Ruth Kempson, Eleni Gregoromichelaki & Stergios Chatzikyriakidis

Joint Utterances in Greek and English: Implications for Linguistic Modelling

Abstract

Joint utterances, in which individuals can take over the construction of structure from one another, are widespread in conversational dialogue, universally available, and manipulated by very young children. Such hand-overs from one speaker to another can distribute all syntactic/semantic dependencies across more than one speaker, and, as a result, are highly problematic for sentence-based grammars. In this paper we show how by the simple move of defining a grammar formalism around a core notion of incremental tree growth (Dynamic Syntax; Cann, Kempson & Marten 2005), this phenomenon is directly expressible. Syntax is defined as the articulation of mechanisms for the monotonic build-up of content relative to context; syntactic, lexical and morpho-syntactic specifications are all defined in terms of actions defining and constraining incremental interpretation; and production is presumed to symmetrically manipulate exactly the same update actions with just the addition of a subsumption check for all partial trees against some richer tree. A sketch of Modern Greek in these terms is introduced, with specifications of general and lexical actions; and the phenomenon of joint utterances is shown to emerge in consequence. The paper closes with reflections on the significance of this move.

1. Introduction: The Challenge of Conversational Dialogue

In this paper, we are advocating a grammar formalism which departs from standard assumptions in that it incorporates one essential reflection of language processing in real time: incrementality, and the twinned concepts of underspecification of structure and its subsequent update. "Syntax", on this view, is nothing more than articulation of mechanisms for the building up or realisation of representations of content relative to context, where both content and context are taken to involve incrementally evolving sequences of structured binary-branching representations in tree format.

The argument of this paper is that a shift of this sort is not a mere alternative to other types of grammar formalism, but an essential change of perspective; for, without such a change, the systemic context-dependency of language remains inexpressible at the appropriate level. The evidence we present comes from conversational dialogue. In dialogue, in sharp contrast to what standard grammars would lead us

This paper has benefited from input and data from Stavroula Tsiplakou, without whom the pleasure of working on Greek might never have got started. Thanks for her input of data, ideas, and most importantly fun over the years.

to expect, the emphasis on complete sentence units per individual speaker is very little in evidence. Instead, there is interactivity between participants, in which one interlocutor freely joins in on what the other person is saying, adding to it, modifying it, often leading the conversation in directions that might not have been intended by either party, hence the creativity and innovation that underpins social cognition.

Such phenomena cannot be taken as some peripheral dysfluency of language performance which should be ignored: they are systematically displayed in every language and they support the interactive activities in which children join in from a very young age. Yet these data have been taken not to fall within the remit of the grammar because orthodox definitions of grammatical constraints are restricted to capturing structural dependencies displayed in sentence-sized units. Accordingly, this simply precludes explanation of data associated with real-time unfolding. But, as we shall see, unless we include such data within the remit of grammar, any model will be seriously incomplete as a characterisation of the data. Moreover, with just the one shift of our methodological assumptions to a grammar formalism that reflects the online dynamics of language processing, the data can be seen as wholly unproblematic. In this paper, we survey the data from two languages, English and Greek, and then show how the modelling of such data is naturally expressible within this novel perspective on natural language syntax.

2. Split Utterances: The Challenge of Incrementality

Switching of roles between production and comprehension, across and within sentential structures, is characteristic of dialogue. People show a surprising facility to switch roles even within the setting out of some phrasal expression:

[1] Conversation from A and B, to C:
A: We 're going to
B: Bristol, where Jo lives.

[2] A: I suppose I need a a
B: mattock. For breaking up clods of earth.

[3] Daughter: Oh here dad, a good way to get those corners out
Dad: is to stick yer finger inside.
Daughter: well, that's one way. (Lerner 1991)

This to-and-fro exchange occurs throughout a conversation. New contributors can join in, often with a fragment; and each such contribution can add unproblematically to whatever partial structure has been set out so far:

[4] Hugh: We 're going to London
Alex: To see Granny.
Eliot: With the dogs?
Hugh: If you promise to keep them under control.
Eliot: Out in the garden?
Alex: Unless it pours with rain.

Such exchanges may take place at sentence boundaries, so that both antecedent and fragment can be construed as sentences, with the fragment trading on some proposition being recoverable from context; but this need not be so as [5] shows:

[5] A: I 've just got back
B: From?
A: Finland.
C: By yourself?
A: This time.

Take-overs mid-sentence may, as in [6–7], be ways of self-consciously aiming for collaborative exchange of information:

[6] A: Have you read …
B: any of your chapters? Not yet.

[7] Gardener: I shall need the fork.
Home-owner: The…
Gardener: fork. For digging the rose-bed. [BNC]

But this is by no means necessary as [8] demonstrates, where the overall proposition that results from the exchange is certainly not one that has been entertained by the initiator of the exchange at the outset, and, arguably, not by either party until relatively late in the exchange:

[8] A: What this shows is
B: that you have completely missed the point.

Example [9] shows that dependencies that can be split up can be arbitrarily complex, indeed any composite set of interlocking dependencies, seamlessly in some sense shared between the parties:

[9] A: Has every student handed in
B: their homework?
A: or even any assignments?

Here, it is not merely a pronoun to be interpreted as bound by the quantifying term uttered by A in [9] but also a negative polarity item in A's further response, equally requiring the 'affective' element provided by the initial utterance of the exchange. Moreover, that continued response by A is not one that A could have had in mind prior to B's reply as it is a response to the completion provided by B, both an extension and a departure from whatever A originally had in mind. The overall phenomenon is that in language use, participants in a discourse can take arbitrary structures, complete or partial as context, and use these as the point of departure from which to switch into speaking and so overtly contribute to the ongoing utterance. In an exactly converse pattern, a speaker can take where they have got to in producing some utterance as the background context relative to which they can understand what is then said to them. It is not a matter of going into reverse in either switch: it is simply a matter of keeping going from where one has got to.

The seamless fluency with which individuals take on or hand over utterance responsibility raises foundational issues for language modelling at a number of levels. The problem for syntacticians and semanticists is that these split points can separate sub-parts of a structure including dependencies of every sort that linguists have pinpointed, whether syntactic or semantic. In some sense, these two parts apparently have to be joined together to determine a whole syntactic/semantic structure; yet the result even so may not constitute a wellformed sentence string:

[10] (Context: with smoke coming from the kitchen)
A: Have you burnt
B: Myself? No, fortunately not.

Here, the two parts put back together yield *Have you burnt myself?* which is ungrammatical, yet they nonetheless constitute a perfectly wellformed exchange. In our view, it is clear that such exchanges between speaker and hearer are not about putting word sequences together to form, in some sense, some single *string*. Instead, they involve some speaker B taking over from the initial speaker A to provide some extension of the *content* of what was initially put forward by A. In [10] this involves B's taking up the projection of the second part of the utterance in a way that reflects themselves now as the new speaker. This demonstrates that the subsentential projection of structure has not only to represent content and the local dependency based on the identification of some subject with the use of the reflexive pronoun, but also time-linearly context-dependent content as expressed by that local anaphor. Hence we see the extent of the challenge that these data pose: the dependencies have to be defined over semantically transparent structure but, crucially, in interaction with whatever imposed context-relativity is determined by the words chosen at time points defined according to the subsentential unfolding of the utterance. Given this interdependence of syntactic, semantic and contextual factors, perhaps not surprisingly, this phenomenon has been ignored until very recently as simply a mystery for all.

