-act range (here given simply

[21] Except in imperatives, which we ignore here.

as mood-indicator Q).[22] Thirdly, again as complementary to verbs, they are precluded from occurring in the latter stages of development of the propositional structure, i.e past any fixing of predicate-argument structure, this being the secondary prerequisite condition in (44). The effect is that in positions other than the string-initial one, the sequence of actions induced by the auxiliary develops the very same structure but without any such mood indicator. In English, therefore, auxiliaries play an anticipatory role, initiating the first steps in the development of some skeletal structure, which the verb can then later develop. In an intuitive sense, they are thus a "light" verb that opens up a frame for the main verb then to expand with this verb elaborating some of the very same nodes already introduced by the auxiliary. [23]

(44) *did*

$$
\left|\begin{array}{llll}
\text{IF} & Tn(a), ?Ty(t) & & \\
\text{THEN} & \text{IF} & \langle\downarrow\rangle\top & \\
 & \text{THEN} & \text{Abort} & \\
 & \text{ELSE} & \text{IF} & \langle\downarrow_*^1\rangle\langle\downarrow_0\rangle\bot \\
 & & \text{THEN} & \texttt{put}(Q); \texttt{make}(\langle\downarrow_0\rangle) : \texttt{go}(\langle\downarrow_0\rangle); \\
 & & & \texttt{put}(Ty(e_s), Fo(\mathbf{s}_{PAST})); \texttt{go}(\langle\uparrow_0\rangle) \\
 & & & \texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle); \\
 & & & \texttt{put}(Fo(\mathbf{U}), Ty(e_s \rightarrow t)); \texttt{go}(\langle\uparrow_1\rangle) \\
 & & & \texttt{make}(\langle\downarrow_*^1\rangle\langle\downarrow_0\rangle); \\
 & & & \texttt{go}(\langle\downarrow_*^1\rangle\langle\downarrow_0\rangle); \texttt{put}(?Ty(e), ?\exists\mathbf{x}.Fo(\mathbf{x})); \texttt{go}(\langle\uparrow_0\rangle\langle\uparrow_*^1\rangle) \\
 & & \text{ELSE} & \texttt{make}(\langle\downarrow_0\rangle) : \texttt{go}(\langle\downarrow_0\rangle); \\
 & & & \texttt{put}(Ty(e_s), Fo(\mathbf{S}_{PAST})); \texttt{go}(\langle\uparrow_0\rangle) \\
 & & & \texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle); \\
 & & & \texttt{put}(Fo(\mathbf{U}), Ty(e_s \rightarrow t)); \texttt{go}(\langle\uparrow_1\rangle) \\
\text{ELSE} & \text{Abort} & &
\end{array}\right.
$$

Details aside, what these lexical specifications illustrate is how words induce very much more than just access to conceptual content, in fact sometimes no conceptual content is contributed (e.g. complementisers like *that* in English). Some elements can project something intermediate, as do pronouns, and, in English, the auxiliary verbs. The auxiliaries in particular induce both an array of actions building up a set of nodes and underspecified determination of content for those nodes. As a comparison between the lexical specifications of auxiliary and main verb makes clear, defining the projection of structure in terms of procedures for inducing trees provides a rich granularity to the account of how structure is induced, allowing more than one word to jointly build structure in some local domain. There is thus not a one-to-one correspondence between word and node in a tree. To the contrary, words rely on existing structure and in turn induce a partial structure which some subsequent word or phrase may develop.

5.3 Combining actions in sequence: WH auxiliary-subject inversion as a case study

We now have all we need to display a full set of actions determining the parsing or production of the string *Who did Mary hug?* that was sketched in (38). As there, we start with an initial requirement for a propositional tree, a single node annotated as $\{Tn(0), ?Ty(t), \Diamond\}$. The actions in (40) induce an unfixed node

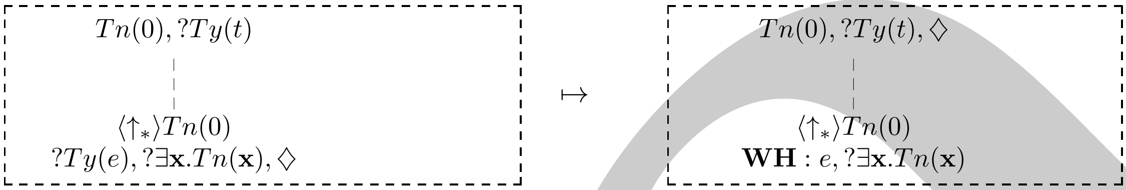
[22] Auxiliary inversion, a relic of V-2 effects from Germanic languages has come in English to be associated with non-assertions like questions, exclamatives, and somewhat archaically also with covert counterfactual conditionals where inversion serves instead of the conditional to convey counterfactuality occurring optionally with negative adverbials such as *never* and *rarely*. Such determination affects mainly the contextual parameters' specification, more specifically, the world of evaluation parameter (Gregoromichelaki 2006, 2011, to appear). In DS, there is no one-to-one mapping between mood indicators and speech acts, see e.g., Gregoromichelaki and Kempson (2015).

[23] Cann (2011) provides detailed analyses and separates the general properties of auxiliaries in their mood-identifying capacity from more idiosyncratic properties of the individual auxiliaries/modals by defining a general but lexicon-internal macro of actions.

with an open term requirement as in the first tree in (45) which allows the parse of *who* to yield an unfixed node annotated with the particularised metavariable $\mathbf{WH}$ (which has no attendant formula requirement: an unanswered question is not grammatically ill-formed) with the pointer returned to the topnode as in the second tree in (45).

(45)

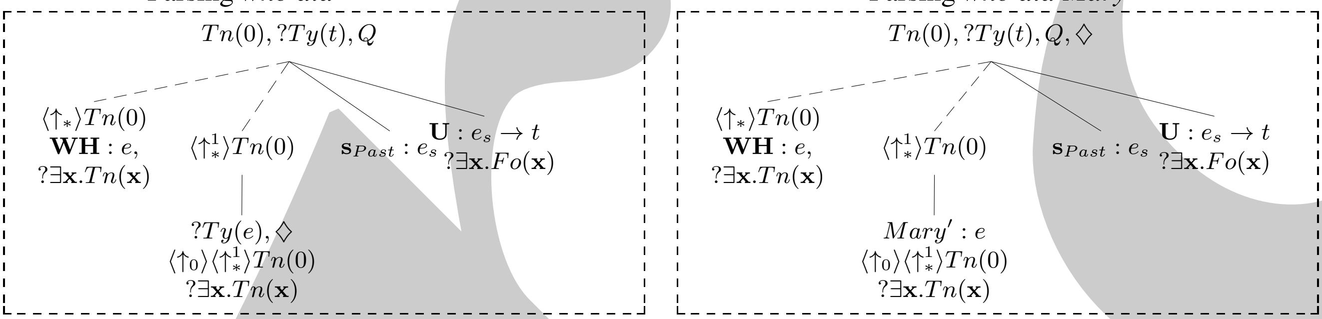
(i)*ADJUNCTION (ii) Parsing *who*



The actions in (44) associated with parsing the auxiliary, as we have already seen, initiate the construction of the propositional template, limited to the construction of a fixed event term and a fixed event predicate annotated with a metavariable. In addition, the topnode is decorated with the mood indicator Q and a locally unfixed node which carries the pointer ensuring that this node is the next to be developed. On the parse of *Mary*, the node is appropriately decorated and again the pointer returns to the topnode, both outputs shown in (46).

(46)

Parsing *who did* Parsing *who did Mary*



As with the finite form, the actions associated with the non-finite main verb *hug* induce the argument structure of the verb, but in a more restricted environment: one in which there is a type complete, but content underspecified, event predicate node that dominates no other node. This set of conditions, shown in (47), is sufficient to exclude the appearance of the form after *have* or *be* (**Mary has hug John, *Mary be hug John*) and to ensure that again the verb appears only after a subject has been constructed, as the modals and *do* always construct or check for a locally unfixed node in the way described above. [24]

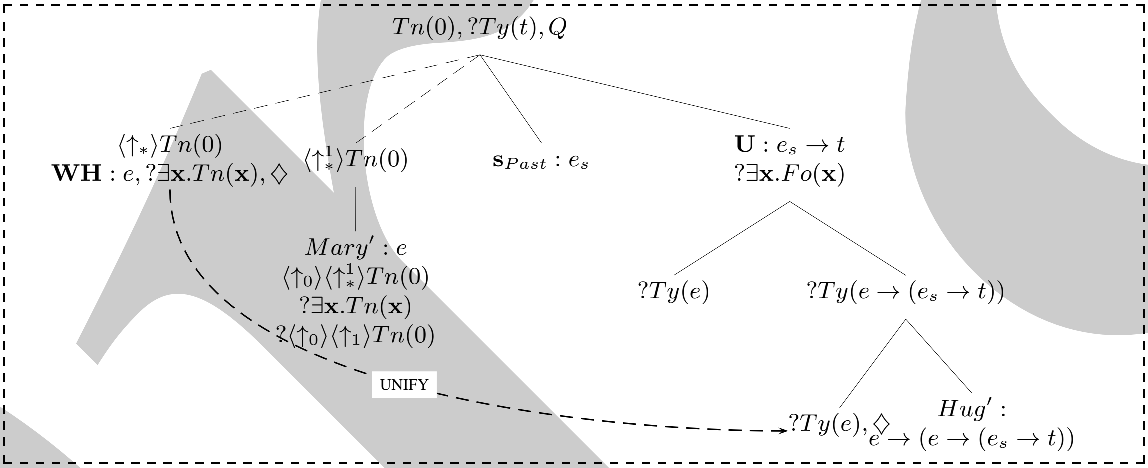
[24] In Cann 2011, main verbs, finite and non-finite, also annotate the event term with information about the Aktionsart of the verb. This refinement is not included here.

(47) *hug* (base form):

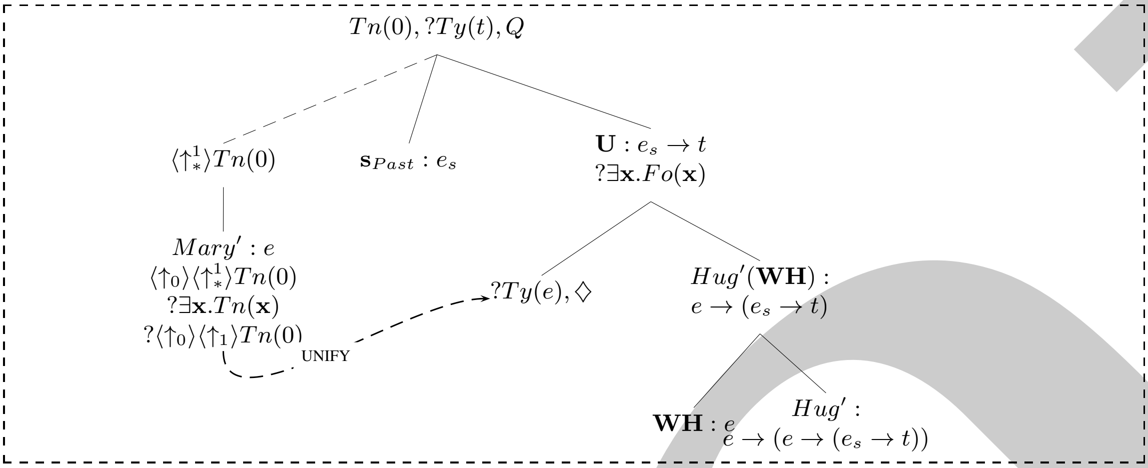
IF	$?Ty(t) \wedge Tn(a)$		
THEN	IF	$\langle\downarrow_1\rangle\langle\downarrow\rangle\top$	
	THEN	$\mathtt{Abort}$	
	ELSE	IF	$\langle\downarrow_1\rangle(Ty(e_s \to t) \wedge ?\exists\mathbf{x}.Fo(\mathbf{x}))$
		THEN	$\mathtt{go}(\langle\downarrow_*^1\rangle\langle\downarrow_0\rangle); \mathtt{put}(?\langle\uparrow_0\rangle\langle\uparrow_1\rangle Tn(a));$ $\mathtt{go}(\langle\uparrow_0\rangle\langle\uparrow_*^1\rangle); \mathtt{make}(\langle\downarrow_1\rangle\langle\downarrow_0\rangle); \mathtt{go}(\langle\downarrow_1\rangle\langle\downarrow_0\rangle);$ $\mathtt{put}(?Ty(e)); \mathtt{go}(\langle\uparrow_0\rangle); \mathtt{make}(\langle\downarrow_1\rangle); \mathtt{go}(\langle\downarrow_1\rangle);$ $\mathtt{put}(?Ty(e \to (e_s \to t))); \mathtt{make}(\langle\downarrow_1\rangle); \mathtt{go}(\langle\downarrow_1\rangle);$ $\mathtt{put}(Ty(e \to (e \to (e_s \to t))), Fo(Hug'))$ $\mathtt{go}(\langle\uparrow_1\rangle); \mathtt{make}(\langle\downarrow_0\rangle); \mathtt{go}(\langle\downarrow_0\rangle); \mathtt{put}(?Ty(e))$
		ELSE	$\mathtt{Abort}$
	ELSE	$\mathtt{Abort}$	
ELSE	$\mathtt{Abort}$		

The effect of these actions is shown in the tree in (48) with the pointer at the open internal argument node. At this juncture, the unfixed node bearing the **WH**-metavariable can unify its content with that of the open term node, simultaneously satisfying the type requirement on the latter and the treenode requirement on the former, and a subsequent step of unification (after one step of functional application) combines the locally unfixed node with the logical subject node as in (49).[25] The remainder of the type requirements in (49) are satisfied by two further steps of functional application to yield the final tree in (38) with the output formula $((Hug'(\mathbf{WH}))(Mary'))(\mathbf{s}_{Past})$.

