Ruth Kempson, Ronnie Cann, Eleni Gregoromichelaki and Stergios Chatzikyriakidis

Action-Based Grammar

DOI 10.1515/tl-2017-0012

1 Competence-performance and levels of description

First of all we would like to thank the commentators for their efforts that resulted in thoughtful and significant comments about our account.[1] We are delighted that the main thesis we presented – the inclusion of interactive language use within the remit of grammars – seems to have received an unexpected consensus from such diverse perspectives. As each commentary deserves a more detailed consideration than space limitations would allow here, we will aim to address the most urgent issues and clarify what we see as misunderstandings.

1 In particular we thank **Poesio and Rieser** for their thorough reading drawing attention to typing errors that impede understanding. For example, indeed, they are correct that the double pointer on Diagram (48) in the paper is a typo. The double pointer device is only employed as a device to implement distinct perspectives on the parse (Eshghi et al. 2015), something we have not included in this paper. Given that comprehensive rule presentation presupposes the formal details of all the logics that intersect in the definition of the framework (see e.g. Kempson et al. 2001, Ch. 9), we did not include an appendix listing all the rules; instead, we elected to provide intuitive and concise descriptions indicating through references where the formal details can be found.

Ruth Kempson, Department of Philosophy, King's College London, Strand, London WC2R 2LS, UK, E-mail: ruth.kempson@kcl.ac.uk
Ronnie Cann, Department of Linguistics and English Language, University of Edinburgh, Old College, South Bridge, Edinburgh EH8 9YL, UK
Eleni Gregoromichelaki, Department of Philosophy, King's College London, Strand, London WC2R 2LS, UK; Institute of Cognitive Science, University of Osnabrück, Neuer Graben, 49074 Osnabrück, Germany
Stergios Chatzikyriakidis, Department of Philosophy, Linguistics, Theory of Science, University of Gothenburg, 405 30 Gothenburg, Sweden; Open University of Cyprus, B1 33, Latsia 2220, Cyprus

The data which form the focus of the current debate are *split utterances* such as (1)-(3), which we argued have to fall within the remit of grammar:

(1) A: Erm, this weekend, right?
B: right
A: we're going to..
B: to Aber Trynant, no?
C: uh?
A: where Ann,
B: Auntie Ann,
A: has just moved and we'd like to
C (interrupting): with the dogs?
B: if you can keep them under control
C: in the garden?
A: unless it rains
C: which it always does

In connection with these, it is important to note that inclusion of dialogue data within grammar analyses is not unique to us. Ginzburg, **Cooper**, and colleagues and, indeed, **Poesio and Rieser** have long argued for such a perspective (Ginzburg and Cooper 2004; Poesio and Rieser 2010). However, since its inception, Dynamic Syntax (DS, Kempson et al. 2001) has been motivated by integrating general processing mechanisms firstly introduced within semantics/pragmatics, more importantly, underspecification and update, within the grammar formalism. In all its forms and incarnations (ranging from the labelled deductive system (LDS) of Kempson (1996) to DS-TTR of Purver et al. 2010), DS also includes a non-linguistic concept of context subserving both the analysis of monologue and dialogue. So the split-utterance data come as further confirmation of this perspective and we took them as further evidence of the need to incorporate the time-linearity and context-dependency of processing within the grammar framework itself, underpinning both parsing and production. As we argued, in our view, this is what yields linguistic creativity, which is thus fundamentally a result of interaction. Therefore we are proposing that the structure of language crucially reflects its functioning in coordinating human interaction both with the environment and other people.

Nowadays, we are not alone in adopting this broad theoretical orientation, as **Poesio and Rieser** note: modelling incrementality has been a concern of other frameworks from categorial grammar (Steedman 2000; Morrill 2011) and Tree-Adjoining Grammars (Lombardo and Sturt 2002; Mazzei et al. 2007;

Demberg-Winterfors 2010) to functionalist accounts (O'Grady 2005). However, our more radical proposal states that this stance presupposes the abandonment of the analysis of *linguistic* structure on the basis of phrase-structure grammars which impose constituency analyses/derivations on strings expressed as syntactic types (categories, non-terminals). This view is pervasive, in particular, this is the type of grammatical analysis presupposed in Cooper (2014) even though the grammar as a whole eventually categorises speech events instead of abstract string-interpretation pairings. The same applies much more fundamentally to the grammar outlined in **Kobele, Merchant, Kobele and Merchant** where this assumption is essential for the claimed results. Various consequences follow from our differentiation. First of all, if the phenomena being explained by reference to the phrase structure of natural language (NL) strings are explainable without it, as DS sets out to prove, then theoretical parsimony indicates that the assumption of such a string-based organisation is at best redundant. Secondly, we have claimed that such an approach misses important generalisations by targeting the wrong level of explanation (see e.g. Gregoromichelaki et al. 2013; Gregoromichelaki in press). Therefore, eventually, such models are only able to deliver approximations of solutions to the puzzles that do not scale up well. But, more importantly, the distinction that DS aims to highlight is that instead of assigning psychological reality to hierarchical structures on strings, we need to directly attribute structure to meanings, whether these are represented as trees (Cann et al. 2005; 2012), TTR representations (Purver et al. 2010; Eshghi et al. 2013; Hough 2015) or geometrical structures (Gärdenfors 2014; Gregoromichelaki 2017). We then take a processing perspective on "syntax" as directly building/linearising conceptual structure and view these parsing/production processes as adequate to express all generalisations and regularities that appear when examining the apparent structuring of NLs. Thus, by removing the representational level of syntactic structure, we explicitly and formally model "syntax" as *actions*, procedures that reflect the factors that a psycholinguistic model would employ as explanatory. Words and the other stimuli involved in signal-categorisation are then modelled as "affordances" (action-triggers) that participants can take advantage of to coordinate with each other (Gregoromichelaki 2013a, in press; Kempson and Gregoromichelaki in press). As **Cooper** rightly urges, this embeds the theory of syntax, and generally linguistic behaviour, within a theory of human action while trying to preserve the insights gained by previous formal analyses.

For this reason, unlike what **Kobele** assumes, we do *not* continue "to abstract away from memory limitations, distractions, shifts of attention and interest, and errors", so we do not in our practice adopt the standard competence-performance distinction as he describes it. Indeed all such phenomena are

the target of enquiry: disfluencies, corrections, restarts, repairs are all to be modelled within the grammar (see e.g. Kempson et al. 2007; Gargett et al. 2008; Gregoromichelaki et al. 2009; Eshghi et al. 2015; Hough 2015). Within formal grammars, we are not alone in taking this stance (see Ginzburg and Poesio 2016 and references there). Additionally, in line with embodied approaches, we do not exclude the fact that machine/neural implementations might introduce properties of the system that need to be reflected in its architecture (for example, memory upper bounds in processing, see e.g. Honing and Zuidema 2014; Perruchet and Rey 2005; reconsideration of representational constructs, see e.g. Milkowski 2013; cf. Lewis and Phillips 2015). We therefore disagree with **Kobele** in that we believe that in his commentary he confuses legitimate idealisation/abstraction tools and devices (see also Stokhof and van Lambalgen 2011) with the standard interpretation of the competence-performance distinction so that the latter ends up having no empirical content at all (which might explain his puzzlement and conclusions). Therefore, given that we take the competence-performance differentiation seriously, we think it is perfectly sensible to question the theoretical assumptions that underpin its standard formulation.[2] It is not good enough, in our view, to present it as a harmless idealisation having to do just with the selection of relevant data: the distinction has been used to make hypotheses about mental architecture and the nature of "grammar" (Chomsky 1965 and many others following over decades).

Regarding Marr's levels-distinction, from what we can understand from his position, we think we are in agreement with **Kobele**'s conclusion in that we do aim primarily at explanation at the algorithmic level. This is because computational-level explanations, as e.g. his account, view linguistic processing in "black box" terms considering only inputs and outputs and can be implemented by a large number of non-intensionally-equivalent algorithms. Instead of remaining agnostic as regards the mechanisms, we elevate them as crucial both for the explanation of the empirical phenomena and for the parsimony of the general theory. So we are interested in directly modelling *processes*, peering inside the transitions of the "automaton" as **White** so aptly puts it (crucially though incorporating the time-linear dimension too) (see Kempson et al. 2015). Unlike Kobele, we are not aiming at abstract *descriptions* of processes that can be implemented in various ways and, consequently, we also do not claim to model abstract processes of "derivation" or proof-search (Kobele 2015; cf. Morrill and Valentin 2015) because, for those, psycholinguistic insights like incrementality and the attendant underspecification are not essential characteristics.