It is semanticists, however, and, more specifically, those concerned with constructing dialogue models who have taken up the challenge of modelling these data, at least in part (see e.g. Ginzburg 2012). Yet even in such characterisations, the underpinning sententialist methodology remains intact, imposing the assumption that such fragments constitute nonstandard sentence-types (Fernandez & Ginzburg 2002; Ginzburg 2012; and others). This may be a justifiable move for cases for which there is a plausible sentence-based paraphrase, but there are many where there is not, of which [10] provides a particularly striking case. There are also technical problems revealed where the interjected fragment occurs so early in an utterance that there is no appropriate constituent for providing an appropriate propositional reconstruction of its interpretation:

[11] A: They X-rayed me, and took a urine sample, took a blood sample. Er, the doctor
B: Chorlton?
A: Chorlton, mhm, he examined me, erm, he, he said now they were on about a slight [shadow] on my heart. [BNC: KPY 1005–1008]

[11] is notable in this connection. Ginzburg (2012), Fernandez (2006) and others propose highly sophisticated multi-level forms of fragment resolution that follow the classical model-theoretic account of ellipsis (Dalrymple, Shieber & Pereira 1991). In that they seek to retain a syntactically simple account of the elliptical fragment itself, placing the burden of explanation on coercion operations defined on denotational contents and a fine-grained account of context evolution. More specifically, case-specific lambda-abstract forming operations are taken to apply to the established content of a previous turn to provide the requisite propositional interpretation.

This becomes possible within a "constructionist" HPSG-TTR framework enriched with rules that take into account syntactic, semantic and contextual parameters to provide constructs with which the fragment can combine to yield some encoded propositional interpretation. But for [11] the mechanisms provided are not applicable as they require completed proposition-yielding utterances. Moreover, what is required is not any abstraction over the previous context: that would predict an interpretation 'did Chorlton X-ray A, take a urine sample and blood sample from A'? Instead, the clarification concerns the entity described by the word *doctor*, to yield some token of an individual type, and we see no reason to coerce the fragment to yield some sentence/propositional type. And there remains in addition the fact that whatever abstracts are to be constructed on this model will have to involve specification of every type of syntactic dependency as part of that abstract-forming process, so the sense in which this remains a semantic characterisation for what are syntactically simple fragments becomes wholly unclear.

The data are problematic also for pragmatists. The core assumption shared by standard pragmatic models is that utterance understanding involves the hearer's grasping of the proposition which the speaker either has or could have in mind (Grice 1975; Sperber & Wilson 1995). Yet in some cases at least, it would appear that this assumption cannot be sustained and, yet, this does not threaten the success of the exchange, as witness [3], [8] and [12], in which the son is resisting all attempts to get him to be helpful and is certainly not waiting for the third of the commands even if he has bothered to process the second:

[12] (A mother, B son)
A: This afternoon first you 'll do your homework, then wash the dishes and then
B: you 'll give me 10 pounds?

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

[13] A: Are you left or
B: Right-handed.

[14] Lawyer: Will you choose your son as your attorney or
Client: My wife

To the contrary, interactive exchanges at the subsentential/subpropositional level

is what determines whatever content it is that comes to be established by either party. Moreover, these data, far from being dysfluencies which the child has to ignore in order to come to a control of wellformed sentences of their language, emerge in and support the earliest stages of language learning:

[15] Carer: Old McDonald had a farm… On that farm he had a
Child: cow.

And, as with adults, the need to recognise the other party's intended content seems a heavy burden for almost all carer-child exchanges. Indeed, in exchanges with young children, carers may rely on the child's ability for interaction even though possibly fully aware that her audience is not yet at a stage of development in which they can be expected to have any capacity for recognition of higher-order intentions:

[16] A (teacher to each child in turn in the class): And your name is …
B (child): Mary

In [16] the child is merely completing the template set out by the carer with what is answer to a question as well as the completion of a statement, which they can do without any need to identify a given thought on the part of the questioner. If these data provide evidence that recognition of the content of other people's intentions is not a necessary condition on successful acts of communication, as we suggest they do, then foundational assumptions of Gricean pragmatics need to be reconsidered (Gregoromichelaki et al., 2011).

3. Split Utterances in Greek

Contrary to parametric and "constructionist" explanations, such data are not specific to single languages or some epiphenomenon confined to properties of the English lexicon. Within the confines of their syntactic properties, all languages display the same array of data to a very high degree of parallelism. Interruptions for clarificatory or other purposes can take place mid-constituent in Greek just as in English:

[17] Clarification: two gardeners

A:	Thelo to…	'I want the…'
B:	skalistiri, gia na spasoume tus svolus	'mattock. To break up the clods'

Narratives shared between speakers can be indefinitely extended:

[18] Extending narrative

A:	molis girisa	'I am just back'
B:	apo…	'from'
A:	ti Filandia	'the Finland'
B:	me to filo su?	'with your friend?'
A:	Fetos oxi. Isos tu xronu.	'Not this year. Maybe next year.'

[19] Anaphoric dependency and conditional split (Context: visiting grandmother)
Child: Giati na pao? Thelo na katso spiti. 'Why should I go? I want to stay home.'

Mother:	Monos su me ta skilia?	'On your own, just you and the dogs?'
	Afu otan pame eki pernas kala	'But when we go there you have a good time.'
Child:	An dulevi i tileorasi	'If the TV works'

As [19] provides preliminary indication of, there is no necessity for conversational partners to be in agreement or take a cooperative stance with respect to each other's stock of assumptions, as such interactivity can persist across fierce disagreement without any threat to the conversational interaction itself:

[20]	A: ki afto dixni	'and this shows'
	B: oti ise ilithios	'that you are an idiot'

And whatever multiple speech-act communication a split exchange may constitute, it can be participated in by young children just as freely as between adults:

[21]	Teacher to Ana:	se lene	'they call you'
	Ana to Teacher:	Ana	'anna'
	Teacher to Nikos:	ki esena	'And you'
	Nikos to Teacher:	Niko	'Nick'

Just as in English, this ability at switching between one party and another is no respecter of linguistic categorisations as to structure. In Greek, where a reflexive anaphor is used, a switch of participants roles can unproblematically apparently distribute the dependency across more than one speaker:

[22]	A:	Exo ta xalia mu.	'I feel rotten'
	B:	Giati den exis ebistosini…	'because you don't have faith…'
	A:	Ston eafto mu? oxi, de ine afto.	'in myself? No, that's not it'

[23] Reflexive and quantifier binding

A:	Tha doso afksisi…	'I will give a raise…'
B:	Ston eafto su ke ta filarakia su?	'To yourself and your friends?'
A:	Oxi, eksipnakia, s' esena, sti Maria, se olus	'No, clever clogs, to you, to Mary, to everyone.'

Moreover, there is no limit on the type or combination of dependencies which may be so split:

[24] Quantifier binding and NPI:

A:	Paredosan oli i fitites	'Have all the students handed in'
B:	Tis ergasies tus?	'their term-papers?'
C:	I esto kapies/tipota askisis.	'Or even any assignments?'