(48)



[25] Note that the locally unfixed node cannot unify with the logical object node because of the imposition by the main verb that it must ultimately have the modality of the logical subject node, i.e. $\langle\uparrow_0\rangle\langle\uparrow_1\rangle Tn(0)$.

(49)


So we see that defining update mechanisms for on-line incremental building of representation of content provides a natural basis for expressing both short and long-distance dependencies.[26]

5.4 Word-order variability

It might seem that, without structure being defined over words and constituents and without a concept of syntax independent of the building of contents from strings, the grammar is too liberal to account for traditional concepts of well-formedness and grammaticality. Nor might it seem possible to account for typological variation in word or constituent orders. The framework is – deliberately – liberal with respect to aspects of language use in that it does not *a priori* exclude things that are often considered prescriptively ungrammatical but are nevertheless common in the spoken language such as the common use of resumptive pronouns in relative clauses in spoken English (Cann et al, 2005b). Nevertheless, it is not the case that 'anything goes'. All word order variation and the constraints on the possibilities are determined by (a) the contextual conditions and licensed updates defined by computational rules, (b) the underlying logic in which they are expressed, and, (c) language-specifically, lexical actions associated with words and morphemes.

Take the basic relatively strict SVO ordering of English and its interaction with left dislocation. As noted above, the fact that main verbs have to appear after the syntactic subject is guaranteed by the contextual condition that a locally unfixed node is already constructed disallowing **Hugged Mary John*. Equally, since any locally unfixed node is required to be complete, auxiliary plus main verb clusters are disallowed before subjects: **Did hug Mary John?* or *Who did hug Mary?* with the interpretation 'Who did Mary hug?'. That the locally unfixed node is identified as syntactic subject with the interpretation of logical subject in finite verbs and some logical object for passives is determined by the fact that the parse of the verb fixes the role, and thus the position, of that node, within the emergent propositional structure. Together with the constraints associated with unfixed nodes this ensures that *John, Mary hugged* cannot be interpreted as 'John hugged Mary'. At the same time, auxiliaries are disallowed postverbally because of the condition that no individual predicate node have been constructed for a sucessful parse, thus disallowing strings like **Who Mary hug did?*. Other ungrammatical strings are equally disallowed because of the interaction of the constraints imposed on the successful firing of general rules. For example, **Did who Mary hug?* cannot be parsed/generated at all: the WH-word has to annotate a locally unfixed node since the construction of the event term and event predicate nodes by the auxiliary precludes the firing of the rule of *ADJUNCTION; this in turn means that *Mary* cannot be parsed because a re-construction of a locally unfixed node will collapse with the existing node; the result is an incoherent formula annotation since the WH metavariable cannot

[26] We have not in this paper illustrated "true" unbounded dependencies as exhibited in examples like *Who does Jane think Roger sacked?*. However, such cases are covered by the modality $\langle\uparrow^*\rangle$ which is not sensitive to type specifications and so ranges over all dominance relations, irrespective of their clausal or other properties.

be updated from within its own containing tree. And even if it could so update and had done so, then there would be no means of satisfying the requirement that there be a second argument to *hug*. Hence the ungrammaticality. Additionally, there are strings like *Who Mary did hug* which, while perfectly acceptable as a relative clause, is not possible as a root question. This is ensured by the fact that, without the update provided by the auxiliary taking as input a structure in which there is no already constructed locally unfixed node, the string cannot be interpreted as a question, i.e. the mood indicator Q is not projected onto the containing propositional node.[27]

5.4.1 Verb-final languages

The challenge at the other end of the spectrum is the existence of verb-final languages which allow extremely free ordering of DPs in a simple clause with only the verb positioning as necessarily final being rigidly determined:

(50) *supai-ni, shorui-o, haanarisuto-ga watashi-ta*
spy-IO document-DO journalist-SUBJ hand-PAST
To the spy, the document, the journalist handed [Japanese]

All orderings of the three DPs in (50) are possible, and without necessary rhetorical distinctiveness (Kempson and Kiaer 2010). Informally, it is often reported, such freedom poses the puzzle that the structure these expressions induce doesn't seem to be projected until the verb is processed, so such sequences would seem necessarily to impose a delay in implementing structure, each such pre-verbal expression being in some sense unfixed. They thus might appear to provide a severe challenge to the claim put forward here that syntactic properties of natural language are grounded in incremental processing, which is furthermore constrained to preclude the construction of more than one instance of any tree relation at a time, here more than one locally-unfixed node. However, such languages notably display a rich case system, and, as Nordlinger (1998) observes, it is this morphological sub-system which carries the functional load of inducing structure in anticipation of the verb that follows, so that by the time the verb is processed, the structure it induces serves simply to instantiate and so confirm that already induced by the sequenced case-marked DPs. This insight is straightforwardly reflected in the account proposed here. The substance of the claim turns on the restriction that at any point in the derivation there cannot be more than one instance of any given tree relation. All that is required then to see these sequences as being commensurate with the restriction is to assume that each step of constructing and annotating a locally-unfixed tree node (as required to process the three DPs *supai-ni*, *shorui-o*, *haanarisuto-ga* in (50)) must be necessarily followed by an ("abductive") set of actions which resolves the underspecified relation of that unfixed node to the determinative relation encoded in the case-marker. It is notable in this connection that case-suffixing is obligatorily final in the DP sequence: all NP-modification is rigidly ordered before the nominal plus case-suffix. The case-marking thus serves a two-fold purpose. First it encodes the precise hierarchical position within the emergent predicate-argument structure of the node which the expression it suffixes annotates: i.e. in DS terms, a requirement of $?\langle\uparrow_0\rangle Ty(e_s \rightarrow t)$ for a subject-marker, $?\langle\uparrow_0\rangle\langle Ty(e \rightarrow (e_s \rightarrow t))$ for a direct object marker, and so on. Secondly, being obligatorily final, the case-marker encodes the information that the processing of that sub-structure is complete: nothing more needs to be added. And unless resolution is immediately established, no possible derivation is achievable. All that we then assume here is that case in these languages is indeed "constructive". The dynamic of such end-placed suffixes is thus carried out in successive steps across the

[27] This account incidentally commits us to presuming that even in questions in which it is the subject that is questioned, the subject term initially annotates a regular unfixed node i.e. initially not characterised as inside the local predicate-argument structure, only becoming so by merging with the locally unfixed node once that is projected through actions of the auxiliary so as to license its identification as logical subject in the presence of the verb.

sequence of the derivation. And this means that there will never be more than one underspecified tree relation at a time: each will be built and annotated as locally unfixed and immediately updated, thereby giving the pointer licence to proceed with a subsequent step of Local *Adjunction, leading in similar fashion to an introduced platform node thereupon resolved.

By this route, the succession of DPs progressively induce the substructure which the various case-markers encode, so that upon the processing of the verb, its induction of the full array of predicate-argument structure sufficient to satisfy the adicity of its predicate will be no more than an overlaying of the structure already induced through the processing of the sequence of DPs; and its own tense suffix, being a propositional operator of type $t \rightarrow t$, will thereupon trigger the bottom-up compilation of content. The formal analysis thus reflects the intuition of speakers that case is carrying the functional load of structure projection prior to the processing of the final-placed verb. [28]

5.4.2 Meeting morpho-syntax challenges: morphological "spreading" as tree co-construction

One further application of the co-construction of a node is its applicability to the puzzle, in free word-ordering languages, of multiple encodings of a single grammatical function over a number of different morphemes in a composite expression. The classic example is that of *concord*, agreement of morpho-syntactic indicators like case, number, or gender, which may license freer discontinuity effects than more rigid word-order languages allow. For example, in the Latin example in (51), nominative case is morphologically encoded not only on the end-placed subject noun *ocelli* ('eyes') but also the apparently left-dislocated adjectival modifier *turgiduli* ('swollen') positioned before the verb *rubent* ('be-red'):

(51)	*Flendo*	*turgiduli*	*rubent*	*ocelli.*
	weeping-ABL.SG	swollen.DIM-NOM.PL.MASC	be.red-PRES.3PL	eye.DIM-NOM.PL.MASC

'(Her) little eyes are red, swollen with weeping' [Catullus, *Carmina* 3]

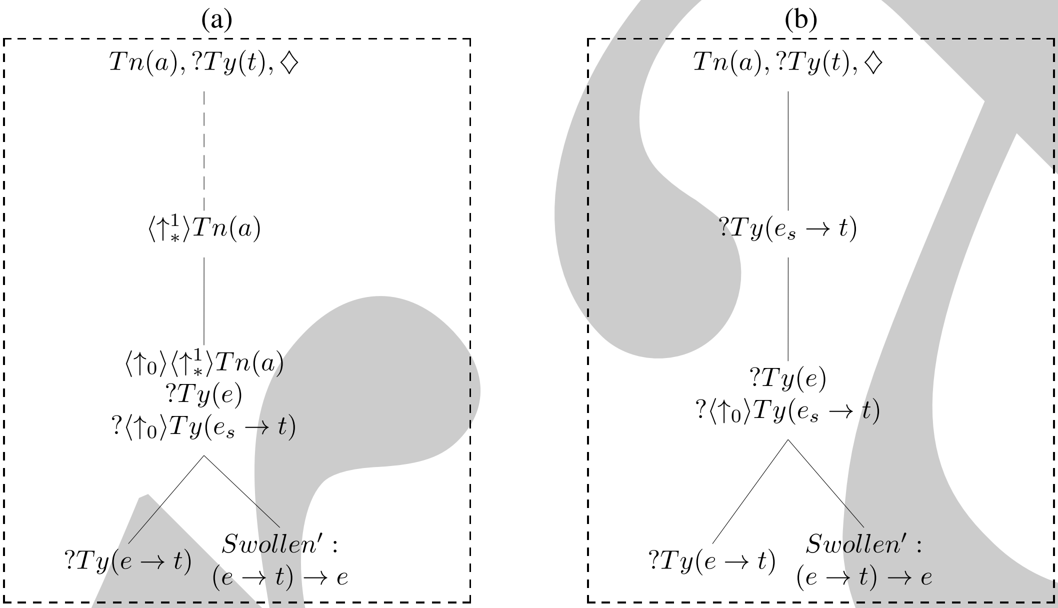
A fairly standard account of this type of case agreement would have the head noun imposing the case on the containing noun phrase with any agreeing modifier copying that value to be morphologically realised, yielding a situation in which case-marking may have a grammatical function (on the noun) or may have no force at all (on an adjective) (see for example Carstens 2000 and Baker 2008 among many others). This can then be taken as an argument for an independent level of morphosyntactic restrictions with no semantic impact whatsoever. However, the underlying tree logic of DSinvariably exploited for functionally-driven purposes during processing. The case-marker has a unitary grammatical function, no matter what type of expression realises it morphologically: all such marked expressions impose constraints on their containing term node which collapse harmlessly into one, no matter how many words express that case. This specification can now be exploited for pragmatic purposes: in a case-rich language like Latin, noun and modifier can be separated for emphasis with the case-marker now assuming the processing function of directing the hearer's predictions in a certain direction so as to expect a particular type of upcoming head noun. In the same way, the same case-marker can not only restrict what the upcoming linguistic input will be but also indicate what conceptual structure the uttered noun-phrase needs to fit in, i.e. direct the hearer to a specific interpretation.