2 If Kobele and colleagues have already rejected the standard competence-performance distinction we would be very pleased to hear that we are in agreement.

2 The nature of the grammar

Unlike DS's action articulation, the grammar presented by **Kobele** here and in various places remains the analysis of a system of "signs" with syntactic categories being the glue that holds together two independent systems of sound (phonology) and meaning (semantics). Unlike DS, syntactic processes do not have direct access to the semantic structure so underspecification and update can only be modelled by means of empty strings, string-constituency manipulations etc., and an attendant multiple 'spurious ambiguity' problem. Categories, or so-called nowadays *syntactic types*, still remain sets of strings expressing hierarchical constituency assumptions only now paired with semantic interpretation. We do find accounts based on Categorial Grammar and the general framework of Minimalist Grammars (MG: Stabler 1997; Kobele 2012a, 2012b, 2015) much closer to ours than standard syntactic accounts like the traditional minimalist/HPSG ones we mentioned in the paper. Additionally, we find Kobele's (2015) account of ellipsis and **Kobele and Merchant**'s comments more convincing than their previous accounts (Kempson et al. forthcoming). However, even though Kobele's current account has significant advantages over standard syntactic accounts, and despite the similarity of the formal tools employed there to our analysis of VP-ellipsis (first appearing in print in Purver et al. 2006), conceptually, it still maintains standard assumptions regarding syntax which, from our point of view, lead both to empirical and to meta-theoretical difficulties that we itemise in the following.

Firstly, the Kobele and Stabler formalisms ground their assumptions about syntactic categories, syntactic operations, and syntactic constraints (the "tecto-grammar") on standard assumptions that still maintain the 'autonomy of syntax' hypothesis, however diluted this has become there and however much they have contributed to the move towards a less "syntactocentric" framework. Even if such grammars aim at transparency between syntactic and semantic rules, even if Move is assimilated to Merge, or local and global constraints are pushed into the category system (Graf 2013), the fact remains that they take the parser/generator model as relying on such arbitrary syntactic foundations instead of the other way round (for example, Stabler (1997) and **Kobele** here argue explicitly against the usefulness of insights from parsing for grammar design). In contrast, DS takes "syntax" as characterisable on the basis of general processing mechanisms (see also Bickhard 1993, 2009).

Secondly, such accounts adopt a standard competence-performance distinction, interpreting Marr (1982), in seeing the grammar as providing "a description

of what is being computed" (Kobele, 2012b), i.e., propositional knowledge-that, with the parser/generator implementing algorithms based on the grammar. DS explicitly rejects such a view of the distinction between computation vs. algorithm which is based on a view of language as a code, to be analysed on the model of formal languages (Gregoromichelaki 2013a). This is not just a matter of theoretical preference. We believe that these conceptual differentiations have also empirical effects, as we argue in the paper: computational-level accounts cannot provide satisfactorily natural accounts of the split-utterance data. In addition, we think that they miss some major generalisations in modelling syntactic mechanisms (e.g. long-distance dependencies) as sui generis, instead of qualitatively similar to underspecification-and-update, familiar from semantics/pragmatics, which becomes wholly natural when taking a processing perspective (we thank **Rieser and Poesio** for acknowledging this as a major theoretical DS contribution).

Thirdly, we believe that in our models we need to differentiate between *ambiguity* vs. *underspecification*. Therefore an account that postulates multiple underlying derivational analyses cannot be competitive with respect to one which provides a constructive presentation of a new antecedent for an elliptical fragment. This is what our analysis of VP-ellipsis and other phenomena such as sluicing require, even though we do adopt a *recomputation* instead of a *reuse* technique as **Kobele and Merchant** put it. Besides the conceptual issue of misclassifying underspecification as ambiguity, which leads to missing a significant amount of generalisations, in a model like ours, where processing constraints are taken into account and considered as part of the explanation for linguistic form/meaning, any unjustifiable spurious ambiguity (e.g. "memoizing" an open-ended amount of structures, Kobele 2012a) causes cascading explanatory difficulties. Moreover, as we aim to show, the same mechanism of being able to address the rerunning of parsing/generation actions has much wider applications, for example, unifying explanations for quotation phenomena, code-switching, and demonstrations of multimodal stimuli (Gregoromichelaki in press).

Beyond the phenomena of ellipsis, split-utterances, and dialogue, we believe that ignoring the processing perspective in a grammar creates problems that generalise as we argue in the paper. Even the most flexible constituency assumptions break down not only in dialogue but also in monologue as shown by the parenthetical examples (18)–(21) in the paper and many other phenomena like extraposition, scrambling, non-constituent coordination, etc. For a grammar of the type sketched by **Kobele**, despite some of its pleasing properties, it does not matter whether the parser/generator operates left-to-right or right-to-left. Both are possible in principle since the structure is analysed holistically by breaking down a "sign" to its components. In contrast, the procedures

implementing actions defined for DS are directional in the sense that their order of deployment makes a huge difference to the outcome. As **Cooper** points out, there are transparent relations to chart parsing here, given the composite dynamic of partially top-down, partially bottom-up processing. In fact, DS incorporates as part of the grammar a complete graph of the potential parsing/generation paths probabilistically ranked (Sato 2011; Hough 2015). It is unfortunate that reasons of space prevented us from presenting the whole current model leading to justifiable objections by Cooper in this respect. Moreover, indeed, as **Cooper** notes, DS is similar to a (parallel) ATN architecture but only at a superficial level. One difference is that structure building (recognition or generation) in DS is not driven by syntactic categories but by processes that model psycholinguistic insights (for example, the device of "unfixed nodes" models memory structures, processes, and limitations). As Cooper himself notes, another significant feature of DS is that the trees are *not* records of computation, showing how a logical expression has been constructed or parsed, but rather objects of computation themselves. Phonologically-analysed stimuli ('words') are one means to manipulate such objects (among others, Gregoromichelaki in press, 2017).

It is also not the case, as **Cooper** complains, that the system freely evolves in line with each and every newly constructible understanding, yielding in effect a separate grammar for each language (although, in fact, this might prove not to be a disadvantage as **Eshghi and Lemon** note, when coming to model language "in flux" to use Cooper's terminology). Like many other frameworks, DS assumes a universal set of elementary actions and computational rules, grounded in domain-general psychological processes, and a set of lexical rules which are language-specific. At this abstract level, there is nothing exotic and different about DS, distinguishing it from other approaches in highly-lexicalised frameworks like CCG, HPSG or LFG. It is also similar to even more computationally-oriented and less lexicalised frameworks like Ranta's Grammatical Framework (GF, Ranta 2011), where a distinction between an abstract syntax (a set of rules common to all languages) and concrete syntax (the linearisation rules for each language) is made. However, the nature of these rules is another issue and a number of differences can be found there. Unlike categorial grammar, in DS, by having purely conceptual structure trees (or indeed TTR representations, Purver et al. 2010) and a *pointer* which models attentional focus to indicate currently active states, word order can be manipulated independently of semantic types and/or combinatorial rules. Thus the conceptual types postulated can have universal application with word-order differences delegated solely to time-linear pointer movement.