In Greek, there is of course a rich agreement system, and in all such split fragments, the take-over by a second party must involve suitable selection of agreement specification:

[25] Morpho-syntactic gender split

A (female):	Tin proigumeni vdomada imuna poli…	'Last week I was very…'
B (male):	arosti? ill.FEM	

Under certain circumstances, there may even be distributed switch of speakers across a clitic cluster:

[26] Completing a clitic cluster:

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

Finally, as in English, such split fragments may do double duty as completing a question and providing an answer, giving rise to multiple indications of speech acts, even by a single constituent fragment (see also [25–26])

[27] A: Klotsas me to aristero su i me 'Do you kick with your left or with'
B: to deksi mu (podi) 'my right (foot)'

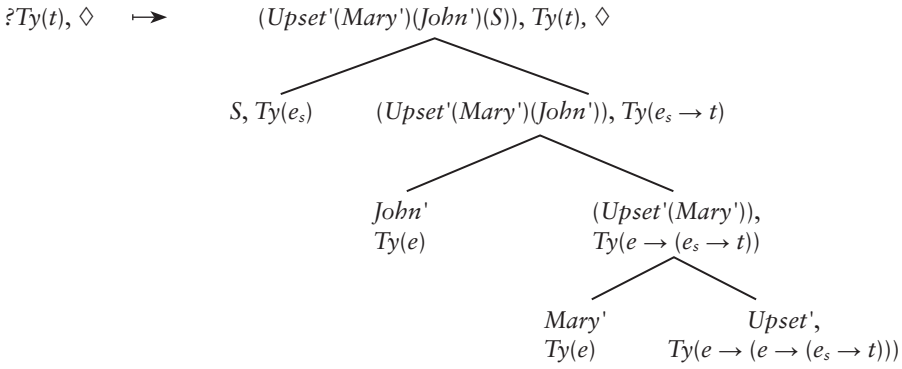
In short, the dynamics of split utterance phenomena runs in strikingly parallel fashion right across different languages, and different language families. The dynamics of interaction achievable by individual languages is constant across them all. This is not a language-individual phenomenon. Quite generally, humans in interaction take the structures so far constructed for verbalisation or understood in parsing as the background to the next processing task, whether this be one of production or comprehension. Moreover, this phenomenon affects all linguistically-determined dependencies. Hence, howsoever the phenomenon is to be characterised, these data *have* to fall within the core remit of grammar. If these data remain excluded from grammar-internal specification, *not a single* phenomenon will receive complete characterisation within the grammar unless an explosion of deletion/reconstruction/coercion operations is defined, in our view, an indefensible stance. What we now turn to is a perspective in which these data become naturally modellable, indeed a predictable consequence of the framework, without either any additional constraints on performance external to the framework or framework-internal construction-specific stipulations.

4. Dynamic Syntax

Dynamic Syntax (DS) is an action-based grammar formalism of which the core notion is structural development of representations relative to context. With the dynamics of structural growth built into the core grammar, natural-language syntax is seen as a set of principles for developing the mapping of such structures to strings of words and vice-versa. The syntactic mechanisms, that is, are procedures that define *how* parts of representations of content can be incrementally introduced/produced or updated with all such development being relative to context. Accordingly, context is just as structural and dynamic as the concept of content with which it is twinned, constituting a record not merely of the (partial) structures built up, but also the procedures used in constructing them.

The general process of parsing is taken to involve building as output a tree whose nodes reflect the content of some uttered formula – in the simple case of a sentence uttered in isolation, a complete propositional formula. The input to this task is a tree that does nothing more than state, at the root node, the goal of the interpretation process to be achieved, namely, to establish some propositional formula. For example, in the parse of the string *John upset Mary*, below in [28], the output tree to the right of the arrow, constitutes some final end result: it is a tree in which the propositional formula itself annotates the top node, and its various sub-terms appear on the dominated nodes in that tree rather like a proof tree in which all the nodes are labelled with a formula and a type. The input to that process is an initial one-node tree which states the goal of the interpretation process which is to achieve a formula of appropriate propositional type (hence the *requirement ?Ty(t)*, the*?* indicating that this is a goal not yet achieved, the *S* representing the final Davidsonian event/situation argument of type e_s:[1]

[28] Parsing: John upset Mary.



These DS trees are invariably binary, reflecting functor-argument structure, and, by convention, the argument always appears on the left branch, and the functor on the right branch (a *pointer*, ◊, identifies the node under development). Each node in a complete tree is annotated not with words, but with *formulae*, i.e., terms of a logical language, these being subterms of the resulting propositional representation. In order to simplify presentation, the formula values are displayed here as though in English, represented here as *Mary'* and so on, but it must be kept in mind that the actual representations consist of concepts constructed through instructions encoded in the words of the language.[2] The parsing task is to use both lexical input

[1] Details suppressed for simplicity· see Cann (2011).

[2] This convention of using words to display concept annotations will be sustained throughout the paper, as in Greek tense the morphology is not reliably syncretic and morphological forms may map in a complex way needing lexical specification, item by item. This is not a problem given the mapping onto a logical structure that is not inhabited by the word forms themselves, but it is a problem we largely suppress in this paper. We also suppress all issues of quantification and tense specification, taken to project an additional argument (event) term in each case; see Cann (2011), for a detailed account of

and information from context to progressively enrich the input tree to yield such a complete output following general tree-growth actions.

4.1 Formal Properties of Trees

Though the annotations on trees may be determined idiosyncratically and in part through lexical itemisation, the trees themselves and the mechanisms for their growth are language-general, and so to be defined as part of the formal meta-vocabulary for grammar writing. In order to enable explicit formulation of how such structures grow, trees themselves cannot be taken as primitive and have to be formally defined, together with a vocabulary for describing actions that induce the requisite tree growth. Following Blackburn & Meyer-Viol (1994), DS adopts a (modal) logic with two basic modalities. Firstly, there is the modal operator $\langle\downarrow\rangle$: $\langle\downarrow\rangle\alpha$ holds at a node if α holds at its daughter (with variants $\langle\downarrow_0\rangle$ and $\langle\downarrow_1\rangle$ for argument and functor daughters respectively). There is also its inverse $\langle\uparrow\rangle\ \alpha$ which holds at a node if α holds at its mother, equally with argument and functor variants indicative of the status of the daughter relation so identified. There are also LINK ($\langle L\rangle$) relations between trees, with their inverse $\langle L^{-1}\rangle$.

The core pair of concepts driving forward the tree-growth process is that of underspecification with an attendant *requirement for update*: both are essential in getting appropriate reflection of the time-linearity involved in building up (partial) trees in stages. There are different types of underspecification: underspecification of some putatively final tree, hence underspecification of formula content or type, underspecification of type of node and even of the relation of that node to others in the tree. For every node, in every tree, all aspects of underspecification are twinned with a concept of *requirement*, *?X*, for some annotation *X*; and these are constraints on how the subsequent parsing steps must progress. Such requirements apply to all types of annotation: there are type requirements, *?Ty(t)*, *?Ty(e)*, *?Ty(e → t)* etc.; tree-node requirements, $?\exists x.Tn(x)$ (associated with underspecified tree-relations in need of update); formula requirements $?\exists x.Fo(x)$ for any expression which, though typed, is only a placeholder for the content to be fixed independently of its lexical projection; and modal requirements expressing future developments, for example, $?\langle\uparrow_0\rangle$ *Ty(e → t)* captures accusative case-marking in the form of an output requirement that a node so annotated be immediately dominated by a node of predicate type. In each case, these requirements drive the subsequent tree-construction process: unless they are eventually satisfied, the parse will be unsuccessful.