To sketch how this works, let us (somewhat simplistically) assume, for the Latin case, that nominative signals that the case-marked noun phrase will provide the highest individual argument (the subject) of the

[28] Confirmation of this account comes from what looks at first to be problematic. It is observed in the minimalist literature that more than one instance of Move α is possible in these languages though tightly constrained by such expressions being required to be construed as within the same local sub-structure. This too is just as the DS account would predict, as with nothing precluding a feeding relation between *Adjunction and Local *Adjunction, the sequenced actions of building a locally-unfixed node and immediately updating it can apply within a substructure which is itself unfixed within the containing emergent structure, but the effect of such a sequence would indeed be the constraint that two such apparently long-distance moved expressions have to be interpreted as within the same local domain (Kempson & Kiaer 2010).

predicate, a property that can be ensured within a DS tree by marking a term node with the requirement that the conceptual value it bears must combine as the first argument after the event node, expressible as the requirement $?\langle\uparrow_0\rangle Ty(e_s \rightarrow t)$. Our focus here will be exclusively on expressing the discontinuity of an adjective-noun sequence both nominative-marked, and accordingly both to be interpreted jointly as associated with the subject of the main verb which, in (51), intervenes between them.[29] The action of parsing first the adjective *turgiduli* ($Swollen'$) with its nominative marking is defined to induce a locally-unfixed term node marked as being required to be a logical subject, of which it itself projects a predicate modifier, as in (52a).[30] This requirement enables a tree-development where the unfixed node is updated as being dominated by a node of the right type, thus satisfying the nominative requirement, giving the structure in (52b).[31]:

(52)



The parse progresses with the verb *rubent* ('be-red') projecting a propositional template. The final noun *ocelli* is then parsed through the construction, again, of a locally unfixed term node again annotated with the nominative requirement through the case on the noun, (53a). This node then unifies with the argument node already fixed in processing the adjective, a harmless super-imposition exactly analogous to the co-annotations which auxiliary and verbal annotations induce, yielding the tree in (53b).[32]

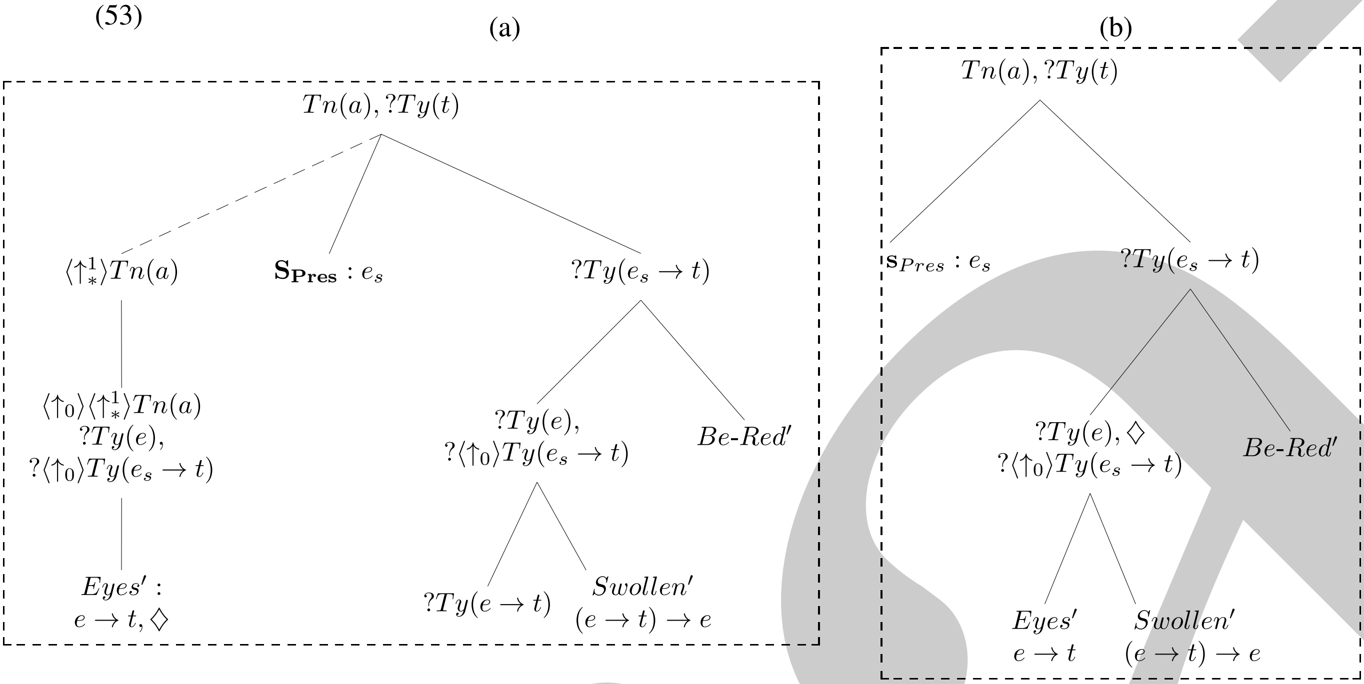
[29]The analysis ignores altogether the contribution of the gerund *flendo*, and also the internal structure of the term phrase, for simplicity. The typing assigned both to the nominal as of predicate type and its predicate modifying adjective as of type $(e \rightarrow t) \rightarrow e$ are ad hoc simplifications for expository purposes only, as is the omission of any epsilon binder.

[30]On trees here we show the English form of concept representation, eg $Swollen'$, $Eyes'$, as, without their morphological affixes, Latin stems are hard to decipher, threatening the transparency of the displays.

[31]Requirements upon a node's development are removed if operations have taken place to determine their satisfaction: see Kempson et al (2001) ch.9 for formal details of the various components of the formalism that take care of such "bookkeeping" activities. In such case-rich languages, we also assume that case-licensed processing steps induce the inferential strengthening of extant underspecified relations (see Cann & Kempson 2008; Gregoromichelaki 2013a).

[32]Projecting the required epsilon binder, omitted here for simplicity, is achieved by defining both adjective and nominal as including actions which induce such term substructure relative to whether or not a binder has already been introduced, this equally yielding the effect displayed here of apposition via superimposition (see Cann et al 2005a, ch.7 for an account of apposition).

(53)



Exactly the same type of account can be given for other, highly problematic case-number-gender discontinuities in morphologically-rich languages, such as Greek, while at the same time explaining the existence of morphological marking and concord as enabling context-sensitive processing, as illustrated in (54), and earlier in the (26)-(27) cases:

(54)	*Ena*	*grizo*	*epsaxna*	*kanape*	*(alla*	*den*	*vrika*
	a-ACC.MASC.SG	grey-ACC.MASC.SG	seek-PAST.1SG	sofa-ACC.MASC.SG	(but	not	I.found
	pouthena).						
	anywhere).						

‘I was looking for a grey sofa (but I couldn’t find one anywhere I looked)’ [Modern Greek]

It is because DS representations are domain-general integrating content derived by various modalities, not just language, that indexically-construed case-marked subsentential expressions can exploit their own case-marking to induce the structural context licensing their integration within an appropriate conceptual structure (see Gregoromichelaki 2002, 2013a, 2016). Dynamic Syntax can thus model multiple, discontinuous, and indexical morphological marking and, through it, provide an explanation for the existence of such marking, not as an arbitrary syntactic mechanism, but through its function in facilitating the conceptualisation of the communicated eventualities in an efficient, incremental and predictive manner.

5.5 Capturing locality constraints

With lexical macros inducing content incrementally through both absorbing constraints from the linguistic and non-linguistic context and subsequently also adding further constraints to it, we can now reconceptualise standard locality restrictions via the tree logic describing partial binary-branching trees.

The local tree for a simple predicate-argument structure, for example, will always contain an unbroken chain of functor steps along its right edge, as in (38)(iv). This is because the first non-functor node above any node annotated by the predicate will be the type t node, and each further unfolding of a pair of functor-argument nodes follows this pattern. Constraints invoking concepts of locality can thus be defined as constraints on tree-context accessibility, as we’ve seen with case annotations. Reflexive pronouns, for example, provide a metavariable requiring substitution by the value provided by an argument node within a tree determined by an unbroken functor path from the immediately dominating node which the reflexive itself has annotated – i.e., in DS terms: $\langle\uparrow_0\rangle\langle\uparrow^1_*\rangle\langle\downarrow_0\rangle Tn(a)$ ('$Tn(a)$ holds along one argument-relation up, plus a possibly empty sequence of function-path relations plus one argument relation down'):[33]

(55) *myself:*

IF	$\langle\uparrow_0\rangle\langle\uparrow^1_*\rangle\langle\downarrow_0\rangle Fo(\mathbf{x})$
	$Speaker'(\mathbf{x})$
THEN	$\texttt{put}(Fo(\mathbf{x}))$
ELSE	`abort`

Principle B restrictions on pronouns are equally definable relative to a complementary constraint. Thus we see that locality constraints, often thought to require constraints specifically of syntactic form, are straightforwardly definable over the emergent structural representations of content.

At the other end of the locality parameter, adjunct trees can be built externally to the tree currently under construction as part of the incremental building process yet without necessarily being part of the tree over which compositionality of content is defined, in such cases serving instead to enrich the context as that emergent content is being established. Such paired trees are secured through a process which determines the sharing of a term in the two structures, yielding a pair of the tree containing the point of departure and some secondary so-called LINKed tree. The canonical exemplar of such paired structures is relative clauses, whose content is overtly indicated by the relative pronoun, this being the expression which imposes the sharing of such a term across the two structures:[34]

(56) I approached John, who was at the front of the queue.

This mechanism is of very general applicability, not being type-restricted: it can be used to license nonpropositional as well as propositional, context-enriching devices, amongst which are apposition devices as in (57)-(58), and also the many different types of attitudinal adjuncts, politeness indicating adjuncts, respect markers, and so on:

(57) John, the leader of the team, who had initially advocated change, subsequently voted against it.

(58) John, the idiot, tragically, ignored the rest of the team.

Indeed, we would add to this list the array of explicit speech-act indicators which may accompany what are otherwise simple predicate-argument structures, and the explicit representation of attributes of the utterance event itself, which display canonical adjunct-like behaviour in not being obligatory:

(59) The bathroom light, please/if that's OK by you/I insist, should be left on at night, as Tommy is scared of the dark.

[33]Lower case bold variables are rule-level variables (not metavariables) unifying with specific values on the current tree at which they are processed (see Kempson et al 2001).

[34]Given the analysis of quantification as invariably involving variable-binding term operators of type e, a natural model of restrictive vs nonrestrictive relative clauses becomes available. On the one hand, a variable of type e is introduced as a platform on which to construct a possibly complex restrictor given by the noun (of predicate type), and, on the other hand, the resulting type e term once the variable with its attendant restrictor is duly bound by the epsilon operator. On this basis, we can anticipate at least two distinct types of relative clause arising from the stage of the construction process at which some LINK transition is induced: either, at the point of constructing the variable, building a LINKed structure sharing that variable (the *restrictive* relative construal), or, at the point of having compiled the resulting epsilon term, building a LINKed tree containing a copy of that bound term (the *nonrestrictive* construal) (see Kempson et al 2001, Cann et al 2005a for details). Here we illustrate only with nonrestrictive forms.

These auxiliary tree-extending sequences can thus yield both propositional and non-propositional content ((56), (57)-(59)), with both types contributing to providing the context for the formulation or satisfaction of further predictions.

Being inherently a context-extending mechanism, the LINK modality provides a means of explaining certain types of conversational behaviour. One such is the way that people can extend utterances apparently indefinitely, providing incremental extensions to their conceptualisations of an eventuality (Guhe 2007):

(60) John came back. Late last night. Extremely tired and frustrated.

In (60), there are two adjunct extensions to the initial apparently complete sentence: the first modifying the event of returning, the second modifying the subject *John*. Standard frameworks at this point require a complex system of backtracking to parse such an utterance in order to embed the adjuncts into their defined syntactic positions within the syntactic structure. In DS, the LINK modality provides a simpler means of accounting for the data. A parse of the initial utterance provides a complete propositional tree that acts as the context in which the subsequent utterances are to be interpreted. Parsing (and producing) the second utterance in (60) involves constructing a new propositional tree LINKed to the event term. That event term is copied over to provide the argument of the more specific temporal specification. The third utterance is then interpreted using another propositional structure, this time LINKed to the individual subject argument with the formula annotating that node carried over to become an argument of the two predicates. We thus end up, after a process of tree evaluation (Cann et al 2005a: 209-10) with the coordinate logical form in (61):

(61) $[(Return'(John'))(s_{Past})] \wedge [(Late\text{-}last\text{-}night')(s_{Past})] \wedge [(Tired' \wedge Frustrated')(John')]$

In principle, this sort of extension can be carried on indefinitely and across multiple participants without requiring the constant backtracking and reconstruction of some representation of constituent structure within the output string, which would be necessary if tree-building operations could only license fixed tree-relations.

5.5.1 Defining local processing domains

Nonetheless, these side-sequences, being auxiliary and structurally external, provide a natural barrier outlawing the possibility of discontinuous dependencies being resolved outside the local domain they define, an effect familiarly known as *strong island constraints*:

(62) *Who did you see John [who likes __]?