More fundamentally, all DS procedures implement notions of initial underspecification and subsequent update (ranging from predictions ensuing without any linguistic input to very specific constraints imposed by 'grammaticalised' forms of underspecification, e.g. reflexive anaphors, case-markers, and island restrictions). Ordering the resolution of such a large amount of pending underspecifications is delegated to the management of the position of the *pointer* at each parse state. Assuming that the pointer position dictates the possible continuations in each particular language thus captures the philosophical insight that knowing a language is "knowing how to go on" (Wittgenstein 1953, para. 154), i.e. *acting coherently*, within a language game (see also **Eshghi and Lemon; White**; and **Cooper**). The emphasis here is on acting without presupposing mental state coordination as a prerequisite (cf. **Rieser and Poesio**; and Poesio and Rieser 2010), or propositional structure building as an essential motivation for acting. This explains how various speech-acts can be accomplished by just going on with the processing of the string, without unnecessarily introducing (meta)descriptions (conceptualisations) of the acts involved as in (2)-(3):

(2) 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)]

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

In such cases the opening of DS-syntactic dependencies is the motivating factor underlying the extension of subpropositional parse-states not some already envisaged propositional whole. And this holds for both comprehension and production: sub-propositional parse states are also involved in production (in testing for parsing consequences before generating) while the parse states themselves self-induce structure-generation to accommodate predicted inputs in comprehension. So we assume that the grammar architecture models both production and comprehension as closely linked, rather than isolating them from each other. In the paper, our analysis of the split-utterance phenomenon crucially involves the generation of predictions about immediately forthcoming actions so no overall propositional goal needs to be imposed for well-formedness, i.e. coherence (as mentioned in footnote 9 of the paper). This has consequences both for grammar design and philosophical considerations as we explain below.

2.1 Helpful and unhelpful ellipses

Kobele, for example, dismisses our claim that the competence/performance dichotomy needs to be revised, on the grounds that the sketch of a model that delegates the analyses of such phenomena partly to the grammar and partly to a probabilistic 'language model' will successfully model split utterances where the interrupting participant is intending to match and complete the propositional content which their interlocutor might have had in mind. From our point of view, it is conceivable that such an analysis is possible but only if the remit of explanation is restricted to those interruptions in which the hearer is trying to articulate what the speaker has in mind to utter, so-called helpful completions (see Lerner 2004). In these terms, the phenomenon might be seen as parasitic upon derivations of string-interpretation pairings reflecting sentence-based knowledge as held by some individual idealised speaker. In fact, such an analysis has already been provided by Poesio and Rieser (2010) (as **Rieser and Poesio** rightly point out), only in much more sophisticated, thorough, and convincing terms, even though their account involves calculation of joint intentions instead of a 'language model'. Indeed, an extension of Kobele's style of explanation might arguably be conceived as implementing the alternative analysis that Poesio and Rieser present in terms of Pickering and Garrod's Interactive Alignment model, i.e. as managing the probabilistic adaptation of each participant's resources to the other interlocutor. If that's all there is in the split utterances phenomenon, as Poesio and Rieser demonstrate, such cases might not provide grounds for challenging the pragmatic characterisation of "successful communication" as based on utterances in which the hearer is able to retrieve through the utterance some interpretation which the speaker could have had in mind to express (Sperber and Wilson 1995).

However, the problem which (1), and other data in the paper, aim to bring out is that such "helpful completions" are only a subset of the phenomenon. The problem that is highlighted with such examples arises when the initiating participant has no specific overall propositional goal in mind before they start speaking as is familiar both through introspection and psycholinguistic evidence (Kleist 1951; Goodwin 1979; Schober and Brennan 2003; Guhe 2007; Dember-Winterfors 2010). Moreover, speakers usually produce their utterances in "instalments" or "trial constituents" (Clark 1996) checking as they go along, and hearers provide acknowledgments, clarifications, or disagreements subsententially as they process word-by-word, many times leading to the abandonment of the initiated string when it has served some locally-defined purpose. Moreover, an interrupting participant may have no intention to use the provided interrupted fragment as providing evidence for their interlocutor's mental state

(beliefs, intentions, etc. as have been modelled in exquisite detail by Poesio and Rieser 2010; Ginzburg 2012, Larsson 2003), instead steering the action towards their own purposes (see e.g. (3)). In both types of case, there need be no overall propositionally-articulated contents, let alone sentences, which the initiating speaker has in mind prior to starting articulating their utterance. In our view, these data thus demonstrate that while mind-reading might be a sufficient condition for determining that successful communication takes place, it is not a *necessary* condition.

Now the grammar sketched by **Kobele** (based on De Groote 2001's Abstract Categorial Grammars, ACG) consists of two parallel homomorphically-related structures, hierarchically-structured strings and semantic values. In this formulation, the computational primitives introduced are indeed minimal; and elsewhere Kobele and colleagues, who have done a lot to move current research towards more representationally-minimal and less "syntactocentric" assumptions, claim that such a type of account does not introduce a syntactic level of representation (Kobele 2012a).[3] We acknowledge that this is true. However, Kobele's use of the ACG formalism here and similar devices elsewhere still maintains a view of syntax that is independent and distinct from other cognitive abilities and, more closely to home, semantics and pragmatics (which is exactly what DS is designed to refute as **White** perceptively points out). Here and in other formulations Kobele and colleagues assign psychological reality to syntactic categories, operations, and global constraints (e.g. the Shortest Move Constraint) which are assumed to be specifically linguistic and not reducible to general cognitive constraints (see also Stabler 2011; Stabler 2013). They are then imposed in the grammar most often nowadays hidden within complex syntactic categories. The semantic reflexes of such complications are then inevitably of dubious psychological grounding taking us back to claims of the 'autonomy of syntax' hypothesis.

What is more relevant to the issue at hand, this has also empirical consequences. Firstly, if you define a licensing grammar on the basis of string-constituency, there needs to be a distinguished non-terminal (more generally, a *distinguished type*), namely, a sentence (and, derivatively or primarily, a proposition). But then, incomplete, interrupted, or abandoned strings, duplication of words or corrections (see e.g. (2) earlier), which are not only common but also essential for dialogue (Clark 1996), will never receive any licensing. Secondly, despite Kobele's suggested interpretation of production rules as "actions", these

3 See also **Cooper** mentioning Cooper (2014, in prep) where, however, the minimal coverage of presumed "syntactic" phenomena is not adequate for us to judge whether the claimed epiphenomenal nature of the assumed "syntactic" aspect of the sign can be maintained.

are not intended as models of human parsing/generation actions unless a parsing/generation regime is imposed so that they necessarily operate in a time-linear manner (this is the result of the account being couched at the computational level). For example, despite Kobele's claims to the contrary, a compositional semantics does *not* give you online incrementality "for free" (even though, as he rightly points out, the *strict competence hypothesis* can be thus maintained). Despite the fact that the grammatical knowledge assumed intrinsically includes the potential of breaking down sentences to interpretable parts, again, in particular cases, this has to be imposed by the parsing/generation control structure and is not given as part of the syntactic apparatus.

Thirdly, as regards Kobele's comparison with the general DS account, it is possible to reformulate DS under a proof-search paradigm. This is shown, in a more succinct and more suitable for our purposes manner, by the sketch of a translation by **White**, who rightly points out that various explorations have been undertaken within DS for articulating conceptual structure. However, what unites all these approaches is the philosophy underlying all DS theoretical choices, namely, that syntax models the psychological processes of parsing/generation. So all versions of DS maintain a commitment to keeping the semantic types as much as possible grounded to plausible conceptualisation interpretations. On the other hand, syntactic operations (e.g. unfixed nodes, LINKed trees, lexical access etc.) model memory operations and articulation and directly manipulate the assumed conceptual objects (simplified in the paper as trees) as **Cooper** points out. Underspecification and update, notions that operate across cognition, then account for all apparent syntactic restrictions (e.g. the Shortest Move Constraint). The consequences of abandoning this stance are more evident in **Kobele**'s attempt to mimic the DS analyses of (some) split utterances in the sketch of a framework that he outlines. Like Poesio and Rieser, but presumably more on the lines of Anderson (1990), he has to invoke probabilistically-based "rationality" assumptions external to the grammar. It is well-known that such assumptions eschew the "mechanistic" interpretations of cognitive psychology (Anderson 1990). But such mechanistic explanations are exactly the point of modelling the phenomena through DS. More specifically, the DS account, by including production and predictive processing within the grammar, is able to substantiate in more explicit and explanatory terms the account that Kobele sketches: it is because the grammar itself operates via predictions that a hearer who assigns high likelihood to an upcoming conceptual content can attempt lexical access and vocalise a continuation. They do not have to "guess" or metarepresent mental states regarding the other interlocutor. Kobele might argue that such "rationality"/optimisation constraints are not to be regarded as part of the grammar since he wants to identify a specifically linguistic "causal

power" that explains linguistic structuring (but cf. Kobele et al. 2013). However, if one is prepared to accept a run-of-the-mill probabilistic grammar (Hale 2006), what is the justification for saying that processes that operate at more comprehensive levels should in principle be excluded from the model? Are we not missing a generalisation here?