4.2 Growth of Trees

Actions for tree growth are then defined as actions for building such tree relations, involving a procedural vocabulary with sequences of actions consisting of, e.g.,

English tense and auxiliary specifications for in depth illustration of how lexical specifications may interact to determine composite effects.

make(X), go(X) and put(Y) operations, where X and Y are tree relations and node-annotations (labels) respectively.

Computational Actions constitute generally available strategies for tree-growth; they play a role in determining movement of the pointer, ◊, around some partial tree under construction; they are also in charge of removing requirements once they are satisfied, and performing beta-reduction operations when possible. The lexicon and lexicon-internal specifications are where language-specific generalisations are expressed.

Lexical Actions are actions associated with words, and conditional in form. Given a certain trigger, these induce an unfolding macro of tree-growth actions that lead to some tree update from the partial tree containing that trigger. For example, from a requirement of the form *?Ty(t)* as the triggering condition, verbs project a skeletal propositional template projecting a node for a predicate and attendant arguments as determined by the adicity and typing of the predicate, which are annotated as part of the action sequence with either the requirement *?Ty(e)* or with a typed place-holding annotation exactly in the manner of anaphoric expressions. And this is where there is an array of options available across different languages.

Argument nodes projected by English verbs are all associated solely with a requirement of the form *?Ty(e)*, thus ensuring that in each case, there has to be some further step of language-input processing in order to satisfy type requirement and project some concept formula of that type. In a language such as Greek, on the other hand, with its mixed subject pro-drop object non-pro-drop specifications, there is asymmetric specification of the various argument nodes. The object argument is defined in like manner to English as the weaker form of annotation needing further linguistic input, annotated with the requirement *?Ty(e)*. The subject argument, however, is specified as quasi-pronominal, with a type-specific placeholder, a meta-variable indicated as **U** with requirement solely for a formula value, its type already being determined. And the event term argument too is specified as a placeholder for both verbal, auxiliary and adjunct specifications to add to its articulation of the appropriate event term. The lexical specification of the verb induces the entire propositional template and such attendant annotations through provided actions of make(X), go(X) and put(Y):

[29] ksero 'I-know'

IF	*?Ty*(*t*)
THEN	make ($\langle\downarrow_0\rangle$) : go($\langle\downarrow_0\rangle$);
	put(*Ty*(e_s), *Fo*(U), + *PRES*, ?∃*x*.*Fo*(*x*)); go($\langle\uparrow_0\rangle$)
	make($\langle\downarrow_1\rangle$); go($\langle\downarrow_1\rangle$); put(?*Ty*($e \rightarrow (e_s \rightarrow t)$));
	make($\langle\downarrow_1\rangle$); go($\langle\downarrow_1\rangle$);
	put(*Fo*(ksero'), *Ty*($e \rightarrow (e_s \rightarrow t)$))
	go($\langle\uparrow_1\rangle$); make($\langle\downarrow_0\rangle$); go($\langle\downarrow_0\rangle$);
	put(?*Ty*(*e*); go($\langle\uparrow_0\rangle$); go($\langle\uparrow_1\rangle$); go($\langle\uparrow_1\rangle$));
	make($\langle\downarrow_0\rangle$); go($\langle\downarrow_0\rangle$);
	put(*Fo*(V), Ty(e), ?∃x.Fo(x))
ELSE	Abort

In the set of actions above, first an event term is taken to be projected by the form *ksero*.[3] Then the predicate-internal structure is constructed along with the concept associated with *ksero*, with its two attendant arguments plus that of the event term. As this specification sets out, the difference between so-called pro-drop specifications and non-pro-drop specifications turns on the number of arguments which the verb itself projects. It is no coincidence on this view that the verbal morphology in many languages gives every appearance of projecting a verb plus pronominal arguments (albeit in reduced form). On the DS view of the matter, this is precisely the formal specification – the projection by the verb of metavariable annotations to such argument nodes.

Anaphoric expressions themselves can be of various types, also projecting a metavariable content as a place-holder licensing either indexical construal or identification within the construction process: the notation involves defining pronouns as projecting a formula metavariable (*Fo(**U**)*) with the accompanying requirement for formula update, *?∃**x**.Fo(**x**)*. It is also possible to express what it means for a word to annotate a terminal node: the lexical specifications associated with such expressions impose a requirement that the node in question cannot be further expanded, a restriction which strong forms of pronoun impose along with other words. Agreement specifications do not impose any such restriction, thereby giving rise to further development, with clitic pronouns lexically varying as to whether they are more like word or agreement forms of specification.[4]

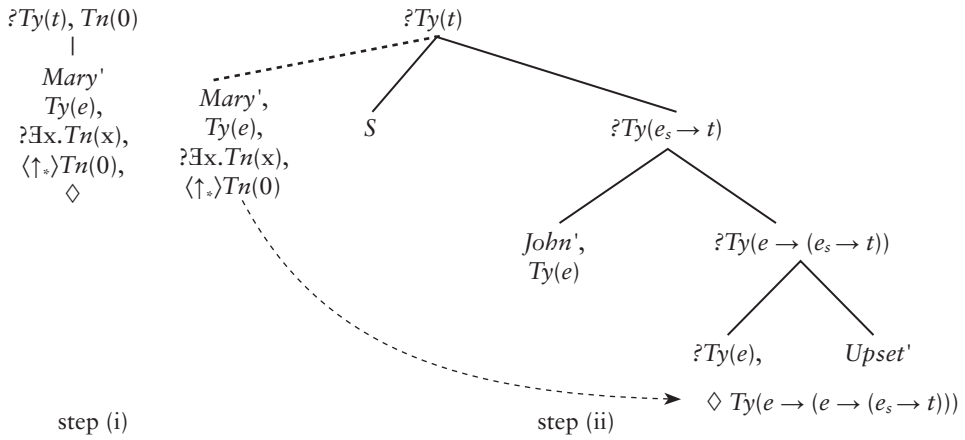
With the concept of underspecification and update being the core notion on which the grammar specification turns, there is a structure-building analogue of the content underspecification, familiar from anaphoric specifications; and amongst the computational actions there are processes inducing underspecified structural relations, local and non-local. These weak structural relations are defined using Kleene star operators: all such weak tree relations have an associated requirement for future provision of a fixed tree relation, i.e., a fixed tree node address: *? ∃**x**.Tn(**x**)*. For example, $\langle\uparrow^*\rangle$ *Tn(a)* is defined as an annotation on a node indicating that there is at least one future development in which the node with address *a* bears a sequence of mother relations to the present node. This relatively weak tree relation is taken to express long-distance dependency effects in terms of structural underspecification and update. For example, when first processing the word *Mary* in [30] below, which is initially construed as providing a term whose role isn't yet identified, the parse is taken to involve the application of a computational action which introduces from the initial root node annotated with *?Ty(t)*, a relation to that top node which is *underspecified* at this juncture, identifiable solely as dominated by the top node (indicated by *Tn(0)*), and requiring type *e*, specified by a *?Ty(e)* annotation. This enables

[3] Here we suppress all details, giving only the stipulation *+PRES* as a promissory note for a properly formulated account; see Cann (2011), for a detailed exposition of English auxiliary, modal and verbal specifications to yield appropriate results.

[4] For simplicity these restrictions are omitted here.

the expression *Mary* to be taken to annotate this node: this is step (i) of [30]. The accompanying requirement for a fixed tree node position $?\exists x.\ Tn(x)$ induces the update to this relatively weak tree-relation which in this derivation takes place after processing the subject plus verb sequence to yield the two-place predicate structure in step (ii) of [30].[5] Provision of a formula annotation for the object argument node and update of the unfixed node initially introduced is given by the *unification* step indicated there, an action which satisfies the update requirements of both depicted nodes.