These constraints have been argued to involve a sui-generis syntactic condition on structure that is not reducible to any semantic or functional property, a putative counter-argument to the DS stance. However, within DS, these are modelled as constraints on incremental processing, intrinsic to the structural framework. On this analysis, all such structures involve temporary interruption of the processing of an already initiated tree-structure, for example, in order to provide elements that may have just emerged as essential during utterance planning. For this reason, the pointer indicating the current node of development has jumped via a LINK transition to start a second LINKed structure with its obligatory shared term. Formally, the LINK relation is a tree modality, indicated as $\langle L \rangle$, distinct from the general dominance relation $\langle \downarrow_* \rangle$ which is defined over functor and argument nodes. In consequence, long-distance dependencies, which involve relations of arguments to their functors, cannot get resolved within LINKed constructions (Cann et al 2005a; Gregoromichelaki 2006).

Stepping back from the details, we have thus seen, that by incorporating within the grammar a formal reflex of the ongoing time-linear processing, we have removed the need for additional syntactic and morpho-syntactic levels of string-structure representation, thereby reducing the degree of formal and notational complexity of the grammar.

6 Cross-turn syntactic licensing: split utterances

For all the claimed success of the DS framework in matching analyses against which other frameworks have been evaluated, the most striking indication of its potential explanatory force as a grammar formalism lies in the emergence of split utterances as an immediate consequence, given that more conventional frameworks have no means of expressing this phenomenon in any straightforward way. Recall that DS mechanisms model the grammatical licensing involved both in the comprehension of strings (parsing) and linearisation of conceptual structures (generation). In both cases, licensing involves the satisfaction of predicted goals; both comprehension and production manipulate the same representation and are equally incremental; and the only distinction between them is that production must always be at least one (subpropositional) step ahead in terms of conceptual structure. Every licensing of a word in generation has to be checked against a partial *goal-tree* representing how the speaker is going to extend the context next in order to make sure that the lexical macro employed in the next step complies with an extension towards this goal-tree. Thus, on this view, production is predicted to be slightly slower than comprehension and to be later to emerge in language acquisition.

This bidirectional modelling of processing results in the mirroring of speaker and hearer actions, with both attempting to satisfy predicted goals in the next processing step. Lexical access in DS relies on such predictions to retrieve lexical macros, appropriately conforming with the goal-tree, from the lexicon (Purver et al, 2006, Purver et al 2014, Hough 2015, for formal details). In consequence, and with production slightly lagging behind, a possibility licensed by the grammar is for the hearer to elect to verbalise a lexical item that they have retrieved from their own lexicon, satisfying joint or diverging predictions, instead of waiting for the previous speaker's selection. This does not present any processing problems for the previous speaker, who midway can shift seamlessly into becoming the hearer, because, with production assumed to be incremental, they can integrate the input of the interrupting co-participant instead of continuing with their own lexical search and verbalisation (although the latter remains a possibility as can be seen below):

(63) Dan: but it seemed to be, to Ken at least
Roger: the wrong kind
Dan: another kind of distinction. [modified example from Lerner (2004)]

(64) Ken: He said the married couple will walk- walk down the street and they will be all dressed up and people will come by with ...
Louise: rice.
Ken: rice, petals or anything they think is suitable. [modified example from Lerner (2004)]

To see in detail why such transfer of control in producing an apparently joint utterance is so seamless, let's take first the split utterance of (3) repeated here:

(3) A: Did Jo . . .
B: stumble? I hope not.

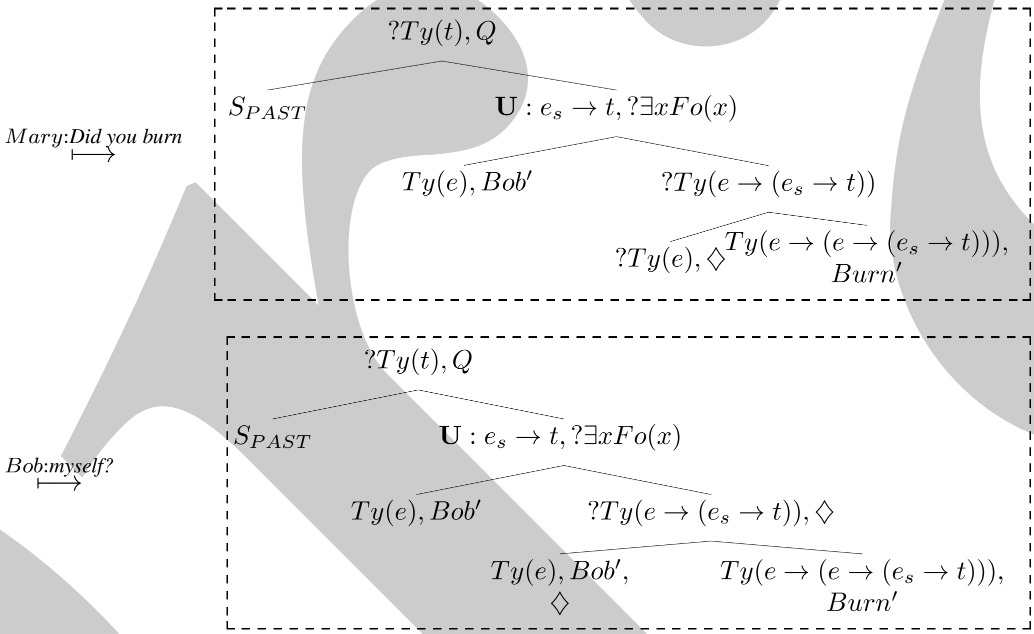
As we've already seen, the sequence of an initially placed auxiliary plus DP gives rise to a very partial top-down structure in which one locally unfixed node labelled through the processing of subject is annotated, as here by the term Jo' within a structure with a past-tense-specified event term. Otherwise the predicate structure of type $e_s \rightarrow t$ is wholly undeveloped, annotated just with predictions for an upcoming verbal expression. This underspecified structure will be common to both A as speaker and B as hearer at this juncture, though in A's case with some goal-tree as the intended development in mind though not necessarily with all decisions for its realisation fully established (the goal tree may be only partial). At this juncture, B may retrieve some lexical item satisfying the predicted continuation and choose to intervene to develop the current common underspecified structure in ways commensurate with B's own preferences, whether or not B envisages this as what A might have had in mind. In relinquishing control of the form of the utterance, A has merely to abandon his/her lexical search (unless A chooses to query B's offering) in order to develop his/her own pending structure on the basis of what B has offered (see Eshghi et al 2015 for the modelling of clarification questions in DS terms). All in all, interpretation and production by either party, in either role, is achievable in a wellformed and wholly straightforward way.

To now see how morphosyntactic restrictions can, and must, be licensed cross-turn and distributed across speakers, even in a mostly analytic language like English, let us take up the split utterance of (65), a simplified variant of (28), at the point of Bob's continuation:[35]

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

(66) displays the partial tree upon Mary's utterance having been processed. At this point, the metavariable projected by Mary's utterance *you* has been substituted by a conceptual representation of the current hearer, Bob. Given this as a context, Bob is now licensed to complete the utterance but 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:[36]

(66)



[35]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, see Eshghi et al 2015.

[36]We omit here the representation of the contextual parameters' shift at each stage, e.g. speaker-hearer roles (for details, see Kempson et al 2007, Gargett et al 2008, Gregoromichelaki & Kempson 2015, Gregoromichelaki, to appear, Purver et al 2010, Eshghi et al 2015). As pointed out by a reviewer, the involvement of non-ratified participants e.g. an interruption by a speaker assuming the perspective of a 'bystander' can cause illformedness when agreement information clashes with current contextual parameters:
A: *You usually*
C (not previously speaker/hearer): *burn(s) himself?*
This endorses the DS claim that contextual parameters are relevant for grammatical wellformedness and indicates that not all continuations are predicted to be felicitous (see Eshghi 2011, Eshghi & Healey 2015 for the relevance of various types of conversational participants' contexts and Gregoromichelaki, to appear for perspectival parameters included in the contextual representation).

The fact that the structure may have been induced in part by the other participant is irrelevant for well-formedness. This phenomenon, as here modelled, is specific to the construal of the reflexive pronoun, said to be subject to a Principle A constraint (Chomsky 1981); 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 2005a, Gregoromichelaki 2006, 2013a). In both cases, the morphosyntactic restrictions impose constraints circumscribing the goal-directed search for achieving the interpretation of linguistically-underspecified elements.

The point illustrated here is general. 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. Shifts of roles from speaker to hearer within the domains of all syntactic and semantic dependencies are unproblematic in this framework, discontinuities across more than one participant thus displayed even in cases of syntactic dependencies which are commonly presumed to be expressible only through a syntax-specific vocabulary:

(67) A: Mary, John ...
B: trusts?
A: Well, he clearly likes her.

In this case what A and B severally as speaker and hearer have built at the point of switch is a partial structure with one unfixed node pending (via actions induced by the word *Mary*) and one locally unfixed node (via the word *John*) (as the only way to legitimate two such bare DPs within a single structure in English). Both trees also include predictions for upcoming verbal expressions to provide positions for the unfixed nodes. Following the switch, the actions induced by the particular form of the verb will fix the interpretation of $John'$ as subject (in virtue of the property of English finite verbs determining that the locally unfixed node already in the structure under construction must be construed as logical subject). 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; and the shift of roles is anticipated correctly to be seamless.

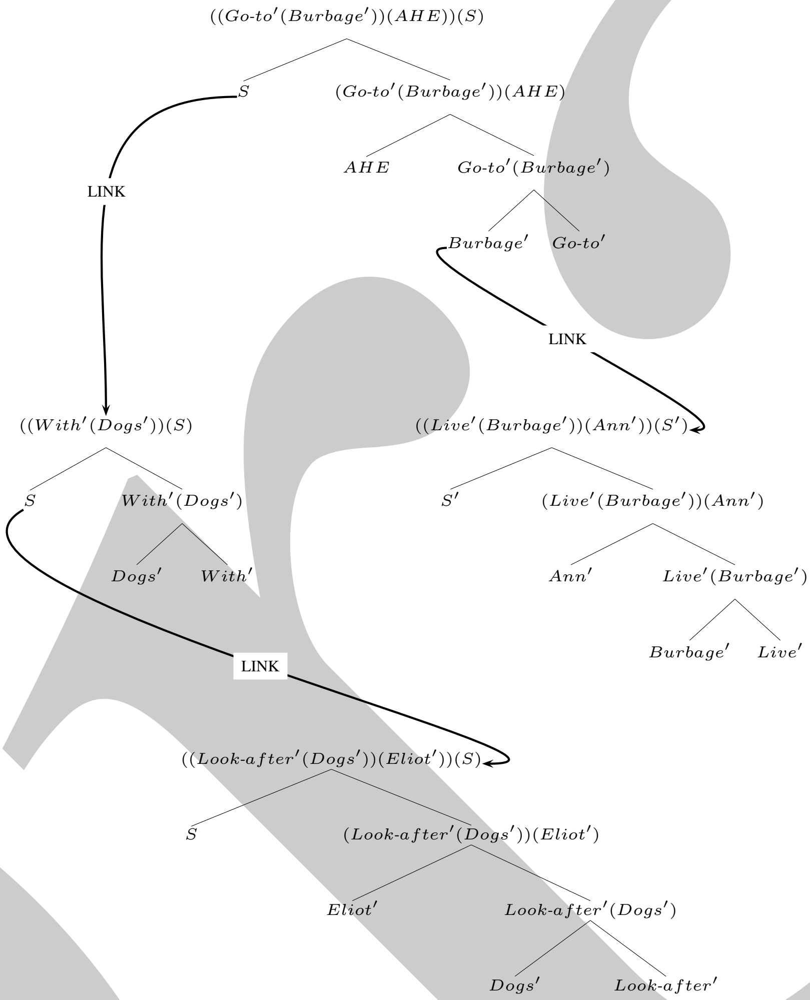
We can now give a sketch of one of the more complex examples set out initially (section 2), (6) repeated below:

(6) Alex: We're going to ...
Hugh: Burbage, where Auntie Ann lives
Eliot: with the dogs?
Hugh: if you look after them.