Notice also that, contra to **Kobele**, such "fragments" do not necessarily have to be parts of sentences since they can be perceived and understood well enough but nevertheless abandoned midway (see e.g. (1)). Moreover, there are also the cases where the two " split" strings may add up to an ungrammatical string, because of the change of speaker, and yet constitute a perfectly well-formed exchange. This was illustrated in the paper by example (28), among others, repeated below:

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

In describing our analysis, Kobele seems to have misunderstood the whole point of our account and indeed the framework: we do not "make the objects speakers and hearers are generating more abstract" because interlocutors are not assumed to produce/comprehend syntactic structures. So it is *not* syntactic structures that are extended during the processing of split utterances. Even though Kobele attempts to mimic our solution formulated in his terms, he never answers the crucial question: how is this string-interpretation ever licensed in his grammar since the phonological string cannot be generated (licensed, "proved") as a whole? Is the 'language model' implementing the rationality assumptions able to override the grammar? This is of course a possible account but it seems to us very unnatural compared with the naturalness of the example.

3 Ambiguity vs. context-dependence

Kobele and Merchant focus on the subpart of our paper in which, as part of a critique of standard frameworks, we argue that the assumption of a starting symbol for a phrase structure grammar that licenses all linguistic structures would force sentential (or at least constituent) analyses for various types of "fragments" we identify. This is not such an extraordinary assumption as they present it to be nor a "straw man" especially given that we have not been given

an answer as to how their own grammar comes to license non-sentential strings.[4] Even Ginzburg (2012), who otherwise soundly and forcefully eschews a syntactic account of ellipsis, assigns NPs-projecting-to-S analyses to non-sentential utterances for theory-internal technical reasons having to do with the structure of the independent syntactic component assumed (for more orthodox syntactic analyses see Ludlow 2005; Marti 2006; Weir 2014 a.o.). We are very pleased to hear and applaud the fact that **Kobele and Merchant** would not have to resort to such analyses; and we might be convinced that they can potentially handle the types of "fragments" that Merchant mentions. But, under their analyses sketched in various places, we do not understand Merchant's complaints having to do with the flexibility of our system: surely, allowing for a variety of sui-generis mechanisms (deletion, type-shifting, complex categories, etc.) to implement resolutions for a variety of fragments (**Merchant** and Merchant 2004; Merchant 2010) is a clear presentation of a multitude of explanatory resources at their disposal too (and, still, data such as in (4) cannot be captured). We ask then in turn: for each particular analysis cited in **Merchant** and **Merchant and Kobele** what prevents any of the other postulated mechanisms from applying in that case and overgenerating analyses and licensing?

Indeed, we insist that the flexibility of resources assumed in DS is justified by the complexity of the data being analysed. For example, **Merchant and Kobele** claim that the hardest cases are those such as *The information could have been released but Gorbachov chose not to* where ellipsis reconstruction involves "deviations from identity"; and they suggest that these "seem likely to pose a problem for KCGC". We believe that this might reflect a misunderstanding: to the contrary, the DS account of ellipsis is based on a parallelism with anaphora, namely, again, employing initial underspecification and subsequent update. One type of update is copying contextual content directly rather than rerunning actions to construct structure anew (i.e. unlike what happens in sloppy/paycheck readings cases). This mechanism (also employed for the processing of bound pronouns) in the passive surface non-identity cases allows reconstruction of *content* by replicating the content of the antecedent directly, rather than by making reference to the words. The supposedly problematic reconstruction of an ellipsis site from the predicate of the antecedent thus turns out to be precisely what is needed in such cases since, at the level of conceptual structure, the surface syntactic subject of the antecedent VP, being a passive form, will end up providing content at the object argument node. However, notice that the DS account employs exactly the

4 In contrast, Barton (1991) provides such an answer although not an analysis as fine-grained as will be needed for our data, while Minimalist analyses like Progovac (2006) allow for non-sentential licensing but only for a subset of "fragments"; both do not embrace incrementality.

same mechanisms for ellipsis resolution here as those that are the basis of implementing anaphora resolution with its restricted linguistically bound (e.g. anaphors and bound variables) and freer indexical construals so that all such phenomena appear unified.

In contrast, **Merchant** argues that even ellipsis data, including split utterances, do not constitute a natural phenomenon (as do **Rieser and Poesio**). In confronting indexical situation-recoverable ellipsis, which is problematic for syntactic accounts if they aim at uncovering generalisations, he bites the bullet and proposes a type-shifting rule for interpreting predicates occurring without their arguments overtly. It seems to us that this semantic operation with no string-related consequences is of a completely different nature to his otherwise structural account of ellipsis, thus contra-indicating the generality of that form of analysis. On the other hand, with situation-recoverable ellipsis which nonetheless requires appropriate choice of morphological form, he has to invoke yet another licensing mechanism, that of the presence of an implied "script", which, nevertheless, has to be linguistically-articulated in order to impose "structural case" (cf. Barton 1991; Progovac 2006). These analyses add to the panoply of accounts of ellipsis with the unacceptably high levels of ambiguity for each linguistic element occurring in each particular context. We believe that, despite the apparent flexibility, rather than simply create a list of different types of data, our analysis to the contrary strives to take these data as signalling the systemic nature of context-dependence in natural language. And we plead guilty to **Rieser and Poesio**'s charge that we attempt to unify apparently diverse phenomena like that of split utterances, against the backdrop of that much more general problem. Take for example the following, which, Merchant argues, is not elliptical in any way:

(5) A: The doctor
B: Chorlton?
A: Yeah Chorlton, he insisted that I have further tests.

We agree with his conclusion. However, as a case of clarification ellipsis, it would be analysed as sentential (an NP that projects S) and propositional in the most prominent account of clarifications, namely, Ginzburg's, especially since in other languages (e.g. Greek) such NPs will have to bear the appropriate case-marking in parallel with Merchant's examples of indexically-resolved script-related fragments.

Under our analysis, A's first utterance is an opening towards a continuation, anticipating some form of propositional completion. To this initial utterance, B

provides a follow up clarification which, depending on the intonation, can be perceived either as clarification request or clarification offer. A then replies with a further non-sentential element in agreement. This then anaphorically links in to the structure that follows through the overt pronoun *he*. The issue in this type of example is not so much whether these are ellipses of a sentential sort or not (though arguably the opening utterance would by standard assumptions be seen in such terms, needing a predicate). More significant is that what the interpretation of such fragments shows is how the immediate preceding structure is the crucial context relative to which the fragment itself is construed as an update (Gregoromichelaki et al. 2013). The process both of inducing structure and the attribution of content to that structure are thus essentially context-dependent. And it is this puzzle, with its additional problem of apparently needing some processes of higher order inference on the part of the participants in the form of mind-reading, which the DS analysis aims to defuse.

This is achieved by adopting a low-level mode of explanation in which such fluent role-switching interactivity is a side-effect of grammar-internal mechanisms determining that structure/interpretation is built up on an incremental word-by-word basis for both parties. For the case of data like B's utterance in (5), we have at least two options corresponding to distinct construals: for cases where this is an offer of clarification, we will analyse those as simply appositional (as **Rieser and Poesio** observe). For the cases where B's utterance is analysed as a clarification request, it is essential for its explanation to include mention of the assumption, present in DS at least since Sato (2011), that the full architecture incorporates a full model of the grammatical *context* as probabilistically-ranked options of all the next states that the grammar licenses at each word-processing step. This is illustrated in the form of a directed acyclic graph (DAG) (Eshghi et al. 2011; Eshghi et al. 2013; Hough 2015; Hough and Purver 2014; 2017). Clarification requests and their responses are then modelled not as part of the semantic structure via modelling some presumed conceptualisation of speech-act descriptions but, procedurally, as backtracking movements along the DAG potentially leading subsequently to distinct previously disfavoured processing paths (Eshghi et al. 2015). The same mechanism is employed for self-corrections and other disfluencies (Hough 2015) and the explanation of how puns and other word-plays operate (Gregoromichelaki in press). For our purposes, this implies that the split-utterance data in all their complexity fall out as an immediate consequence, without having to invoke any externally imposed higher-order mode of inference or "grammaticalised" speech-act characterisations, only the assumption of a shared grammar that enables interlocutors to go on with processing in the hope that trouble will be resolved downstream. In fact, as **White** points out, local representations of the shared context might not be

isomorphically related. This does not seem to impede communication[5] as long as the low-level actions are "choreographed" appropriately so that they coordinate.[6] More broadly, in fact, in our view, the non-equivalence of the 'wide' content (**Rieser and Poesio**) of the conceptual structures locally assembled by each participant is the source of all the significant effects that interaction confers (Healey 1997; 2008; Healey et al. 2014) and the basis of linguistic creativity and conceptual abstraction (Mills 2011; **Cooper**). But, we suggest, we should not confuse the function with its result: natural languages are not phonology-semantics mappings, instead they are mechanisms for performing such mappings ad hoc in each context (Bickhard 1993; 2009; Gregoromichelaki 2017; Gregoromichelaki in press; and cf. Cooper in preparation 2017).