[30] Mary, John upset.



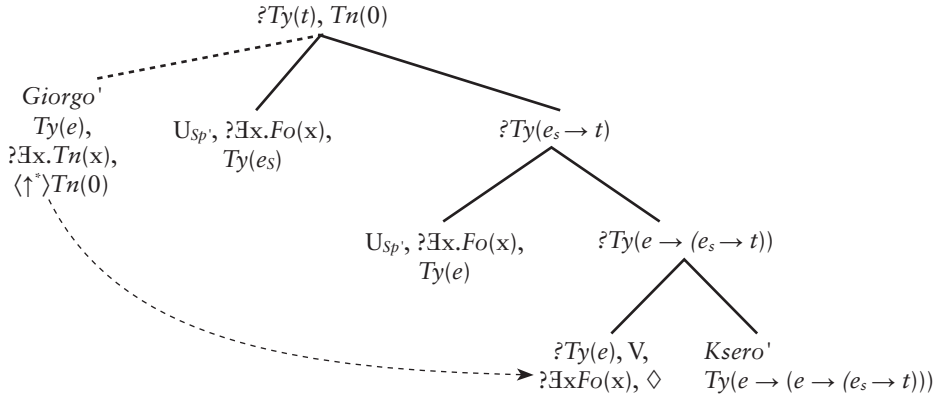
Such a basis for expressing long-distance dependency is available in Greek, either with or without an attendant clitic pronoun:

[31]	Ton	Giorgo	(ton)	ksero.	'George, *I* know.'
	the.ACC	George	(him.CL-ACC)	know.1SG	

It is notable that no distinct derivation is needed in the presence of a clitic, because the specification of a pronoun as projecting a placeholder for a term of type *e* is fully commensurate with a step of unification which secures the addition to that object node of the feature specifications, which up to that point have been associated with node whose relation to the remainder of the tree remains unspecified. It is through such a step of unification that both node specifications are enriched, exactly as in English:

[5] Accounting for left-peripheral expressions in these dynamic update terms is not contentious as a parsing strategy: what is innovative is its application within the grammar-mechanism as the basic underpinning to syntactic generalisations.

[32] To Giorgo (ton) ksero



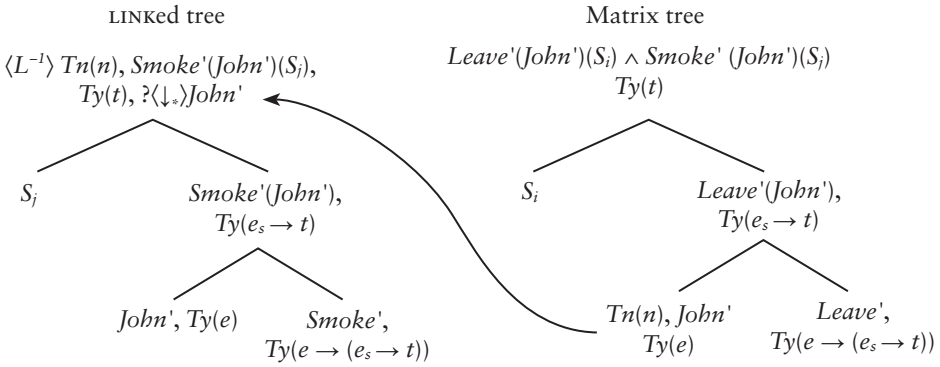
As regards dialogue, the tight interleaving of parsing and production is essential. Production in DS is presumed to follow the very same process as parsing, with one further assumption: at every step in production, there must be some richer tree, a so-called *goal tree*, which the tree under construction must *subsume*, in the sense of being able to be developed into that goal tree by rules of the system (Purver & Kempson 2004). So parsers and producers alike use strategies for building up representations of content in association with predictive mechanisms in the form of the generation of requirements that constrain the process.[6] The only difference is a direct reflection of the intuition that in production, the speaker must have some richer concept in mind, not necessarily fully propositionally-complete. It is this close parallelism of parsing and production that will stand us in good stead when it comes to the modelling of split utterances.

4.3 Compounding Trees through Term-Sharing

All that is needed to round out this DS account to achieve a basis for reflecting the full array of compound structures is a licence to build paired trees, so-called LINKed trees, with one (possibly partial tree) constituting the context constructed for the relative construction of the other. Such trees are linked together solely by the sharing of terms, established for example by encoded anaphoric devices such as relative pronouns. This is an encoded context-construction device. Consider the structure derived by processing the string *John, who smokes, left* (omitting details of tense specification):

[6] This view is commensurate with that of Pickering & Garrod (forthcoming), but does not involve any grammar-external superimposed performance constraints.

[33] John, who smokes, left.



The arrow linking the two trees depicts the so-called LINK relation. The tree whose node is pointed by the arrow is the LINKed tree (read ⟨L^{-1}⟩ as 'linked to'), semantically, a tree providing contextual grounding for some aspect of the main structure. Within the process of constructing any one such LINKed tree, the full range of computational, lexical and pragmatic actions is available; and with this flexibility to allow the incremental projection of arbitrarily rich compound structures, the result is a formal system combining lexical, structural and semantic specifications, all as constraints on the growth of trees, with both context and content structure buildable in tandem. This concept of LINKed structures is definable type-generally, and has very general application, across Hanging Topic Left Dislocation structures, clausal and phrasal adjuncts, and so on.[7] This is particularly pertinent to Greek, where the combination of left-peripheral expression and clitic doubling provides a rich array of apparent argument-duplication effects, which within the DS formulation of structure projection is unproblematic. Nodes may freely allow annotation from more than one lexical input, as long as these meet the monotonicity constraint that all tree growth within an individual derivation involves progressive enrichment of the emergent tree.

Conversely, since the grammar framework articulates a set of strategies for tree development, there may be more than one sequence of actions that can lead from the output requirement to some wellformed end result. *Clitic doubling* is a case in point, as in [31]. The left-peripheral object case-marked expression may well be taken to annotate an initial object argument that is initially unidentified as to its position within the configuration, an underdeterminacy that is resolved first with the tree growth achieved by the clitic and its annotation and then, once there is in hand the template structure projected by the verb, by unifying the updated object node with the object-requiring node established within that template. But equally well, that left-peripheral object-marked expression may be taken to annotate an initially projected LINKed tree structure containing just such a node, which then imposes on

[7] The advantage of this articulation of compound structures as independent LINKed trees is that it provides a natural basis for expressing the so-called "syntactic island" constraints.

the emergent tree structure under construction a requirement that some copy of the term initially constructed be created in the development of the main tree, a requirement which the clitic can then duly satisfy (see Gregoromichelaki 2013, for detailed examination of left- and right-dislocation effects in Greek from a DS perspective). There is indeed in principle no uniquely determined sequence of derivations per string-interpretation pair: the grammar is an articulation of mechanisms of processing which can apply in a number of ways to determine an appropriate string-interpretation pairings relative to context, as one would be expecting of an optimally flexible parsing-directed system.