Here the conversational exchange turns out to be a conditional only relatively late on in the exchange, with a construal modifying only a subpart of the structure presented. As we have seen, adjuncts in DS are taken to induce the building of auxiliary structures defined across a LINK transition definable from any node under development within a given structure. So, firstly, the relative *where Auntie Ann lives* can be induced as a modifier of the named village, Burbage (the relative pronoun *where* being taken to trigger the construction of an appropriately second occurrence of the concept *Burbage'* within this new emergent LINKed structure) ensuing in a propositional assertion. Eliot's addition to the utterance of *with the dogs*, on the other hand, is taken as a modifier of the event term already initiated in processing the main clause *We're going to Burbage* uttered jointly by Alex and Hugh, a move which induces a distinct LINK transition from the event term of the proposition constructed from that main clause. This then yields an expanded event specification 'We're going to Burbage with the dogs' construed as a query. Conditional antecedent clauses project a complex modifier to the event specification of the consequent clause, the content of the antecedent processed as

LINKed to this event term via the term copied into the LINK structure induced by parsing *with the dogs* (Gregoromichelaki 2006, 2011; for parenthetical conditionals and other word orders see Gregoromichelaki 2006, Ch. 4). So here the newly developed event term with restrictor 'We are going to Burbage with the dogs' can now be modified by Hugh's utterance of *if you look after them* via yet a further LINKed structure presenting the conditional restrictor to that evolving event term. The pattern is exactly that given for (61) – a network of structures severally LINKed to one another through a succession of shared terms in each pair of such structures each addition serving potentially a distinct speech-act function. The effects of all these processes put together is shown very schematically in (68), with the 'core' proposition extended by a number of LINK structures whose content has been variously contributed by the three interlocutors:

(68) Schematic analysis of (6):



And, finally, in the early intervention case of (24), simplified here as:

(69) A: The doctor
B: Chorlton? A: Mhm. He said...

the handover is none other than the regular device of apposition as could have been used by the speaker herself as *The doctor, Chorlton, (Chorlton) (he) said that......* This kind of structure is also analysed as a LINK transition, this time from one type e node to induce the construction of another to be paired with it as providing an extension to the restriction of the epsilon term under construction, which, in not having developed here the epsilon term account of quantification, we leave as a promissory note.[37] Its significance at this juncture is that both parties will separately have constructed the input structure licensing the building of a LINKed structure notably across a nonpropositional transition; and the shift of roles causes no interruption in the incremental accumulation of content.[38]

Thus, the perspective of an action-based explanation of syntactic/semantic constraints blends seamlessly with action coordination as achieved in human interaction without having to invoke reasoning mechanisms operating on the conceptualisation of metadescriptions of such actions (Gregoromichelaki et al 2013a), i.e. without any recourse to mind-reading or externally imposed steps of inference. Furthermore, in being a consequence of the general dynamics enshrined in the framework, this potential for extending non-sentential constructs in conversational exchanges is predicted to be universal.

7 Language as mechanisms for action-coordination

7.1 Grammar as constraints on content underspecification

With this characterisation of split utterances to hand, in terms of interactively getting structures to grow, we finally wish to probe its explanatory strength by asking whether core individual mechanisms of such a grammar are themselves grounded in the potential they provide for joint construction of meaning, rather than this being simply some side-effect of little consequence.

7.1.1 Pronominal anaphora

The pronoun-antecedent relationship provides the most familiar case of collaborative building of structure: antecedent recovery for pronouns is uncontentiously taken to be licensed irrespective of who was responsible for the antecedent of the pronoun. It is thus this canonical pattern that we take as the base case for our point of comparison. There are three different sources for pronoun-antecedent recovery – anaphoric, cataphoric, and indexical. First, there is the provision made available by previous linguistic material. Such pairings can be distributed cross-sententially and sub-sententially across speakers, whose success may indeed induce a jointly-achieved speech act as in (70), an effect notably modellable only if an incremental perspective is assumed:

(70) A: He
B: or she
A: yes, they would do their utmost to cause us trouble.

Secondly, again a familiar phenomenon, is that pronouns may provide a licence for later (cataphoric) provision of a value, a license for delay, which is made possible on the DS account by the lexical provision of a place-holding metavariable. As with backwards anaphora resolution, this again can either be resolved within one person's utterance, or distributed across more than one person. The core instances of these are the so-called *expletive pronouns*, like *it* in English, which have become conventionalised for this specifically anticipatory function:[39]

(71) A: It's obvious that I'm wrong.

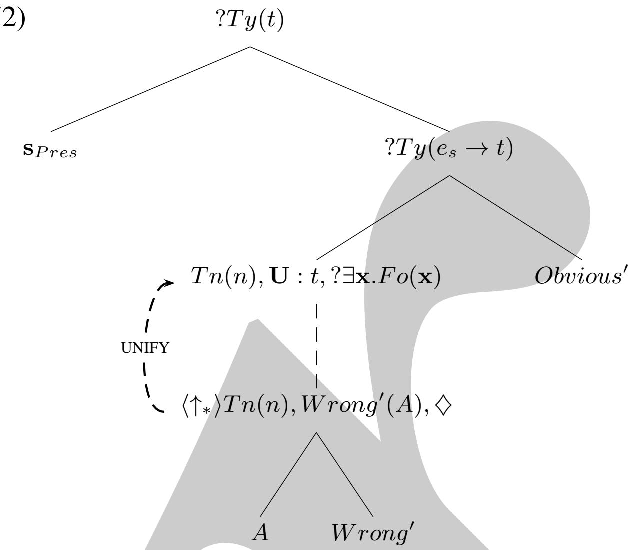
[37] An evaluation of two such paired type e terms involves a re-binding of the extended term to create this result, see Cann et al (2005a) ch.7.

[38] Clarifications may of course be the basis for pinpointing disagreement in the structures severally under construction by either party, which will involve adjustment of structure to yield a commonly established result (see Hough 2015, Eshghi et al 2015 for detailed analysis of such examples).

[39] There are other instances of cataphora: see Gregoromichelaki (2006), Ch. 5 for more details of the DS analysis, and Gregoromichelaki (2013a) for grammaticalised cataphoric effects in Modern Greek clitic constructions.

As in (65), there is a locality constraint (in Ross 1967 known as the "Right-Roof constraint") here determining the environment within which the value of that initial pronoun must be established. This expletive pronoun we assume here is underspecified for type, licensed to range over either t or e (see Cann et al 2005a, Chapter 5).[40] Given DS assumptions, this constraint is a simple consequence of the emergent compositionality of the resulting tree: license to delay provision of a value to a metavariable (as provided by a pronoun like *it*) is possible but only until that point at which the various parts must combine together to yield an interpretation for the whole.[41] This point is reached in expletive structures as soon as the sister predicate to that pronoun has been constructed (in (71) following the incorporation of the content of *obvious*). It is at this point therefore that the pointer will return to the subject node so that combination with its sister yield a value for their mother. In such cases, a provisional metavariable place-holder for a formula value annotates that sister subject node (earlier provided by *it*), so further steps have to be taken to develop it (recall, the sequencing of the words in the string does not stand in one-to-one correspondence with the structure established from the string). In virtue of the presence of the metavariable, a locally-unfixed node expanding the subject node is the only option that will lead to a well-formed result. Such a node is duly built and duly annotated with the content of the clause; it is then immediately unified with the subject node above it:

(72)



As before, nothing in the system determines who has to be responsible for making available such provision of the value of the expletive, and the take-over by another party (with due shift in the person marking on the indexical pronoun) is fully well-formed:

(73) A: It's obvious
B: that you are wrong.

Thirdly, there is the indexical interpretation of anaphoric expressions:

(74) [Context: A is contemplating the space under the mirror while re-arranging the furniture and B brings her a chair]
A to B: That's/it's perfect. Thanks.

[40] With a fully developed theory of event terms as epsilon terms binding the entire content of the propositional structure, this underspecification of typing for the expletive and the nonhomogeneous characterisation of the adjective *obvious* as allowing a type $t \rightarrow t$ classification is avoidable, but we leave all these details aside (see Gregoromichelaki 2006 for detailed exegesis, including rules converting any type t to being the restrictor of an epsilon term hence of type e).

[41] The only other DS delaying mechanisms are associated with late incorporation of unfixed node content, see Gregoromichelaki 2006, Gregoromichelaki 2013a

And, as we would now expect, there may also be the potential for interplay that we saw earlier in a case rich language such as Greek ((27) is repeated here):

(75) [same context] A: That? You must be joking.

(27) [same context]
A to B: tin karekla tis mamas? / # i karekla tis mamas?
the-ACC chair-ACC of mum's? / # the-NOM chair-NOM of mum's?
(Ise treli?) (Are you crazy?)
[clarification, Modern Greek, Gregoromichelaki 2012]

What both (75) and (27) illustrate is not only the task of identifying the concept to be constructed (as antecedent to the demonstrative pronoun in (75) and the definite DP in (27), but also the effect of "constructive case" (see section 5.4.1) in morphologically-rich languages that indicates the function of the non-sentential utterance within the utterance scenario (Gregoromichelaki, 2016) with an additional indexically-recovered speech-act function described here as 'clarification' (for such clarifications as involving contextual backtracking in DS see Eshghi et al 2015). What lies at the heart of modelling the fluent interaction between participants and their utterance situation is the domain-general vocabulary in which the DS context model is expressed, for it is this which allows such seamless feeding of information culled from the utterance scenario to the language processing activity.

7.1.2 Predicate ellipsis

This pattern carries over to the construal of VP ellipsis. The dynamic patterning again falls into three major categories. In English, it is elliptical fragments with an auxiliary for an open propositional development which display the textual anaphoric property (the so-called VP-Ellipsis). Given the lexical specification of auxiliaries as place-holders in exactly the same manner as pronouns, with indication of the event term to be constructed to aid the selection process, this parallelism in establishing anaphoric dependence is as we would expect; and the interactive effect, too, is expected:[42]

(76) A: You were tired yesterday, weren't you?
B: I wasn't.
A: You were. I saw you nodding off in the back row.

Secondly, there is also the potential for construal from a subsequent source, as we saw in detail in section 5.3, with the derivation in which the auxiliary could project a skeletal architecture leaving an open propositional metavariable which the following verb, projecting the same overall architecture could go on to develop. We can now see how the projection of the place-holding formula projected by the auxiliary directly follows the anticipatory pattern of expletive pronouns and is subject to the very same locally restricted form of anticipation.

[42]The VP-ellipsis pattern of strict and sloppy interpretations, the one involving an identical copy of the predicate form at the ellipsis site, the other a parallel but distinct interpretation at the ellipsis site, applies equivalently to pronoun resolution, in what have been called *lazy pronouns* as in (ii) (Karttunen 1969):
(i) Bill checked his paper for typing errors, and so did Harry.
(ii) My parents always do their tax returns during the summer, but we never do them until the last minute.
On the DS account, sloppy ellipsis construals involve reiterating the action sequence attributed to the antecedent in context (see Purver et al 2006, Kempson et al 2015 for details), the re-application of such an action sequence relative to the new subject (in the second conjunct) determining the parallel but distinct interpretation. Lazy construal of *them* in (ii), in similar manner, reiterates the set of actions associated with its antecedent, reconstructing, as the antecedent had, a term containing a metavariable as a subterm which in being resolved in the new environment duly yields the distinct construal of *them* as "our tax forms" (whoever *we* refers to).

And finally the case of indexical construal, familiar for pronouns and as anticipated, also VP-ellipsis fragments, given their characterisation as projecting a metavariable of open-propositional type. VP-ellipsis fragments can indeed be interpreted/produced directly from a shared context as in (77) with A offering to B an open propositional skeleton to develop relative to their joint observation of the scene in front of them, the particular choice of modal serving to narrow down the choice of open propositional formula to be constructed:

(77) {A standing beside son with surfboard in hand checking the waves}
A: I wouldn't if I were you. Notice the red flag.

So again we see the pattern of interactive anaphoric, cataphoric, and indexical resolution being available for ellipsis resolution.

7.2 Syntactic mechanisms as coordination devices

7.2.1 Unfixed node building: a mechanism for interaction

We turn now to the building of unfixed nodes, the mechanisms taken to underpin what are standardly taken to be classic cases of (long-distance) discontinuity and, supposedly, an irreducibly syntactic phenomenon. If, as we claim, grammatical devices have a common domain-general source, namely, underspecification and update, unfixed-node constructions should behave similarly to anaphora and ellipsis. Moreover, analysing the patterns that result from the employment of unfixed nodes as sui-generis mechanisms misses a major generalisation in the universal articulation of natural language grammars.

So we now turn to examine whether the building of unfixed nodes and their resolution allows the pattern displayed by canonical instances of joint construction such as achieved by pronouns and predicate ellipsis. If so, it too should be resolvable distributively across more than one participant preceding it in the exchange, subsequently in the exchange, and immediately from context that has previously been constructed, or indexically from the more general context of the utterance scenario.

This time we take forward resolution of such an unfixed node from within the construction process, and across a jointly-achieved structure, as this is the hallmark of long-distance dependency:

(78) A: The books, I'm told are not worth us insuring.
B: The Assyrian horse ...
A: it's obvious you must insure.

Here, arguably, the expression *the Assyrian horse* is taken to annotate an unfixed node. This node's subsequent resolution is not only subject to whatever constraints are imposed by the need to resolve the expletive whose substituend it is contributing to develop, but also needs to occur after a change of context as speaker/hearer roles switch mid-utterance (as indicated by the indexical *you*).