As **Rieser and Poesio** indicate, we should therefore confirm here that the DS account turns on not representing propositional descriptions of speakers' attitudes to the content of what is uttered as an essential part of the output interpretation or planned utterance (unlike the Gricean account of $meaning_{NN}$) (Grice 1975). We accept that natural languages are the par excellence expressions of a symbolic system of conceptualisation and this is what our trees reflect (in simplistic form, for more elaborate models and variations, see Purver et al. 2010; Hough 2015; Gregoromichelaki in press). But, further than this, we do not believe that every nuance of meaning has to be reified and expressed as a conceptual structure element in the meaning representations. As **Eshghi and Lemon** very convincingly argue, meaning depends on the language games the language users are engaged in. Their model instantiates how we conceive the analysis of (implicit) speech act characterisations (see also Gregoromichelaki and Kempson 2015). Language games as types are malleable learned routines involving complementary non-linguistic and linguistic actions that constitute knowledge-how that language users accumulate. The processual coherence afforded within such games has to be constructed each time subsuming the actions that linguistic stimuli trigger (hence our analysis of morphological cases and other structural indications of indexical construals as not necessitating the presence of linguistically-derived contents). So at each re-enactment of such a game not only is retrospective context provided but also the contents assigned to physical actions, environmental stimuli, and linguistic elements are recomputed (Mills and Gregoromichelaki 2010). So, to answer **Rieser and Poesio**'s question

5 For arguments that they are not *necessarily* distinct, as Ginzburg (2012) argues, see Gregoromichelaki and Kempson (2015).

6 For more such relevant, in our view, assumptions pertaining to the issue of synchronising processes from distinct sources, see Rieser (2016) where a calculus is proposed for emergent choreographing of the various channels of communication.

regarding propositions, our analyses of non-sentential utterances presuppose neither linguistic elements (e.g. verbs) nor even conceptual elements (e.g. predicates) in the context because we do not assume that language users necessarily explicitly conceptualise *descriptions* of the routines they are able to participate in (contra **Cooper** and unlike what **Merchant** refers to as "scripts" if we have understood his claim correctly). People develop coordinative patterns of behaviour, learn 'how to go on', but following a 'rule' is distinct from 'knowing' this rule (Wittgenstein 1953, para. 202). Hence non-sentential utterances are shareable without meta-representational assumptions because they integrate within such shared games and there is no need for explicit conceptualisation of descriptions of the dialogue moves made. Unless there is explicit indication of illocutionary force, we therefore eschew fine-grained categorisations of speech acts and attitudes. Instead, we attempt to provide mechanistic characterisations of the routines involved (see, for example, Eshghi et al. 2015). On the other hand, we do not doubt that interlocutors can and often do conceptualise such routines and the content of their actions (as **Cooper** indicates). This is shown by the fact that there are lexicalisations of descriptions of such speech acts (performatives). In such cases, the representation of meta-level attitudes can indeed be modelled as adjuncts to the structure under construction, as so-called LINKed structures (Purver et al. 2010; Gregoromichelaki and Kempson 2015), but such explicit representation of meta-level attitudes is *not* a sine qua non of successful communication. And given the ease with which very young children can participate in split utterance activity, this suggests, at the least, a robust alternative to explore (see Gregoromichelaki et al. 2011).

In one sense the phenomenon of ambiguity in natural language is deeply familiar: it is everywhere, something so common that it is generally taken merely as diagnostic of natural language, a challenge for the child to grapple with in coming to acquire a grammar, to whom innate capacity is attributed in order to render the acquisition task possible at all. But, on this view, the mystery of first language acquisition deepens year by year as the numbers of ambiguities multiply – with the unearthing of yet another type of ellipsis, yet another speech act, yet another type of implicature, yet another type of relative clause, etc. However, as **Rieser and Poesio** pointed out, according to the DS perspective, such dependency of natural language expressions on the attendant context is even interwoven with the structural architecture of language itself: it is the very heart of what language is about, the clue to its flexibility, its lack of determinate fixing of boundaries. In this, we suggest, language is like vision and other forms of cognitive processing which, according to Clark (2013, 2016), are all active processes which are invariably determined relative to the specifics of the context, our physical stance, what we have just been thinking about, our emotions, our physical state.

Context on this understanding ranges over conceptual representations, whether constructed in virtue of the processing of some immediately previous string, from memory, or as a reflection of objects in or related to the current scenario and participation in some language game, or even the processing actions just executed as parser or producer. Given the DS stance on 'syntax' as the progressive manipulation of domain-general representations, there is no arbitrary division to be made between grammar-internal and grammar-external representations.[7]

At this juncture, it is important to stress that we do not refuse, as a matter of ideology, ever to sustain claims of ambiguity for phenomena: the dispute whether given phenomena constitute ambiguity or underspecification is not a simple dichotomy. To the contrary, natural languages are evolved systems subject to both cultural and psychological constraints which determine routinisations of usage leading over time to calcified, distinct lexical entries, so claims of ambiguity for individual cases may always be justifiable. But we urge any such move has to be on a case by case basis (which is a process superbly illustrated in **Merchant**'s work despite our disagreements about the outcomes) reflecting the many psychological, physiological and sociological considerations that can impinge on how languages change.

4 Comparing alternative dynamic frameworks and extensions

In his evaluation of the importance of articulating dynamic alternatives to the more conventional static frameworks, **Cooper** suggests ways in which DS effects might be transferred in independently justified frameworks, specifically TTR; and these clearly warrant further probing. TTR, as far as we can see and at least in its bare form, has not been developed to capture the rich array of empirical results established to date within the DS framework[8] (but cf. Ginzburg 2012; Ginzburg et al. 2014). There are,

7 Despite his embodied stance on the nature of mind/brain, A. Clark remains agnostic over whether such representations are epiphenomenal. Like him, we remain open on the issue of whether such representationalism may turn out to be ultimately eliminable.

8 For example, Kempson and Kiaer (2010) for Korean multiple scrambling, Gregoromichelaki (2013b) and Chatzikyriakidis (2010, 2012, 2017) for various clitic related phenomena in Greek and dialects of Greek, Bouzouita (2008) for the diachrony of the Spanish clitic system, Cann (2011) for idiosyncracies of the English auxiliary system, Cann and Wu (2011) for idiosyncracies in Chinese passive forms, Sehraku (2013) on Japanese clefts and the restricted optionality of case marking, Gibson (2015) for idiosyncratic left-peripheral word orderings in Rangi, a typologically exceptional Bantu language.

without doubt, reasons to find constructive type theory attractive, especially concerning sublexical meaning and interfaces with multimodal inputs and outputs (Larsson 2011), thereby indicating a TTR extension of what DS offers, as long as the time-linear dimension of analysis that DS imposes is respected, which we do not doubt is feasible given the expressivity of TTR as a formalism (see Purver et al. 2010; Eshghi et al. 2013; Gregoromichelaki in press).