4.4 Concepts of Locality

The advantage of the DS framework is that with all lexical and computational actions defined in the same tree-growth terms, full interaction between them is predicted. It is straightforward, for example, to identify the requisite concepts of locality constraining antecedenthood for both reflexives and pronouns, enabling a natural feeding relation of such specifications into computational actions of various sorts. Arguments local to a given predicate can all be defined as meeting the characterisation $\langle\uparrow_0\rangle\ \langle\uparrow_1^*\rangle$ *Tn(a)* (*Tn (a)* being up one argument-relation plus a possibly empty sequence of function-path relations from the node so characterised). Accordingly, reflexive anaphors can be characterised as projecting the action specified in [34]:

[34] Actions associated with reflexive pronouns:

IF	?*Ty*(*e*), *Fo*(α)	
THEN	IF	$\langle\uparrow_0\rangle\langle\uparrow_*^1\rangle\langle\downarrow_0\rangle$ *Fo*(α)
	THEN	put(*Fo*(α), *Ty*(*e*))
	ELSE	Abort
ELSE	ABORT	

And, conversely, the substitution process of non-reflexive pronominals excludes as antecedent any formula annotating a node standing in such a local relation.

In sum, the familiar array of syntactic phenomena – discontinuity in syntactic dependency, binding specifications for anaphoric construals, agreement constraints on such construals, and so on – can all be expressed by adopting the intuitive step of shifting perspective to allow grammar formalisms articulating concepts of real-time growth of semantic representation. Such syntactic dependencies can all be redefined as resulting from mechanisms licensing various forms of underspecification whose update processes then interact with both general and idiosyncratically imposed constraints to incrementally determine an array of emergent partial structures reflecting content more or less transparently. All that is then needed to express arbitrarily complex structures is the assumption that the local predicate argument structures so projected can be extended to form composite pairs of local structural domains through mechanisms of variable sharing; and the result is a rich basis from which to set out novel solutions to syntactic puzzles.

5. Split Utterances across Languages

The striking bonus for the DS perspective is that the phenomenon of split utterances is directly expressible. Generation, recall, makes use of exactly the same tree growth mechanisms as parsing, and with exactly the same commitment to incrementality in association with the predictivity afforded by the imposition of requirements. This means that in all exchanges both parties will be building up a tree to serve either as the basis for the parsing process or as the basis for the production process.

From the perspective of modelling the point of switch of roles, two properties of the generation mechanism are pertinent. First, there is nothing to prevent speakers initially having only a partial structure to convey, i.e., the goal tree may be a *partial* tree: this is unproblematic, as all that is required by the formalism is monotonicity of tree growth, and the subsumption check is equally well defined over partial trees. Indeed, as we have seen, the system is set up to project partial trees as context, hence as input to some next point of departure, through the encoding of transitions across from a node in one tree to initiate the construction of an adjunct LINKed tree with the carry-over of some content formula from that first partial structure as a requirement on the unfolding of the LINKed tree. Second, the goal tree may change during generation of an utterance, as long as this change involves monotonic extension.

Such a model under which the speaker and hearer essentially follow the same sets of actions, each incrementally updating their semantic representations, allows the hearer to mirror the same series of partial trees as the producer, albeit not knowing in advance the content of the unspecified nodes. For example, for dialogues such as [35] below, Mary as the speaker reaches a partial tree of what she has uttered through successive updates, while Bob initially as hearer, follows the same updates to reach the same representation of what he has heard: they both apply the same tree-construction mechanism which is none other than their effectively shared grammar.[8] This provides Bob with the ability at any stage to become the speaker, interrupting to take over Mary's utterance, repair, ask for clarification, reformulate, or provide a correction, as and when necessary. What then looks like distributed construction of some structure across distinct participants is in fact no more than switch within the tree-development process from parsing without any goal tree for subsumption purposes to the very same tree-construction process, except that after the switch into production there is a goal tree providing the basis for the further subsumption check. But the dependencies thereby apparently so distributed remain constructed and resolved within a single locally-definable tree domain for an individual agent. So, according to DS assumptions, repeating or extending a constituent of someone else's utterance is licensed only if the hearer, now newly turned speaker, entertains a message to be conveyed (a new *goal tree*) that matches or extends in a monotonic

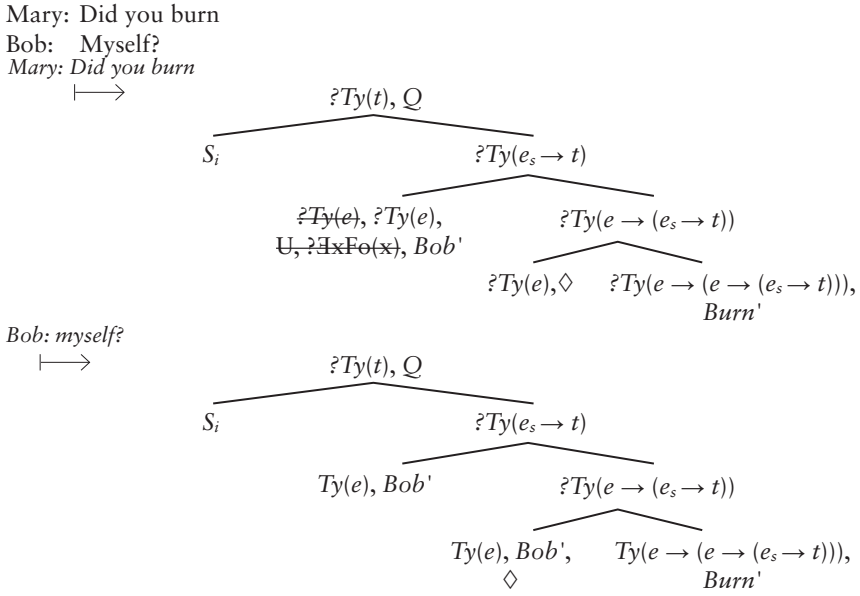
[8] A completely identical grammar is, of course, an idealisation but one that is harmless for current purposes.

fashion the parse tree of what they have just heard. This formal device matches the intuition that in taking over as producer, the speaker must have something, however partial, which they wish next to convey.[9] By way of illustration, we take a simplified variant of [10]:

[35] Mary: Did you burn
Bob: myself? No.

Here, despite the fact that the string **Did you burn myself?* is unacceptable, under DS assumptions, with representations only of informational content, not of structure over strings of words, the switch of person is straightforward and leads to a wellformed result. [36] displays the partial tree induced by parsing Mary's utterance, *did you burn*, which involves a substitution of the metavariable projected by *you* with the name of the interlocutor/parser.[10] At this point, Bob can complete the utterance with the reflexive, for what such an expression does, by definition, is copy a formula from a local co-argument node onto the current node, just in case that formula satisfies the conditions set by the person and number of the uttered reflexive. These are encoded in the lexical actions induced by *myself*. This will have the effect of picking out the *current* speaker (cf. [34]):

[36] Mary: Did you burn
Bob: Myself?



[9] Reflecting such freedom with respect to how much one might have to add, this message (tree) may of course be partial, as in [11], where, very early in the utterance response made by B, a request for clarification is made by adjunction of a nominal constituent to a term just processed of the same type, in order to get such early clarification. Within DS, appositional adjunction is modelled by the construction of a LINKed structure of same type as the input node (Gregoromichelaki et al. 2011).