Secondly, the building of such an unfixed node underpins *stripping/bare-argument ellipsis*, the form of ellipsis in which the fragment is just a bare argument expression which itself triggers the inducing of a predicate from antecedent linguistic context:

(79) A: I think Jane would make a good choice for President.
B: or Sue.

Here one interlocutor provides the actions leading to a complete proposition, which the other then makes use of. In DS terms, the unfixed node initiated by B in (79) is resolved by reiteration of actions stored in context to yield 'Sue would make a good choice for President' following the general strategy pursued in processing VP-ellipsis, the only difference being that in the VP-ellipsis cases, the speaker provides not merely an unfixed node to be used in building up an interpretation at the ellipsis site but also a modal or auxiliary which triggers the top-down construction of partial structure for that site (see section 5.2.2) and in particular a sketch of an event term (notice that B's reply in (79) could have equivalently taken the form *or Sue would* (see Kempson et al 2015). Moreover, we anticipate mixed effects as in (80) below, notably a minor variant of (78):

(80) A: The books, I'm told are not worth us insuring.
B: The Assyrian horse . . .
A: it's obvious we must.

B reiterates the pattern of constructing an unfixed node for the bare argument to annotate, for which A follows up with what is sufficient to create its point of resolution without needing to provide the fully explicit word sequence displayed in (78). And here we can see how the existence of sequences of speaker/hearer coordinated actions is what is diagnostic of conversational dialogue, i.e., the fact that explicit linguistic forms can be omitted with jointly-derived content nevertheless assured.

Thirdly, the set of assumptions thus adopted, if correct, would lead one to anticipate that bare argument expressions decorating an unfixed node should allow enrichment of that unfixed relation indexically. And indeed this is the phenomenon we have seen with the cases of stereotypical scenarios involved in (9)-(10) earlier. Not only is the anaphoric fragment identified from the utterance scenario, but so too is the predicate to be constructed with it, enabling the initially presented term to be integrated into some larger whole. The success of the enterprise notably turns on a number of factors: the domain-generality of the vocabulary, enabling free recovery of information on a cross-modality basis, the intrinsic dynamics of the system of actions which constitute the language, and the fact that all parties to the exchange are engaged in building and developing underspecified representations of content and context. Because of this, such activity doesn't have to be a complex mind-reading affair, involving necessarily meta-representational abilities: it is available to all language users. As soon as the mechanism in however simple a form becomes available, we expect that it will be exploited in interaction through inviting some other interlocutor to contribute complementary actions to facilitate coordination. And, indeed, this very activity lies at the heart of language acquisition. Children in the so-called *one-word utterance* stage rely to a very large degree on the adult caregiver providing what they lack at that stage, namely, the requisite conceptual break-down of a holistic situational representation so that new linguistically-expressible concepts can be acquired:

(81) 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]

Expressed in DS terms, here, the child employs a bare argument in order to trigger via the initiated syntactic dependency the goal-directed context search (in the mother's memory) via the action-processing system (the grammar). This search then retrieves an eventuality of the previous day where Eliot and father had taken the boat back to a central mooring, leaving the empty space pointed to. The mother then verbalises the conceptualisation of this past eventuality, broken-down in word-sized concepts, in order to provide a learning opportunity for the child. The structure induced by such fragments is analogous to the trigger provided by predicative-ellipsis fragments. In both cases, just a single element is provided without a fixed configurational role determined, the remainder to be constructed via access to the context. The only difference is that in the case of the predicative ellipsis fragment, what is provided in addition is a constraint on the event term as encoded in the particular choice of auxiliary a complexity not manipulable by the child's emerging linguistic/conceptual resources at this very early stage.

So the pattern of interactively making use of actions previously used, hence stored in context, or anticipating actions shortly to follow, or allowing free construction from the scenario applies not only to familiar contextual phenomena such as anaphora and, by extension, to ellipsis, but also to the very mechanism that underpins long-distance dependency.

7.2.2 Clitic pronoun placement

This pattern of anaphoric, cataphoric and indexical resolutions of pronouns even turns up in sublexical morphosyntactic phenomena which, in DS, are analysed as involving *grammaticalisation* into a fixed 'macro' of a sequence of previously independent processing strategies. A classic case of such grammaticalisation concerns weak object pronouns (*clitics*) in languages like Modern Greek or Spanish. According to the DS account of clitics (Bouzouita 2008; Chatzikyriakidis & Kempson 2011; Gregoromichelaki 2013a; among others), the strict positioning and interpretational constraints of such otherwise regular pronominals can be seen as a reflex of earlier often-used structure-building strategies associated with certain (weak) pronouns so that the effects of the general strategy become routinised and stored as a chunk. A number of strategies can be reflected in this diachronic calcification, one of them, the dative, inducing the building of locally unfixed structure. The reasoning behind analysing the lexical entries of dative clitics as inducing the construction of locally-unfixed nodes is that dative clitics can express various semantic arguments, i.e., in DS terms, can lead via a fixing of that underspecified tree relation to nodes in various positions in the tree structure (e.g. direct and indirect objects, possessives, benefactives etc.). Nevertheless, such decisions must be resolved locally, hence their characterisation as inducing a locally-unfixed node. The positional determination of such locally-unfixed nodes and construal of their clitic pronoun can be achieved (see pronouns in bold): a) cataphorically, as part of the subsequent construction process; b) anaphorically, from material already in context; or c) indexically:

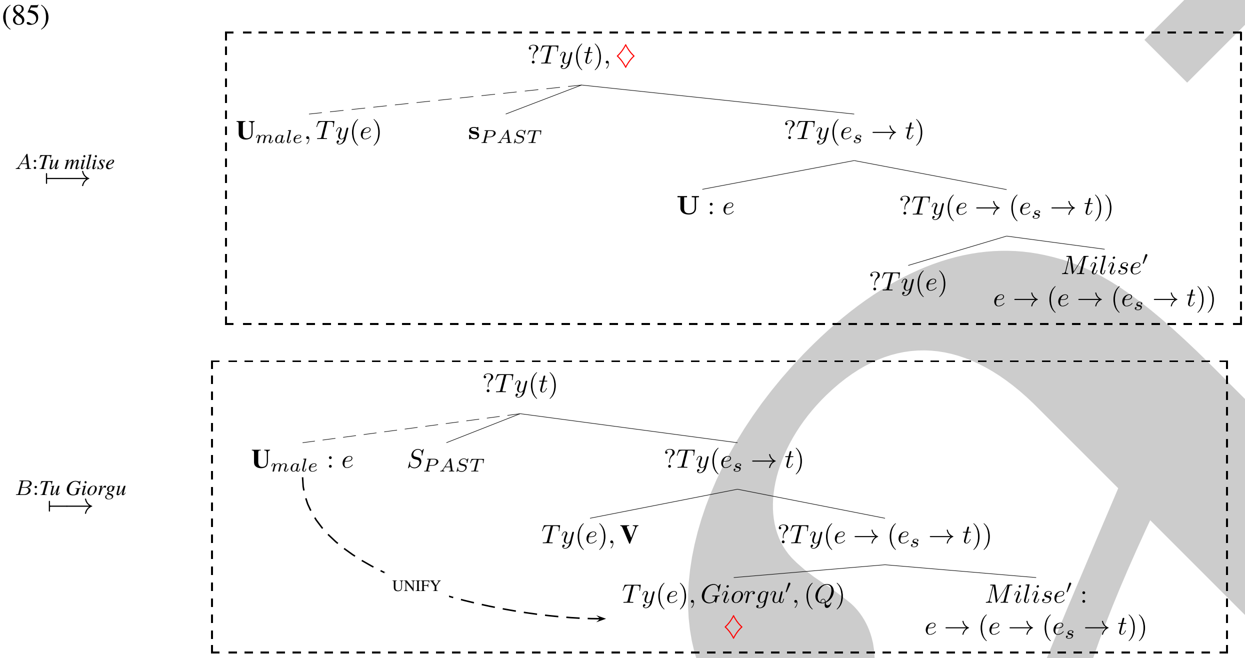
(82) *A:* **Tu** *milise* *xtes...* *B:* *Tu* *Giorgu* *(?)*
A: him.CL-DAT talked-3SG yesterday... B: the-DAT George-DAT (?)
' A: Yesterday he talked to him. B: to George (?)'

(83) *A:* *Tu* *Giorgu* *oposdipote* *B:* **Tu** *milise* *(?)*
A: the-DAT George-DAT definitely B: him.CL-DAT talked-3SG (?)
' A: To George definitely, B: He spoke to him(?)'

(84) {A seeing John talking to Mary}
A: **Tis** ... *B:* *milai* *tora* *(?),* *ne*
A: her.CL-DAT ... B: talking-to-3SG now (?), yes
'A: To her ... B: he talks to her now, yes./ Now he's talking to her? yes'

In the first case, (82), the clitic is projected onto a locally unfixed node. The verb then projects the propositional skeleton, a metavariable on the subject node and a type requirement on the object node. At this point, the second speaker takes over, providing/suggesting a formula value for the object via the word *tu Giorgu* instead of just allowing indexical provision. At this point, the unfixed node and the object node are able to unify:

(85)



In the anaphoric case, (83), the NP is first parsed on a long-distance unfixed node, followed by parsing of the clitic and verb to yield an incomplete local structure containing a locally unfixed node. These two unfixed nodes do not collapse since they are of a different kind (long-distance and local respectively). The long-distance unfixed node can unify with the object node, and the locally unfixed node with that same node too, thus providing a value for the pronominal. In the indexical case, (84), the unfixed node projected by the clitic is able to unify with the object node provided through the interlocutor's production of the predicate. The metavariable is updated to a value provided by the context, since none of the interlocutors feels the need to verbalise the provision of a value.

Confirming the unified nature of not only subsentential but also sublexical morphosyntactic dependencies, a cross-participant utterance can even involve a split affecting clitic clusters (Kempson et al 2013):

(86) *A: Irthe xtes o Giorgos ke tis... B: to edose?*
A: came yesterday the George and her.CL-DAT B: it.CL-ACC gave (?)
'A: Yesterday George came and to her ... B: he gave it to her (?)'

In such cases it is essential that the shared clitic-sequence is parsed as a unit since various positional and formal restrictions, e.g., the person restrictions associated with clusters (the *Person Case Constraint* (PCC), Chatzikyriakidis and Kempson 2011), have to be respected. According to this constraint, a first or second person singular clitic cannot co-occur with a dative clitic:

(87) *A: Irthe xtes o Giorgos ke tu... B: *se edixe?*
A: came yesterday the George and him.CL-DAT B: *you.CL-ACC showed (?)
' A: Yesterday George came and to him ... B: *he showed you to him (?)'

This is because, upon DS assumptions, the dative and first/second person accusative clitics like *se* both induce the construction of a locally unfixed node by virtue of being underspecified with respect to their structural position in the tree structure, their co-occurrence therefore being precluded as a hard constraint imposed by the tree logic, namely the fact that no more than one unfixed node of the same type is possible.[43] It is striking how, unless uses in dialogue, here the split-utterance data, are taken into account properly, i.e. as two parts of a single emergent structure instead of two elliptical independent units, there would be no way to explain the ungrammaticality of such sequences.

What we are thus seeing across this span of parallel distributions, is how the dynamics of relying on or anticipating contextual provision by one or several speakers is not an idiosyncrasy displayed solely by either anaphoric or elliptical expressions, but a wholly general phenomenon underpinning the operation of the grammar. Split-utterances are thus not "elliptical sentences" of a particular sort, but simply exploitations of a general incremental and predictive context-extension mechanism. As such the mechanics of their processing enables, and is being enabled by, all the lexical/morphological, syntactic, semantic, and contextual dependencies as they arise to channel the behaviour of participants towards joint attempts at action coordination.

8 Language as a skill for interaction: novel horizons

We opened this paper with the challenge of modelling the most familiar data of all, that of informal conversational interaction; and we have argued that conversational interaction through language can only be naturally modelled within a grammar framework which adopts the dynamics of language processing as axiomatic. Central to the DS framework which we have taken to substantiate this claim is the articulation of procedural mechanisms that induce progressive joint development of partial structures, enabling natural languages to be seen as tools for interactively establishing behavioural and mental coordination (cf Pickering & Garrod 2013).

We have chosen to illustrate this framework with an array of phenomena taken to be standard challenges relative to which grammar frameworks should be evaluated – the modelling of discontinuities, long-distance dependencies, agreement phenomena, word-order, etc., as well as addressing the novel challenge of the split-utterance phenomenon. In view of the analyses adopted, we note in closing that the shift in perspective allowing the unification of all such phenomena under a single architecture also demands the abandonment of familiar reifications of linguistic processes as independent phenomena of study as well as their grounding in abstract propositional knowledge of rules and representations. Instead, we claim, the dynamics of how people manipulate underspecified contents and partial strings in interaction with others, the environment, and even ourselves (in the form of 'inner speech'), is universal to all languages. Particular patterns that characterise individual natural languages emerge from the routinisation and storage of lexical and computational 'macros' (see also e.g. O'Grady 2013), that nevertheless remain open to modification and context-adjustment since their initial conditions of application or their output states can be narrowed down or extended depending on contextual pressures. It follows that all arguments for the "autonomy of syntax" need, in our view, to be reconsidered, in particular that syntax is independent of semantics, and sealed off ("encapsulated") from other cognitive processes.