Similarly, **White** suggests a high-level point of view from category theory that, in a developed form, is certainly a promising perspective from which we might be able to glean some deeper truths underlying the specifics of our analyses. Moreover, the issue of exploring the effects of seriously thinking about *concurrency*, along with incrementality (*sequentiality*), is a revealing approach to be taken we believe, a perspective that indeed has been taken up in recent work by Rieser (Rieser 2016), admittedly for other purposes. Speculatively, we suspect that therein lie the rudiments of a mechanistic processual account of mental state coordination rather than the imposition of mindreading computations. The implications of the concurrency of comprehension and production processes have not been explored in depth in DS even though, like Hendriks (2014), we do assume that both need to be modelled within the grammar. In this connection, we note that in not positing mindreading as the underlying condition of communicative actions, but instead predictive, goal-directed processing, the DS perspective is commensurate with A. Clark's explication of the entanglement of action and perception in cognition: in preparing for cognitive response, whether in action or perception, low-level proprioceptive preparations for action are involved, not higher order forms of inference (cf. Frith 2007; Tomasello 2008; Friston and Frith 2015). From this point of view, any asymmetries between generation and parsing are directly analogous to the assumption of that framework that all cognitive responses involve actions. It is in virtue of this perspective, that we have reason not to follow Hendriks (2014) in her claim that production and comprehension are radically different processes despite **Rieser and Poesio** drawing our attention to her work as grounds for their own scepticism of our integrative stance (see also Kempen et al. 2012; Kempen 2014). Like Clark, we would see the distinctiveness of parsing and generation as explained at the lower level of the mechanisms themselves in the variant forms of realisation expected (see also Hurley 1998; Gregoromichelaki 2013a). However, we completely agree with Hendriks (2014) in eschewing the view of NLs as codes with fixed form-meaning mappings and, concomitantly, that any comprehension-production differentiations lie entirely within the purview of the grammar.

Relatedly, as **Cooper** rightly complains about, reasons of space did not allow the full presentation of the architecture we assume, which does include speaker-hearer parameters and perspective differentiation. In fact, since Sato (2011), DS has incorporated a full model of context with probabilistically-ranked options of all the next states that the grammar licenses at each word-processing step. As we said earlier, this is illustrated in the form of a directed acyclic graph (DAG) fully articulated in Eshghi et al. (2011), (2013); Hough (2015); Hough and Purver (2014). On the basis of this architecture, Eshghi et al. (2015) model the distinct perspectives of speaker and hearer when issues of feedback arise, namely, backchannel acknowledgments and clarification issuing and repair. This mechanistic account from within the grammar constitutes an explication that does not resort to mindreading as we mentioned earlier.

Moreover, **Cooper**'s other complaint that DS misses the significance of the sense in which knowledge of language can be seen as event-type characterisation would seem to us not exactly accurate. At the most fundamental level, DS models the process of characterising or producing linguistic events so the speech event is modelled directly across the flow from its initiation to its end. The requisite speech-event parameters are also included, having adopted a version of Poesio and Rieser's 'micro-events', a modelling of the individual sub-events that make up that flow of time as identified by word boundaries (Purver et al. 2010; Gregoromichelaki 2017). Where Cooper is right is that we do not impose an abstraction over this process in order to have the grammar just assigning a (static) type to an utterance event (cf. Ginzburg 2012). However, even this potential has been exploited within DS by employing ad-hoc TTR-types to characterise on the fly grammatical performances (DS-action executions) presented as "demonstrations" for purposes of explaining quotation phenomena and code-switching (Gregoromichelaki in press; cf. Ginzburg and Cooper 2014).

Furthermore, recent work within the same architecture shows not only DS's viability as a grammar formalism, but also its fit to include multi-modal data, and the provision it makes available to enable a large amount of dialogue data to be acquired from very small amounts of unannotated data using a combination of DS and Reinforcement Learning (Yu et al 2016; Kalatzis et al. 2016). The model then has notable links with first-language acquisition. Eve Clark (2009) and colleagues emphasise the central role that interaction plays in a child's linguistic development (Clark and Casillas 2016; see also Hilbrink et al. 2015). As **Eshghi and Lemon** point out, the framework supports straightforward learning of speech act conventions simply from the interactive exchanges that take place in informal conversation. Indeed, Eshghi, Lemon and colleagues already have the beginnings of work demonstrating how an agent can learn from very small data sets, given nondeterminism in the selected strategies, the availability of clarification devices

whenever needed, and the ability to evaluate between alternative strategies. With increasing evidence (Eshghi et al. 2013) of such learning, there is therefore emergent research that learning by a child in interaction with other agents is possible with the outcome of yielding a functioning grammar. This in its turn promises to open a window on a wholly new and gradualist account of language evolution broadly in line with Christiansen and Chater (2016), but with communicative success being grounded in interaction rather than in recognition of higher-order intentions. However, Eshghi and Lemon correctly note that the implemented DS framework presumes that lexical specifications are shared between parties, at least in relation to their structural attributes. In this connection though, current work (Cann and Kempson forthcoming) has begun to attempt the formulation of how lexical specifications never constitute more than constraints on concept construction, so that concepts are invariably constructed and predicates developed, rather than being provided directly by lexical encoding. It seems to us that TTR-modelling of ad hoc type construction is also a viable strategy in this direction (Cooper 2012; Gregoromichelaki 2016). Nonetheless, as Eshghi and Lemon rightly point out, there is a big gap from such elementary steps to a full-blown explanation of how child or adult grammars may develop and change in interaction with other agents and in relation to the scenario within which such interaction is taking place. There is also the important issue of the status of *group* within any account of evolution (Sober and Wilson 1998; Wilson 2002). While gaps such as these remain, any account of either first-language acquisition or the corresponding narrative of language evolution will remain no more than promissory notes, and importantly incomplete.

References

Anderson, John 1990. *Cognitive psychology and its implications*. Mahwah, NJ: Lawrence Erlbaum Associates

Barton, Ellen 1991. Nonsentential constituents and theories of phrase structure. In K. Leffel & D. Bouchard (eds.), *Views on phrase structure*, 193–214. Dordrecht: Kluwer.

Bickhard, Mark H. 1993. Representational content in humans and machines. *Journal of Experimental & Theoretical Artificial Intelligence* 5(4). 285–333.

Bickhard, Mark H. 2009. The interactivist model. *Synthese* 166. 547591.

Bouzouita, Miriam. 2008. *Clitic placement in the history of Spanish*. London: King's College, Ph.D. thesis, 22.

Cann, Ronnie. 2011. Towards an account of the English auxiliary system. In R. Kempson, E. Gregoromichelaki & C. Howes (eds.), *The dynamics of lexical interfaces*, 279–317. Stanford: CSLI Publications.

Cann, Ronnie, Ruth Kempson & Lutz Marten. 2005. *The dynamics of language*. Oxford: Elsevier.

Cann, Ronnie & Wu Yicheng. 2011. The *bei* construction in Chinese. In R. Kempson, E. Gregoromichelaki & C. Howes (eds.), *The dynamics of lexical interfaces*, 239–378. Stanford: CSLI Publications.

Cann, Ronnie, Ruth Kempson & Daniel Wedgwood. 2012. Representialism and linguistic knowledge. In R. Kempson, T. Fernando & N. Asher (eds.), *Philosophy of linguistics: Handbook of the philosophy of science 14*, 357–401. Oxford: Elsevier.

Chomsky, Noam. 1965. *Aspects of the theory of syntax* Cambridge, MA: MIT Press.

Chatzikyriakidis, Stergios. 2010. *Clitics in four dialects of Modern Greek: A dynamic account*. London: King's College, Ph.D. thesis.

Chatzikyriakidis, Stergios. 2012. A dynamic account of the Cypriot Greek clitic positioning system. *Lingua* 122. 642–672.

Chatzikyriakidis, Stergios. 2017. Afterthoughts in Greek: Gender mismatches under a dynamic framework. *Journal of Linguistics* 53(2). 279–325.

Christiansen, Morton & Nick Chater. 2016. *Creating language: Integrating evolution, acquisition and processing*. Cambridge, MA: MIT Press.

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

Clark, Andy. 2016. *Surfing uncertainty: Prediction, action and the embodied mind*. Oxford: Oxford University Press.

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

Clark, Eve & Marisa Casillas. 2016. First language acquisition. In K. Allan (ed.), *The Routledge handbook of linguistics*, 311–329. New York: Routledge.

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

Cooper, Robin. 2012. Type theory and semantics in flux. in R. Kempson, T. Fernando & N. Asher (eds.), *Handbook of the philosophy of linguistics*. Oxford: Elsevier.