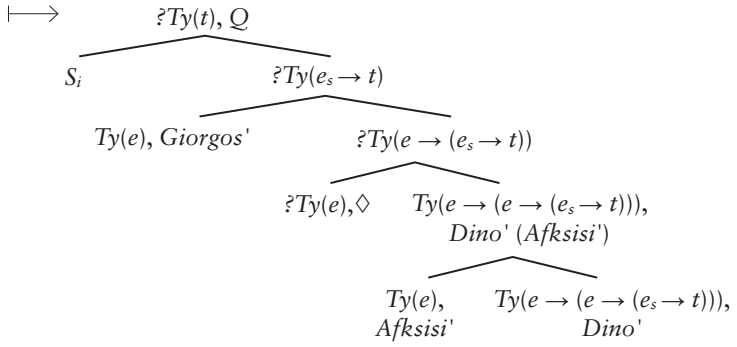
[10] The feature *Q* on a annotated node is not taken to have a fixed speech-act content: given the range of acts achievable by interrogative structures we take interrogative forms to encode a direction by the speaker to the hearer for a particular type of coordination, here notated simply as *Q*.

This illustration is only of the simplest type of split-utterance ellipsis, but the point is entirely general. These seamlessly achieved split utterances can apparently separate off any expression from the syntactic environment it needs for its well-formedness because both speaker and hearer make use of the same mechanisms. So one and the same individual, whether as speaker or as hearer, will invariably have a partial structure on which to rely at the point of participant switch.

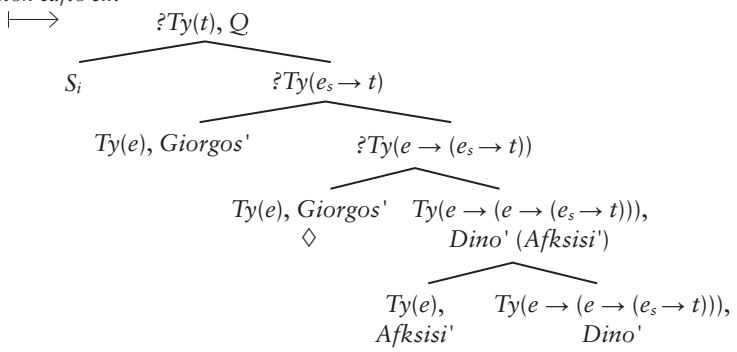
The same principle applies in exactly the same way to Greek, even though the data involve gender mismatching and clitic splitting, neither of which occur in English. Thus similar ideas can apply to cases where the reflexive, which would violate principle A if it weren't part of a split utterance, is any kind of argument or adjunct. Thus, the Greek case in [23] can be dealt with using the very same reasoning as for the English case. The reflexive in [23] takes its value from the speaker of the second turn in the split utterance. If we assume the split utterance to be uttered throughout from the first speaker's perspective, then the second person reflexive *ston eafto su* would render the sentence ungrammatical. But given the split utterance, the second speaker has now become the hearer, hence the second person form of the reflexive:

[37] Giorgos: Tha doso afksisi 'I will give a-raise'
Giannis: Ston eafto su 'To yourself?

Giorgos: Tha doso afksisi



Giannis: Ston eafto su?

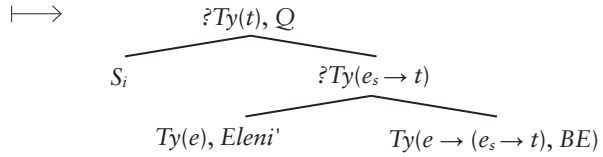


Gender mismatches found in cases like [25] are also explained via the same rea-

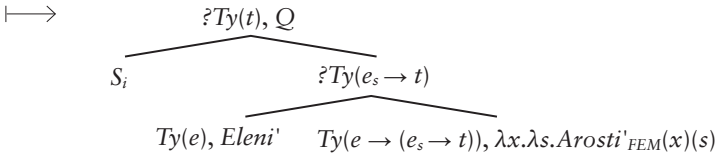
soning, though to different effect. The second speaker in [25] is of male gender; however, he uses the morphologically feminine form of the adjective. If the whole split utterance was uttered as a single utterance by the second speaker, it would be ungrammatical due to the gender mismatch. Yet, under our analysis, such a mismatch does not arise, given that the second speaker's context is based on the context of the first speaker. Given that the speaker is female and the predication is about the first speaker, the form used should be marked as female:[11]

[38] Eleni: Imuna poli … 'I was very'
Stergios: arosti 'Ill?'

Eleni: Imuna poli



Stergios: arosti



Lastly, in [26], the second speaker completes the sentence in the middle of a clitic cluster. One of the things that this possibility might suggest is that clitic clusters are indeed full words.[12] Assuming that clitics are affixes and as such are stored as one lexical entry along with the verb, we would expect that such completions would not be possible, contrary to fact. In the same vein, such examples show that in Greek some of the clitic clusters are not real clusters in the sense that they involve the lexical entries of two individual clitics and not one lexical entry for the whole clitic cluster. This is the stance advocated by Chatzikyriakidis (2010) at least for Standard Modern Greek.[13] In the case in question, the first speaker provides the first clitic. This is a dative clitic, which following DS assumptions (Kempson & Cann 2008; Bouzouita 2008; Chatzikyriakidis 2010; Chatzikyriakidis & Kempson 2011) projects a locally unfixed node. This is the context where the second speaker takes over. The

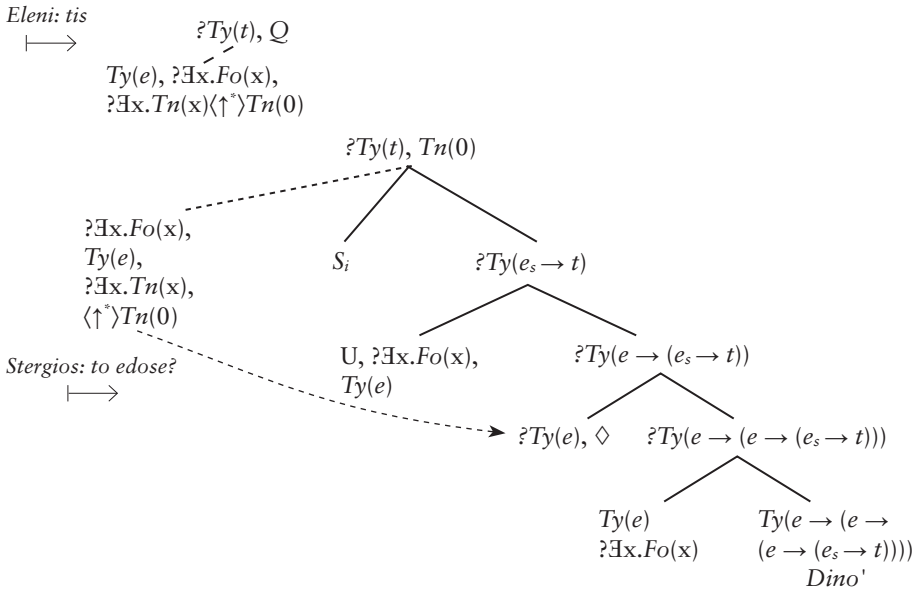
[11] BE is just a metavariable projected by the copula verb. We follow the analysis of the copula of Cann (2011). *Be* in this analysis is an underspecified metavariable. Thus, *BE* is no different from the regular type *e* metavariables, the difference being that it is a predicate metavariable rather than a type *e* one. The restriction *fem* should be thought of as holding for the first argument of the predicate. Thus given a predicate of the type λx.λs P(x)(s), the substituend for *x* should be specified as FEM. The adverb *poli* is not treated here.