This limitation on the remit of syntax, and the grammar more generally, has led other cognitive science researchers and philosophers into arguing for the need for additional code-like systems implementing the supposed rational inferential steps that lie outside of natural language competence specifications. These too, we believe, will need to be reconsidered. Assumptions underpinning pragmatic theorising, for example, have been grounded in the assumption that grammatical specifications, due to underspecification, have to be necessarily supplemented by higher-order reasoning involving *mutual knowledge*, *common ground*, or mind-reading (Grice 1975; Clark 1996; Sperber & Wilson 1986, cf Ginzburg 2003). Instead, under the view

[43] Under this approach, the PCC is neither about person nor about case, but rather turns out to be due to no more than one unfixed node being possible at any one time in the derivation process (assuming that dative clitics do not determine a fixed position, the otherwise licensed in Greek "abductive" fixing of the node is thus not possible). See Chatzikyriakidis (2010) and Chatzikyriakidis & Kempson (2011) for a detailed exposition of how the PCC is captured in these terms.

presented here, underspecification is not a defect displayed by SOME elements in language necessitating recourse to a qualitatively distinct and subsequent pragmatics module for its resolution: rather it is underspecification and update that are central to the grammatical architecture. A crucial means of coping with such underspecification is the generation of contextually-constrained top-down predictions during processing which result to such underspecified mechanisms being constitutive of the linguistic system's operation as a coordination device. So in effect, ALL mechanisms of the system, being goal-directed processing actions, freely interact with more general (associative or probabilistic) heuristics, memory, and cross-modal constraints in their implementation (see e.g. Eshghi et al. 2012, Eshghi et al 2013, Hough & Purver 2014, Horton & Gerrig, in press, Barr et al 2014); in consequence, linguistic elements, both morphemes/words and structures, are now seen as a means of enabling interlocutors to construct within the evolving context a shared interactional reality solely in virtue of the manipulation of these mechanisms themselves (see also Bickhard 2009).

The current view also opens up a new perspective on language acquisition and language change. Instead of analysing linguistic mechanisms under idealised competence assumptions, with performance pressures seen as essentially unrelated and external, a new view emerges here regarding both acquisition and change: As in the modelling of any other acquired skill, they can be attributed to the initial exercise of domain-general capacities involving generalisation, prediction, trial and error, and crucially relying on feedback from the social and physical environment. Such processes then lead to the gradual development and fixation of systems of specialised goal-directed action perception and execution (see papers in Cooper & Kempson 2008, and Clark 2009, Mills & Gregoromichelaki 2010, Larsson 2011, Cooper 2012, Clark & Estigarribia 2011, Mills 2014). On this view, linguistic devices must remain flexible and open-ended (i.e. underspecified) to yield efficiency both within a single instance of interaction and throughout ontogenetic and phylogenetic development. Accordingly, linguistic elements, need to be seen as offering "affordances" for co-constructing rather than encoding temporarily stable contents and structures, which are intrinsically open to adjustment according to various forms of embedding in diverse social practices and uses.[44]

In conclusion, we believe that this non-modular and processual definition of grammatical mechanisms removes one of the obstacles for integrating syntactic mechanisms within a general sensorimotor account of higher cognitive functions (see e.g. Pulvermüller 2010, Gregoromichelaki et al 2013b, Pulvermüller et al 2014). However, the necessity of defining fine-grained interactional-dynamics in order to account for the emergence of meaningful engagements among human agents, their environment, and even their own selves indicates that the boundaries of individual cognition are fluid and indeterminate. Therefore, any account of individual psychological processes will remain incomplete unless the distribution and complementarity of cognitive processes are modelled in parallel (see also e.g. Bickhard 2009, Anderson 2014).

[44]For first attempts at extending the DS model along these lines, see Hough (2015), Eshghi et al (2011), (2015), Eshghi & Lemon (2014), Purver et al (2010), Gregoromichelaki & Kempson (2015), Kempson (2016), Cann & Kempson (2016).

List of Symbols

$\langle\downarrow\rangle$ Immediately dominates; $\langle\uparrow\rangle$ Immediately dominated by; $\langle\downarrow_0\rangle$ Argument daughter; $\langle\downarrow_1\rangle$ Functor daughter; $\langle\downarrow_*\rangle$ Dominates; $\langle\uparrow_*\rangle$ Dominated by; $\langle L\rangle$ The LINK modality; $\langle L^{-1}\rangle$ Inverse LINK modality; $Ty()$ Type label; $?Ty(\mathbf{x})$ Requirement for formula of type $\mathbf{x}$; $Fo()$ Formula label; e Type of individual entity; e_s Type of situation entity; t Type of proposition; $?\exists\mathbf{x}.Fo(\mathbf{x})$ Formula requirement; $Tn(\mathbf{n})$ Treenode label with values in the language $0\{0*, 1*\}$; $?\exists\mathbf{x}.Tn(\mathbf{x})$ Treenode address requirement; $\mathbf{U}, \mathbf{V}$ Metavariables; $\mathbf{WH}$ Specialised interrogative metavariable; $\Diamond$ The 'pointer'

References

Anderson, Michael. J. 2014. *After Phrenology: Neural Reuse and the Interactive Brain*. Cambridge, MA: MIT Press.

Asher, Nicholas & Alex Lascarides. 2003. *Logics of Conversation*. Cambridge University Press. Cambridge.

Asher, Nicholas & Alex Lascarides. 2013. Strategic conversation. *Semantics and Pragmatics*. 6 (2013): 1-58.

Baker, Mark. 2008. *The Syntax of Agreement and Concord*. Cambridge University Press.

Barber, Alex. 2005. Co-extensive theories and unembedded definite descriptions. In Elugardo & Stainton (eds.) 2005, 185-201.

Barr, Dale J., Laura Jackson, & Isobel Phillips 2014. Using a voice to put a name to a face: The psycholinguistics of proper name comprehension. *Journal of Experimental Psychology: General*, 143, 404-413.

Barton, Ellen. 1990. *Nonsentential constituents*. John Benjamins Publishing.

Barton, Ellen & Ljiljiana Progovac. 2005. in Reinaldo Elugardo and Robert Stainton (eds.).

Bickhard, Mark H. 2009. The interactivist model. *Synthese*, 166.3, 547-591.

Bittner, Maria. 2001. Surface composition as bridging. *Journal of Semantics* 18, 127177.

Bittner, Maria 2007. Online Update: Temporal, modal, and de se anaphora in polysynthetic discourse. In Barker, Chris & Jacobson, Pauline (eds.) *Direct Compositionality* pp. 363-404. Oxford: Oxford University Press.

Bittner, Maria 2014a. Perspectival discourse referents for indexicals. In: SULA 7: Proceedings of the Seventh Meeting on the Semantics of Under-represented Languages in the Americas (Cornell University, May 46, 2012) (Hannah Greene, ed.)

Bittner, Maria 2014b. *Temporality: Universals and variation.* John Wiley & Sons.

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.

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

Cann, Ronnie. 2015. Indexicality and binding: construing personal pronouns in English. MS University of Edinburgh.

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

Cann, Ronnie, Tami Kaplan, & Ruth Kempson. 2005b. Grammaticality at the grammar-pragmatics interface: the case of resumptive pronouns in English. *Lingua* 115: 1551-1577.

Cann, Ronnie, Ruth Kempson & Matthew Purver. 2007. Context and well-formedness. *Research in Language and Computation*, 5.

Cann, Ronnie, & Ruth Kempson. 2008. Production pressures, syntactic change and the emergence of clitic pronouns. In Cooper, R. and Kempson, R. (eds.) (2008).

Cann, Ronnie & Ruth Kempson 2016. What do words do for us? Unpublished MS., University of Edinburgh/Kings College London.

Carstens, Vicky. (2000). Concord in minimalist theory. *Linguistic Inquiry 31:2*, pp. 319-355.

Carston, Robyn. 2002 *Thoughts and Utterances*. Blackwell. Oxford.

Chater, Nick, Stewart M. McCauley, & Morten H. Christiansen. 2015. Language as skill: Intertwining comprehension and production. *Journal of Memory and Language* dx.doi.org/10.1016/j.jml.2015.11.004.

Chatzikyriakidis, Stergios. 2010. *Clitics in four Dialects of Modern Greek: A Dynamic Account*. Ph.D King's College London.

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. 1981. *Lectures on Government and Binding*. Dordrecht, Foris.

Chomsky, Noam. 1995. Language and nature. *Mind* 104.413 (1995): 1-61.

Christiansen, Morton & Chater, Nicholas. 2015 The Now-or-never bottleneck: a fundamental constraint on language *Brain and Behavioral Sciences*, http://dx.doi.org/10.1017/S0140525X1500031X.

Clark, Herbert. 1996. *Using Language*. Cambridge: Cambridge University Press.

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

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

Cooper, Robin. 2012. Type theory and semantics in flux. In Kempson, Ruth, Tim Fernando, & Nicholas Asher, (eds.) *Handbook of the Philosophy of Linguistics*, pp. 271-323. . Elsevier.

Cooper, Robin & Ruth Kempson (eds.) 2008. *Language in Flux: Dialogue Coordination, Language Variation, Change and Evolution*. College Publications. London.

van Craenenbroeck, Jeroen. 2012 Ellipsis, identity, and accommodation. Ms. Katholieke Universiteit Leuven/Hogeschool-Universiteit Brussel.

Dalrymple, Mary, Stuart Shieber & Fernando Perreira. 1991. Ellipsis and higher-Order unification. *Linguistics and Philosophy* 4, 399-452.

Dehé, Nicole. 2014 *Parentheticals in spoken English*. Cambridge University Press.

Elugardo, Reinaldo, & Robert Stainton. (eds) 2005. *Sentential and nonsentential speech*. Dordrecht, Springer.

Eshghi, Arash. 2011. Uncommon ground: the distribution of dialogue contexts. PhD Thesis. Queen Mary, University of London.

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, & Yo Sato. 2012. Probabilistic grammar induction in an incremental semantic framework. In *Constraint Solving and Language Processing*, pp. 92-107. Springer Berlin Heidelberg.

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, & Oliver J. Lemon 2014. How domain-general can we be? Learning incremental Dialogue Systems without Dialogue Acts. In *Proceedings of SemDial 2014 (DialWatt) - The 18th Workshop on the Semantics and Pragmatics of Dialogue, At Heriot-Watt University Edinburgh*.

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.

Eshghi, Arash, & Patrick GT Healey. 2015 Collective Contexts in Conversation: Grounding by Proxy. *Cognitive science* doi: 10.1111/cogs.12225.