Cooper, Robin. 2014. Phrase structure rules as dialogue update rules. In V. Rieser & P. Muller (eds.), *Proceedings of the semantics and pragmatics of dialogue (SEMDIAL)*, vol. 18, 26–34. Edinburgh: University of Edinburgh.

Cooper, Robin. 2017. Adapting type theory with records for natural language semantics. In S. Chatzikyriakidis & Z. Luo (eds.), *Modern perspectives in type-theoretical semantics*, 71–94 Cham: Springer International Publishing.

Cooper, Robin. in prep. *Type theory and language: From perception to linguistic communication*. Draft of book chapters. https://sites.google.com/site/typetheorywithrecords/drafts.

De Groote, Philippe 2001. Towards abstract categorial grammars. Proceedings of Association for Computational Linguistics, 39th Annual Meeting and 10th Conference of the European Chapter. 148–155.

Demberg-Winterfors, Vera. 2010. *Broad-coverage model of prediction in human sentence processing*. The University of Edinburgh, Ph.D. Thesis.

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, 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. 2013. Probabilistic grammar induction in an incremental semantic framework. In D. Duchier & Y. Parmentier (eds.), *CSLP lecture notes in computer science 8114*, 92–107. Berlin: Springer Verlag.

Friston, Karl & Christopher Frith. 2015. A duet for one. *Consciousness and Cognition* 36. 390–405.
Frith, Christopher. 2007. *Making up the mind: How the brain creates our mental world*. Oxford: Wiley-Blackwell.
Gärdenfors, Peter. 2014. *The geometry of meaning: Semantics based on conceptual spaces*. Cambridge, MA: MIT Press.
Gargett, Andrew, Eleni Gregoromichelaki, Christine Howes & Yo Sato. 2008. Dialogue-grammar correspondence in dynamic syntax. Proceedings of the 12th Workshop on the Semantics and Pragmatics of Dialogue. (LONDIAL'08).
Gibson, Hannah. 2015. The dynamics of structure building in Rangi: At the syntax-semantics interface. *Cambridge Occasional Papers in Linguistics* 8. 41–55.
Ginzburg, Jonathan. 2012. *The interactive stance: Meaning for conversation*. Oxford: Oxford University Press.
Ginzburg, Jonathan & Massimo Poesio. 2016. Grammar is a system that characterizes talk in interaction. *Frontiers in Psychology, Section Language Sciences*. https://doi.org/10.3389/fpsyg.2016.01938
Ginzburg, Jonathan & Robin Cooper. 2004. Clarification, ellipsis and the nature of contextual updates in dialogue. *Linguistics and Philosophy* 27(3). 297–365.
Ginzburg, Jonathan & Robin Cooper. 2014. Quotation via dialogical interaction. *Journal of Logic, Linguistics and Information* 23. 287–311.
Ginzburg, Jonathan, Raquel Fernández & David Schlangen. 2014. Disfluencies as intra-utterance dialogue moves. *Semantics and Pragmatics* 7(9). 1–64.
Goodwin, Charles. 1979. The interactive construction of a sentence in natural conversation. In George Psathas (ed.) *Everyday language: Studies in ethnomethodology*, 97–121. New York: Irvington Publishers.
Graf, Thomas. 2013. *Local and transderivational constraints in syntax and semantics*. University of California, PhD Thesis.
Gregoromichelaki, Eleni. 2013a. Grammar as action in language and music. In M. Orwin, C. Howes & R. Kempson (eds.), *Language, music and interaction*. London: College Publications.
Gregoromichelaki, Eleni. 2013b. Clitic left dislocation and clitic doubling: A dynamic perspective on left-right asymmetries in Greek. In G. Webelhuth, M. Sailer & H. Walker (eds.), *Rightward movement in a comparative perspective*. Amsterdam: John Benjamins.
Gregoromichelaki, Eleni. 2016. Grammars as mechanisms for fine-grained situated conceptualisation. Talk presented at the 'Word Meaning Workshop' University of the Basque Country, Vitoria-Gasteiz, November 2016.
Gregoromichelaki, Eleni. 2017. *The case of non-sentential speech acts*. Ms University of Osnabrueck.
Gregoromichelaki, Eleni, in press. Quotation in dialogue. In P. Saka & M. Johnson (eds.), *The pragmatics of quotation*. Springer.
Gregoromichelaki, Eleni & Ruth Kempson. 2015. Joint utterances and the (split-)turn taking puzzle. In J. L. Mey & A. Capone (eds.), *Interdisciplinary studies in pragmatics, culture and society*. Heidelberg: Springer.
Gregoromichelaki, Eleni, Ruth Kempson, Christine Howes, Arash Eshghi. 2013. On making syntax dynamic: The challenge of compound utterances and the architecture of the grammar. In I. Wachsmuth, J. de Ruiter, P. Jaeck & S. Kopp (eds.), *Alignment in communication: Towards a new theory of communication*. Amsterdam: John Benjamins.

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

Gregoromichelaki, Eleni, Yo Sato, Ruth Kempson, Andrew Gargett & Christine Howes. 2009. Dialogue modelling and the remit of core grammar. Proceedings of IWCS 2009.

Grice, Paul. 1975. Logic and conversation. *Syntax and Semantics*, 3(S41). 58.

Guhe, Markus. 2007. *Incremental conceptualization for language production*. New Jersey: Lawrence Erlbaum Associates.

Hale, John. 2006. Uncertainty about the rest of the sentence. *Cognitive Science* 30. 643672.

Healey, Patrick. 2008. Interactive misalignment: The role of repair in the development of group sub-languages. In R. Cooper & R. Kempson (eds.), *Language in flux: Language coordination, language variation, change and evolution*. London: College Publications.

Healey, Patrick, Matthew Purver & Christine Howes. 2014. Divergence in dialogue. *PLoS One* 9(2). 10.1371/journal.pone.0098598. http://qmro.qmul.ac.uk/xmlui/handle/123456789/7078

Healey, Patrick G.T. 1997. Expertise or expert-ese? The emergence of task-oriented sub-languages. Proceedings of the 19th Annual Conference of the Cognitive Science Society.

Hendriks, Petra. 2014. *Asymmetries between language production and comprehension*. Dordrecht: Springer.

Hilbrink, Elma, Herideth Gattis & Stephen Levinson. 2015. Early developmental changes in the timing of turn-taking: A longitudinal study of mother-infant interaction. *Frontiers in Psychology* 6, 1492. doi: 10.3389/fpsyg.2015.01492.

Honing, Henkjan & Willem Zuidema. 2014. Decomposing dendrophilia. *Comment on Toward a computational framework for cognitive biology: Unifying approaches from cognitive neuroscience and comparative cognition by W. Tecumseh Fitch. Physics of Life Reviews* 11(3). 375–376.

Hough, Julian. 2015. *Modelling incremental self-repair processing in dialogue*. Ph.D Queen Mary University of London.

Hough, Julian & Matther Purver. 2017. Probabilistic record type lattices for incremental reference processing. In S. Chatzikyriakidis & Z. Luo (eds.), *Modern perspectives in type-theoretical semantics*. New York: Springer.

Hough, Julian & Matthew Purver. 2014. Probabilistic type theory for incremental dialogue processing. In R. Cooper, S. Dobnik & S. Lappin (eds.), *Proceedings of the EACL 2014 workshop on Type Theory and Natural Language Semantics*, 80–88. Gothenburg: ACL 2016.

Hurley, Susan. 1998. *Consciousness in action*. Cambridge, MA: Harvard University Press.

Kalatzis, Dimitrios, Arash Eshghi & Oliver Lemon. 2016. Bootstrapping incremental dialogue systems: Using linguistic knowledge to learn from minimal data. Proceedings of the NIPS 2016 workshop on Learning Methods for Dialogue. Barcelona.

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. 1996. Semantics, pragmatics and deduction. In S. Lappin (ed.) *The handbook of contemporary semantic theory*, 561–598. Oxford: Blackwell.

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

Kempson, Ruth, Ronnie Cann, Arash Eshghi, Eleni Gregoromichelaki & Matthew Purver. 2015. Ellipsis. In S. Lappin & C. Fox (eds.), *Handbook of contemporary semantics*, 2nd edn. Oxford: Wiley-Blackwell.