[12] Even though the affix vs word debate is largely irrelevant to DS. See Chatzikyriakidis (2010, Chapter 3) for a discussion on this.

[13] Of course clusters like *ston*, *stin* 'him to you' and 'her to you' in SMG are treated as one lexical entry. In this case, such completions are not possible and thus these examples do not constitute a counterexample to what we are saying here.

accusative clitic builds the direct object node and annotates it with type *e* and a formula metavariable (the last argument node). The verb then provides the rest of the propositional template. The unfixed node can then be unified with the indirect object node. The second tree presents the situation just before unification.[14] The display below does not show the whole utterance of the first speaker but rather starts after the coordinating conjunction *ke* 'and' (*tis… to edose*):

[39] Eleni: Tis … 'to-her'
Stergios: to edose 'he-gave-it'



Stepping back from the details, note that it is the absence of a 'syntactic' level of representation distinct from that of such semantic representations which allows the direct successful integration of such fragments through the grammatical mechanisms themselves. In particular, on this simple direct projection of predicate-argument structure by both parties, there is no need to analyse the fragment as an independent sentential structure for which there has to be some presumed intermediate inferential step, in order to work out who is being referred to with what as some kind of contextual implication. The incrementality of both parsing and production is the key to the simplicity of the story. Both speaker and hearer are continuing from where they have got to in the utterance processing task, both intuitively and formally. In sharp contrast to the immediacy of the applicability of DS assumptions to the modelling of split utterances, this phenomenon of speaker switch in the middle of a dependency under construction is a major challenge for sentence-based grammar frameworks even to express at all. The grammars simply fail to provide the appropriate

[14] For more information and detailed examples of DS parses see Cann, Kempson & Marten (2005). For detailed specifications of Greek clitics, see Chatzikyriakidis (2010).

units: all dependencies are articulated grammar-internally as defined exclusively over hierarchically defined sentential structures which as a matter of principle are considered solely as statically displaying structural patterns without any concept of emergent growth. So the fact that the data which are so recalcitrant for these frameworks are predicted as a direct consequence of the novel framework provides strong confirmation of the stance adopted. And, though the very simple exchanges of [23], [25], [26] and [31] as modelled here, constitute a minimal collection of analysed data, they are symptomatic of the phenomenon as a whole.

As we look back from the vantage point of the new horizons made available by a dynamic perspective, we are now in a position to re-evaluate the relative lack of success of more orthodox frameworks trapped by their restrictedly sententialist methodology. These data remain intransigent for these frameworks because it is only by taking this step of articulating syntax as the incremental growth of representations of content that we can get any grip on the requisite analysis for language as displayed in its most widespread application, that of informal conversation. What the data display is how speakers and hearers build interpretation relative to whatever is available to them in the immediate context of their current processing task with seamless switching of roles at any point in a conversation, and across any form of dependency, morpho-syntactic, syntactic and semantic. Languages, that is, display the potential for a richly interactive build-up of interpretational effects; and it is only frameworks which express linguistic structure and interpretation in terms of evolving structures that will be able to reflect split utterance data straightforwardly. This story is not expected to vary from language to language, despite the huge variation in update associated with individual words or idiomatic phrases of a language: as long as a speaker and hearer of a language have a sufficiently large set of such update specifications in common, the interactive dynamics of conversational exchange is predicted to be seamless.

Inevitably in putting forward an account adopting such a shift of foundational assumptions, a number of questions tantalisingly beg for answers. The process of utterance interpretation as set out here has made no reference to recovery of any intentional attitudes underpinning the construal process. One might then ask, to what extent does this view of language undermine the very generally presumed centrality of mind-reading for successful communication? What role does recovering the speaker's intentions play in communication, on this new view? And how does language acquisition fit into this more dynamic picture? What is the significance of the domain-generality of the syntactic vocabulary? In particular to what extent are computational actions particular to language, or are they domain-general processing strategies? And, given the lack of encapsulation of the system, what is the precise nature of the interaction between mechanisms internal to the language faculty and the general dynamics of utterance interpretation? There are, no doubt, many more questions waiting to be properly formulated. What is certain, in the meantime, is that addressing issues raised by explaining the systemic use of elliptical fragments

in ordinary conversational dialogues opens up wholly new horizons for exploring concepts underpinning languages and their use.

References

Blackburn, P. & W. Meyer-Viol. 1994. "Linguistics, Logic and Finite Trees", *Logic Journal of the Interest Group of Pure and Applied Logics* 2(1), 3–29.

Bouzouita, M. 2008. "At the Syntax-Pragmatics Interface: Clitics in the History of Spanish", in R. Cooper & R. Kempson (eds), *Language in Flux: Dialogue Coordination, Language Variation, Change and Evolution*. London: College Publications, 221–64.

Cann, R. 2011. "Towards an Account of the English Auxiliary System: Building Interpretations Incrementally", in R. Kempson, E. Gregoromichelaki & C. Howes (eds), *Dynamics of Lexical Interfaces*. Chicago: CSLI Press.

Cann, R., R. Kempson & L. Marten. 2005. *The Dynamics of Language*. Oxford: Elsevier.

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

Chatzikyriakidis, S. & R. Kempson. 2011. "Standard Modern and Pontic Greek Person Restrictions: A Feature-Free Dynamic Account", *Journal of Greek Lingusitics* 11, 127–66.

Dalrymple, M., S. M. Shieber & F. C. N. Pereira. 1991. "Ellipsis and Higher-Order Unification", *Linguistics & Philosophy* 14(4), 399–452.

Fernández, R. 2006. *Non-Sentential Utterances in Dialogue: Classification, Resolution and Use*. Ph.D. diss. King's College London.

Fernández, R. & J. Ginzburg. 2002. "A Corpus Study of Non-Sentential Utterances in Dialogue", *Traitement Automatique des Langues* 43(2), 12–43.

Ginzburg, J. 2012. *The Interactive Stance: Meaning for Conversation*. Oxford: Oxford University Press.

Gregoromichelaki, E. 2013. "A Dynamic Perspective on Left-Right Asymmetries: Clitic Left Dislocation and Clitic Doubling in Greek", in *Rightward Movement in a Comparative Perspective*. John Benjamins, 321–68.

Gregoromichelaki, E. et al. 2011. "Incrementality and Intention-Recognition in Utterance Processing", *Dialogue & Discourse* 2(1), 199–233.

Grice, P. 1975. "Logic and Conversation", *Syntax & Semantics* 3(41), 58.

Kempson, R. & R. Cann. 2008. "Production Pressures, Syntactic Change and the Emergence of Clitic Pronouns", in R. Cooper & R. Kempson (eds), *Language in Flux: Dialogue Coordination, Language Variation, Change and Evolution*. London: College Publications, 179–220.

Pickering, M. & S. Garrod. Forthcoming. "An Integrated Theory of Language Production and Comprehension', *Behavioral and Brain Sciences*.

Purver, M. & R. Kempson. 2004. "Incremental Context-Based Generation for Dialogue", in A. Belz, R. Evans & P. Piwek (eds), *Proceedings of the 3rd International Conference on Natural Language Generation (INLG04)* (Lecture Notes in Artificial Intelligence no 3123). Brockenhurst: Springer, 151–60

Sperber, D. & D. Wilson. 1995. *Relevance: Communication and Cognition*. 2nd ed. Oxford: Blackwell.