Fernandez, Raquel, Jonathan Ginzburg & Shalom Lappin. 2007. Classifying non-sentential utterances in

dialogue: a machine learning approach. In *Computational Linguistics*,33(3): 397-427.
Ferrara, Kathleen. 1992. The interactive achievement of a sentence: Joint productions in therapeutic discourse. In *Discourse Processes,* 15(2), 207-228.
Fiengo, Robert & Robert May. 1994. *Indices and Identity*. MIT Press: Cambridge, MA.
Fodor, Jerry. 1983. . *The Modularity of Mind*. MIT Press: Cambridge, MA.
Gargett, Andrew, Eleni Gregoromichelaki, Christine Howes, & Yo Sato, 2008. Dialogue-Grammar Correspondence in Dynamic Syntax. In *Proceedings of the 12th Workshop on the Semantics and Pragmatics of Dialogue*. (LONDIAL'08).
Ginzburg, Jonathan. 2003. Disentangling public from private meaning. In Jan van Kuppevelt & Ronnie Smith (eds.), *Current and New Directions in Discourse and Dialogue*. Dordrecht: Kluwer, 183-212.
Ginzburg, Jonathan . 2012 *The Interactive Stance: Meaning for Conversation*. Oxford: Oxford University Press.
Ginzburg, Jonathan, & Ivan Sag. 2000. *Interrogative Investigations*. Center for the Study of Language and Information, Stanford, CA
Ginzburg, Jonathan & Robin Cooper. 2004. Clarification ellipsis and the nature of updates in dialogue. *Linguistics and Philosophy* 27, 297-365.
Ginzburg, Jonathan, Raquel Fernndez, & David Schlangen. 2012. On the semantics and pragmatics of dysfluency. *Logic, Language and Meaning* Volume 7218 of the series Lecture Notes in Computer Science pp 321-330 Springer Berlin Heidelberg.
Gregoromichelaki, Eleni 2002. *The Dynamics of Relative Clauses in Modern Greek*. MSc Dissertation, Kings College London.
Gregoromichelaki, Eleni. 2006. *Conditionals in Dynamic Syntax*. Ph.D King's College London.
Gregoromichelaki, Eleni. 2011. Conditionals in Dynamic Syntax. In Kempson, et al (eds) (2011).
Gregoromichelaki, Eleni. 2012. Review of J. Ginzburg (2012) *The Interactive Stance.* Review for Folia Linguistica, Volume 47, Issue 1, 293-316.
Gregoromichelaki, Eleni. 2013a. 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*. Amsterdam: John Benjamins.
Gregoromichelaki, Eleni. 2013b. Grammar as action in language and music. In Orwin, Martin, Christine & Ruth Kempson (eds.) *Language, Music and Interaction*. London: College Publications, 91-132.
Gregoromichelaki, Eleni, to appear. Quotation in dialogue. In Saka, P., & Johnson, M. (eds.)*The Pragmatics of Quotation*. Springer.
Gregoromichelaki, Eleni. 2016. The case of non-sentential speech acts. Ms University of Osnabrueck.
Gregoromichelaki, E., Cann R. & Kempson, R. 2013a. Coordination in dialogue: subsentential speech and its implications. In Goldstein, L. (ed.) *Brevity*. Oxford University Press, 53-73.
Gregoromichelaki, E., Kempson, R, Howes, C. Eshghi, A. 2013b. On making syntax dynamic: the challenge of compound utterances and the architecture of the grammar In Ipke Wachsmuth, Jan de Ruiter, Petra Jaecks, & Stefan Kopp (eds.): *Alignment in Communication: Towards a New Theory of Communication*. John Benjamins, "Advances in Interaction Studies", 58-85.
Gregoromichelaki, Eleni, & Ruth Kempson. 2015. "Joint utterances and the (split-) turn-taking puzzle." In *Interdisciplinary Studies in Pragmatics, Culture and Society.*. Springer International Publishing, 703-743.
Gregoromichelaki, Eleni, & Ruth Kempson. 2016. "Reporting, Dialogue, and the Role of Grammar" In *Indirect Reports and Pragmatics*. Springer International Publishing, 115-150.
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.
Grice, Paul. 1975. Logic and Conversation. *Syntax and Semantics*, 3(S 41):58.
Guhe, Markus. 2007. *Incremental conceptualization for language production*. Lawrence Erlbaum Asso-

ciates.

Hausser, Roland 1989. *Computation of Language..* Berlin, New York: Springer.

Hawkins,J. A. 1994. *APerformanceTheoryofOrderandConstituency..* Cambridge: CambridgeUniversity-Press.

Hilbert, David & Paul Bernays. 1939. *Grundlagen der Mathematik I-II* Berlin, Springer.

Horton, William & Richard Gerrig, in press. Revisiting the memory-based processing approach to common ground. *Topics in Cognitive Science.*

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

Hough, Julian & Matthew Purver. 2014. Probabilistic Type Theory for Incremental Dialogue Systems. EACL workshop Type Theory and Natural Language Semantics (TTNLS), Gothenburg, April 2014.

Howes, Christine, Patrick Healey, Matthew Purver & Arash Eshghi. 2011. On incrementality in dialogue: evidence from compound contributions. In *Dialogue and Discourse 2.1.*

Jaszczolt, Kasia. 2005. *Default Semantics: Foundations of a Computational Theory of Acts of Communication.* Oxford, Oxford University Press.

Jaszczolt, Kasia M., Eleni Savva, & Michael Haugh. 2015. The Individual and the Social Path of Interpretation: The Case of Incomplete Disjunctive Questions. In: Capone, A. & Mey, J. L. (eds) *Interdisciplinary Studies in Pragmatics, Culture and Society.* Springer International Publishing, 251-283.

Kamp, Hans. 1981. A theory of rruth and semantic representation. In Groenendijk, H. Janssen, T. & Stokhof, M. (eds.), *Formal Methods in the Study of Language* , Mathematisch Centrum, Amsterdam, 1981, pp. 277322.

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

Kaplan, D. 1989. Demonstratives: An Essay on the Semantics, Logic, Metaphysics, and Epistemology of Demonstratives and Other Indexicals. In Almog, I, J. Perry, & H.Wettstein (eds). 1989. *Themes from Kaplan*. New York: Oxford University Press. (eds). 481-563.

Karttunen, L. 1969. 'Pronouns and Variables'. In R. Binnick et al. (eds.) *Papers from the Fifth Regional Meeting of the Chicago Linguistic Society*. Chicago Linguistic Society, Chicago, 1969, 108-116.

Kempen, Gerard. 2014. Prolegomena to a neurocomputational architecture for human grammatical encoding and decoding.*Neuroinformatics*, 12, 111-142.

Kempen, Gerard, Nomi Olsthoorn & Simone Sprenger. 2012. Grammatical workspace sharing during language production and language comprehension.*Language and Cognitive Processes* 27:3, 345-380.

Kempson, Ruth. Syntax as the dynamics of language understanding. In Keith Allan (ed.) *The Routledge Handbook of Linguistics*, pp. 135-152. Abingdon, New York: Routledge.

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.

Kempson, Ruth & Kiaer, Jieun. 2010. Multiple long-distance scrambling: Syntax as reflections of processing *Journal of Linguistics* 46, 127-192. doi:10.1017/S0022226709990211.

Kempson, Ruth, Andrew Gargett & Eleni Gregoromichelaki. 2007. Clarification Requests: An Incremental Account. In *Proceedings of the 11th Workshop on the Semantics and Pragmatics of Dialogue* (Decalog 2007).

Kempson, Ruth, Eleni Gregoromichelaki, & Christine Howes (eds.) 2011. *Dynamics of Lexical Interfaces*. Chicago: CSLI Press.

Kempson Ruth, Eleni Gregoromichelaki & Stergios Chatzikyriakidis. 2013. Joint utterances in Greek and English: implications for linguistic modelling. In: Stavrou, M., (ed), *Proceedings of the 33rd Annual Meeting of the Department of Linguistics of Aristotle University 2012*

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

Kobele, Gregory, 2012a. Eliding the derivation: a minimalist formalization of ellipsis. In Müller, S. (ed)

Proceedings of the HPSG 2012 Conference Chungnam National University Daejeon CSLI Publications 307-324.

Kobele, Gregory, 2012b. Ellipsis: computation of. *Wiley Interdisciplinary Reviews: Cognitive Science* 3.3: 411-418.

Kobele, Gregory, 2015. LF-copying without LF. *Lingua* 166 (2015): 236-259.

Lahiri, Aditi & Plank, Frans Phonological phrasing in Germanic: the judgement of history, confirmed through experiment. *Transactions of the Philological Society* 108, 370–398.

Larsson, Staffan. 2011 The TTR perceptron: Dynamic perceptual meanings and semantic coordination. *Proceedings of the 15th Workshop on the Semantics and Pragmatics of Dialogue (SemDial 2011 - Los Angelogue).* 140-148.

Lerner, Gene H. 2004. Collaborative turn sequences. In *Conversation analysis: Studies from the first generation*, ed. Gene H. Lerner, 225-256. John Benjamins.

Lewis, Shevaun, & Colin Phillips. 2015 Aligning grammatical theories and language processing models. Journal of Psycholinguistic Research, 44.1: 27-46.

Marr, David 1982. *Vision* New York : W. H. Freeman.

Merchant, Jason. 2005. Fragments and ellipsis. *Linguistics and Philosophy* 27.6 (2005): 661-738.

Meyer-Viol, Wilfried. 1995. *Instantial Logic*. PhD thesis. University of Utrecht.

Milkowski, Marcin. *Explaining the computational mind.* MIT Press, 2013.

Mills, Gregory J. 2014. Dialogue in joint activity: complementarity, convergence and conventionalization. *New Ideas in Psychology* 32: 158-173.

Mills, Gregory J. & Eleni Gregoromichelaki 2010. Establishing coherence in dialogue: sequentiality, intentions and negotiation. In *Proceedings of the 10th Workshop on the Semantics and Pragmatics of Dialogue, SemDial 10 (PozDial)*. Adam Mickiewicz University, Poznan, Poland, 17-24.

Milward, David. Dynamic dependency grammar. 1994. *Linguistics and Philosophy* 17.6: 561-605.

Morgan, Jerry 1973. Sentence Fragments and the Notion 'Sentence'. In *Issues in Linguistics* ed. by B. Kachru et al., 719-751. Urbana, IL: University of Illinois Press.

Morgan, Jerry 1989. *Sentence fragments revisited.* In: Graczyk, R., Wiltshire, C. (eds.), CLS 25: Parasession on Language and Context. Chicago Linguistic Society, Chicago, 228-241.

Napoli, Donna Jo. 1982. Initial material deletion in English. *Glossa* 16(1). 85-111.

Nordlinger, Rachel. 1998. *Constructive Case: Evidence from Australian Languages* CSLI Press.

OGrady, William. 2013. The illusion of language acquisition.. *Linguistic Approaches to Bilingualism* 3.3: 253-285.

Phillips Colin. 1996. *Order and structure*. PhD thesis. Massachusetts Institute of Technology.

Piccinini, Gualtiero, & Carl Craver. 2011 "Integrating psychology and neuroscience: Functional analyses as mechanism sketches." *Synthese* 183.3 (2011): 283-311.

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

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

Pullum, Geoffrey K., & Barbara C. Scholz. 2001. On the distinction between model-theoretic and generative-enumerative syntactic frameworks. *Logical aspects of computational linguistics*. Springer Berlin Heidelberg, 17-43.

Pulvermüller, F. 2010. Brain embodiment of syntax and grammar. *Brain and language*, 112.3: 167-179.

Pulvermüller, Friedermann, Rachel Moseley, Natalia Egorova, Zuabaida Shebani, & Vèronique Boulenger, 2014. Motor cognitionmotor semantics: action perception theory of cognition and communication. *Neuropsychologia*, 34, 602-619. doi: 10.1016/j.neuropsychologia.2013.12.002

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

Purver, Matthew & Jonathan Ginzburg. 2004. Clarifying noun phrase semantics. In *Journal of Semantics* 21, 283–339.

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 Semantics and Pragmatics of Dialogue 2010*, SemDial 2010 (PozDial), Poznan, Poland, 43-50.

Matthew Purver, Arash Eshghi & Julian Hough. 2011. Incremental semantic construction in a dialogue system. In *Proceedings of the 9th International Conference on Computational Semantics (IWCS).*

Matthew Purver, Julian Hough & Eleni Gregoromichelaki. 2014. Dialogue and compound contributions. In Stent, A. & Bangalore, S. (eds.), *Natural Language Generation in Interactive Systems*, pp 63-92, Cambridge: Cambridge University Press.

Ross, Haj 1967. Constraints on Variables in Syntax. Massachusetts Institute of Technology, Cambridge, MA.

Recanati, Fran cois 2010. *Truth-Conditional Pragmatics* Oxford: Clarendon Press.

Sag, Ivan. 1976. Deletion and Logical Form. Ph.D. Massachusetts Institute of Technology.

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.

Sperber, Dan & Deirdre Wilson. 1986. *Relevance: Communication and Cognition (2nd edition)*. Oxford, Blackwell.

Stabler Jr, Edward P. 1991. Avoid the pedestrians paradox. In Berwick, R.C. et al. (eds.) *Principle-Based Parsing.* Kluwer, Netherlands, 199-237.

Stabler Jr, Edward P. 2013. Two models of minimalist, incremental syntactic analysis. *Topics in Cognitive Science.* doi:10.1111/tops.12031.

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

Stalnaker, Robert. 2014. *Context and Content* Oxford: Oxford University Press.

Steedman, Mark. 1992. *Grammars and Processors* University of Pennsylvania Department of Computer and Information Science Technical Report No. MS-CIS-92-53.

Steedman, Mark. 1996. *Surface Structure and Interpretation.* Cambridge, MA: MIT Press.

Steedman, Mark. 2000. *The syntactic process.* Cambridge: MIT press.

Stokhof, Martin & van Lambalgen, Michiel. 2011. Abstraction and idealisation: The construction of modern linguistics. *Theoretical Linguistics*, 37.1-2: 1-26.

Villiers, Jill G. De, & Helen B. Tager Flusberg. 1975. Some facts one simply cannot deny. *Journal of Child Language*, 2.02: 279-286.

Weir, Andrew. 2014. *Fragments and clausal ellipsis.* PhD Diss. University of Massachusetts Amherst.

Yanofsky, Nancy 1978. NP Utterances. *Papers from the Fourteenth Regional Meeting of the Chicago Linguistic Society*, 491-502. Chicago: Chicago Linguistic Society.