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

Kempson, Ruth & Eleni Gregoromichelaki. in press. Action sequences instead of representational levels: Commentary on branigan & pickering. *Brain & Behavioral Sciences*.

Kempson, Ruth, Eleni Gregoromichelaki, Arash Eshghi & Julian Hough. forthcoming. Dynamics of ellipsis in situated interaction. In J. Craenenbroek & T. Temmerman (eds.), *Handbook of ellipsis*. Oxford: Oxford University Press.

Kempson, Ruth, Wilfried Meyer-Viol & Dov Gabbay. 2001. *Dynamic syntax: The flow of language understanding*. Oxford: Blackwell.

Kleist, Heinrich von. 1951. On the gradual construction of thoughts during speech. *German Life and Letters* 5(1). 42–46. Translated by Hamburger M.

Kobele, Gregory, 2012a. Eliding the derivation: A minimalist formalization of ellipsis. In S. Müller (ed.), *Proceedings of the HPSG 2012 Conference Chungnam National University Daejeon*, 307–324. Stanford: CSLI Publications.

Kobele, Gregory, 2012b. Ellipsis: Computation of. *Wiley Interdisciplinary Reviews: Cognitive Science* 3(3). 411–418.

Kobele, Gregory 2015. LF copying without LF. *Lingua* 166(part B). 236–259.

Kobele, Gregory, Sabrina Gerth & John T. Hale. 2013. Memory resource allocation in top-down minimalist parsing. In G. Morrill & MJ. Nederhof (eds.), *Formal grammar*. Lecture Notes in Computer Science, vol 8036. Berlin & Heidelberg: Springer.

Larsson, Staffan. 2003. The information-state approach to dialogue management. In D. Traum & S. Larsson (eds.) *Current and new directions in dialogue*, 325–353.

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

Lewis, Shevaun & Colin Phillips. 2015. Aligning grammatical theories and language processing models. *Journal of Psycholinguistic Research* 44(1). 27–46.

Lombardo, V. & P. Sturt 2002. Towards a dynamic version of TAG. Proceedings of the Sixth international workshop on tree adjoining grammar and related formalisms. Universit'a di Venezia.

Ludlow, Peter. 2005. A note on alleged cases of nonsentential assertion. In R. Elugardo & R. J. Stainton (eds.), *Ellipsis and nonsentential speech*, 95–108. Netherlands: Springer.

Marr, David. 1982. *Vision*. Cambridge, MA: MIT Press.

Marti, Luisa 2006. Unarticulated constituents revisited. *Linguistics and Philosophy* 29. 135. doi:10.1007/s10988-005-4740-4

Mazzei, Alessandro, Vincenzo Lombardo & Patrick Sturt. 2007. Dynamic TAG and lexical dependencies. *Research on Language and Computation* 5(3). 309–332.

Merchant, Jason. 2004. Fragments and ellipsis. *Linguistics and Philosophy* 27. 661–738.

Merchant, Jason. 2010. Three kinds of ellipsis. In F. Recanati, I. Stojanovic & N. Villanueva (eds.), *Context-dependence, perspective, and relativity*, 152. Berlin: Mouton de Gruyter.

Milkowski, Marcin. 2013. *Explaining the computational mind*. Cambridge, MA: MIT Press.
Mills, Gregory J. 2011. The emergence of procedural conventions in dialogue. *Cognitive Science Society Conference (COGSCI)* 33. 471476.
Mills, Gregory J. & Eleni Gregoromichelaki. 2010. Establishing coherence in dialogue: Sequentiality, intentions and negotiation. In P. Lupkowski & M. Purver (eds.), *Aspects of semantics and pragmatics of dialogue (SemDial)*, 1724. Poznan: Polish Society for Cognitive Science.
Morrill, Glyn. 2011. *Categorial grammar: Logical syntax, semantics, and processing*. Oxford: Oxford University Press.
Morrill, Glyn & Valentin Oriol. 2015. Multiplicative-additive focusing for parsing as deduction. In I. Cervesato & C. Schürmann (eds.), *First International Workshop on Focusing*, 29–54. Electronic Proceedings in Theoretical Computer Science series. DOI: 10,4204/EPCTS.197.4
O'Grady, William. 2005. *Syntactic carpentry: An emergentist approach to syntax.* London: Routledge.
Perruchet, Pierre & Arnaud Rey. 2005. Does the mastery of center-embedded linguistic structures distinguish humans from nonhuman primates? *Psychonomic Bulletin & Review* 12(2). 307–313.
Poesio, Massimo & Hannes Rieser. 2010. Completions, co-ordination and alignment in dialogue. *Dialogue and Discourse* 1. 1–89.
Progovac, Ljiljana. 2006. The syntax of nonsententials: Small clauses and phrases at the root. In L. Progovac, K. Paesani, E. Casielles & E. Barton (eds.), *The syntax of nonsententials: Multidisciplinary perspectives*, 33–71. Amsterdam: John Benjamins.
Purver, Matthew, Ronnie Cann & Ruth Kempson. 2006. Grammars as parsers: Meeting the dialogue challenge. *Research on Language and Computation* 4. 289–326.
Purver, Matthew, Eleni Gregoromichelaki, Wilfried Meyer-Viol & Ronnie Cann. 2010. Splitting the I's and Crossing the You's: Context, speech acts and grammar. Proceedings of the Semantics and Pragmatics of Dialogue(PozDial). Poznan, Poland. June 2010.
Ranta, Aarne. 2011. *Grammatical framework: Programming with multilingual grammars*. Stanford, CA: CSLI Publications.
Rieser, Hannes 2016. *A process algebra account of speech-gesture interaction*. MS. University of Bielefeld. DOI: 10.13140/RG.2.2.30166.24644
Sato, Yo. 2011. Local ambiguity, search strategies and parsing in dynamic syntax. In R. Kempson, E. Gregoromichelaki & C. Howes. *The dynamics of lexical interfaces*, 205–234. Stanford: CSLI Publications.
Schober, Michael & Susan Brennan. 2003. Processes of interactive spoken discourse: The role of the partner. In A. Graesser, M. Garnbacher & S. Goldman (eds.), *Handbook of discourse processes*, 123–164. Mahwah, NJ: Lawrence Erlbaum Associates.
Sehraku, Tohru. 2013. *Clefts, relatives, and language dynamics: The case of Japanese*. Oxford: DPhil.
Sober, Elliott & David Sloan. Wilson 1998. *Unto others: The evolution and psychology of unselfish behavior*. Cambridge, MA: Harvard University Press.
Sperber, Dan & Deirdre Wilson. 1995. *Relevance: Communication and cognition*, 2nd edn. Oxford: Blackwell.
Stabler, Ed. 1997. Derivational minimalism. In C. Retoré (ed.), *Logical aspects of computational linguistics*, 68–95. Dordrecht: Springer.

Stabler, Ed. 2011. Computational perspectives on minimalism. In C. Boeckx (ed.), *Oxford handbook of linguistic minimalism*, 617–642. Oxford: Oxford University Press.
Stabler, Ed. 2013. Two models of minimalist, incremental syntactic analysis. *Topics in Cognitive Science* DOI: 10.1111/tops.12031.
Steedman, Mark. 2000. *The syntactic process*. Cambridge, MA: MIT Press.
Stokhof, Martin & van Lambalgen. Michiel 2011. Abstractions and idealizations: The constructions of modern linguistics. *Theoretical Linguistics* 37(1–2). 1–26.
Tomasello, Michael 2008. *The origins of human communication*. Cambridge, MA: MIT Press.
Weir, Andrew. 2014. *Fragments and clausal ellipsis*. Diss. University of Massachusetts Amherst.
Wilson, David Sloan. 2002 *Darwin's cathedral: Evolution, religion and the nature of society*. Chicago: University of Chicago Press.
Wittgenstein, Ludwig. 1953. *Philosophical investigations*. Oxford: Blackwell.
Yanchao, Yu, Arash Eshghi & Oliver Lemon. 2016. Training an adaptive dialogue policy for interactive learning of visually grounded word meanings. Proceedings of the SIGDIAL 2016 Conference, The 17th Annual Meeting of the Special Interest Group on Discourse and Dialogue, 13-15 September 2016, Los Angeles, CA, 339–349.