

Affordance competition in dialogue: the case of syntactic universals

Eleni Gregoromichelaki
University of Duesseldorf
King’s College London
elenigregor@gmail.com

Stergios Chatzkyriakidis
University of Gothenburg
stergios.chatzkyriakidis

Arash Eshghi
Heriot Watt University
eshghi.a@gmail.com

Julian Hough
Queen Mary University of London
j.hough@qmul.ac.uk

Christine Howes
University of Gothenburg
christine.howes@gu.se

Ruth Kempson
King’s College, London
ruth.kempson@kcl.ac.uk

Jieun Kiaer
Oxford University
jieun.kiaer@orinst.ox.ac.uk

Matthew Purver
Queen Mary University of London
Jožef Stefan Institute, Ljubljana
m.purver@qmul.ac.uk

Mehrnoosh Sadrzadeh
University College London
m.sadrzadeh@ucl.ac.uk

Graham White
Queen Mary University of London
g.graham.white@gmail.com

Abstract

In this paper, we explore the idea that independently developed Dynamic Syntax accounts of dialogue and interaction fit well within the general approach of radical embodied and enactive accounts of cognition (REEC). This approach enables a rethinking of the grounding of linguistic universal constraints, specifically tree structure restrictions, as the outcome of affordance competition, a general REEC sociocognitive mechanism underpinning action selection. Given this subsumption, we argue such an approach opens up a whole new area of language-related dynamic systems research.

1 Introduction

Gregoromichelaki et al. (2019) claim that, in virtue of modelling natural languages (NLs) as actions licensing context-to-context transitions, Dynamic Syntax (DS) fits well within radical embodied and enactive accounts of cognition (REEC)

(Bruineberg et al., 2019; Paolo et al., 2018). At present, there is a considerable effort going on to provide accounts of so-called “higher cognition” processes, like memory and imagination, from an enactive perspective that eschews the necessity of representational content (see, e.g., Froese and Izquierdo (2018); Briglia et al. (2018); Hutto and Peeters (2018); Werning (2020); Bruineberg et al. (2019); cf. Gallistel and King (2009)). However, major trends within REEC still remain underdeveloped with respect to the fine-grained details of the traditionally considered par excellence representational medium, NL. While broad outlines of an account have been provided, what is still missing are more specific models of NL-related phenomena standardly conceived as low-level syntactic and semantic issues. Yet in the absence of any attempt to incorporate NL activity in all its complexity within the REEC account as a general nonmodular view of cognition,

its claims remain seriously incomplete. This paper seeks to do three things: First we set out the general embedding of the DS framework within a REEC perspective, introducing the argument that a grammar formalism incorporating psycholinguistic (or “performance”) insights and an externalist (inter)action-based view of semantic significance not only provides a natural basis for a model of conversational dialogue but also opens up the possibility of achieving an integration of current research in cognitive neuroscience and linguistic theorising. Secondly, regarding the finer details, we give a more detailed demonstration of how the DS framework, with its action-centric dynamic, has the potential to capture straightforwardly familiar morphosyntactic data for which static frameworks had to add auxiliary hypotheses of considerable complexity. But, on this basis, we can also go one step further towards explaining how the analysis of these phenomena carries over seamlessly when they are embedded in the dynamics of ongoing conversational dialogue. This is an advantage of the dynamic approach which is not available to static frameworks or to individualistic accounts both presuming a competence/performance dichotomy. We then address a putative counterexample to the REEC general claim of the dispensability of standard representational assumptions and the primacy of action. It concerns a robust structural constraint said to hold of all NLS ([Kempson et al., 2016 a.o.](#)). We argue that such a restriction does not necessitate a representational explanation and can be seen instead as grounded in a combination of socio-cognitive constraints and general properties of dynamical systems (see, e.g., [Silberstein and Chemero, 2012](#)) hence fully commensurate with the view that all aspects of cognition are grounded in action without mandatory invocation of brain-internal mental representation.

2 Rethinking the nature of NL grammars

Dynamic Syntax, as set out in [Kempson et al. \(2001\)](#); [Cann et al. \(2005\)](#) and much other work since, is a grammar architecture whose core basis is incremental integration of the contribution of word sequences within a surrounding landscape of affordances. Affordances, under this perspective, are relations between possibilities for action provided by the environment (including the social milieu) and agent abilities available in a ‘form of life’ ([Rietveld and Kiverstein, 2014 a.o.](#)). Within

this view, we take the grammar as an integral part of the sociocognitive environment into which humans are enculturated via natural dialogue and interaction. Thus the grammar, as it is being developed and constantly revised, is not seen as qualitatively different from other constraints influencing the environment-engaging behavioural patterns of individuals and groups in the achievement of their aims and goals. It is of course possible to reify these processes of linguistic engagement with the world and assign them abstract structure articulated in representational terms, as in the standard grammatical models in theoretical linguistics or models intended to explicitly teach linguistic skills. Indeed, folk theories of language and thought (at least in some societies) undeniably possess such conceptions of what language is, and we do not doubt that such conceptualisations affect linguistic behaviour. However, we believe that such conceptions are neither basic, nor universal (see, e.g. [Linell, 2005](#)). Instead, they constitute further (meta)affordances available within particular social groups for engaging with the available first order NL resources and constraining more systematically the relevant action possibilities.

Grammar – which for us includes what are standardly distinguished as syntax and semantics/pragmatics – constrains human (inter)action by providing a source of normativity, of what is right or wrong, of what makes sense or not, and, hence, what is perceived as rewarding or not, relative to particular social practices. For this reason, grammar is not a construct encapsulated within an individual brain or mind. Following Wittgenstein’s well-known arguments against the existence of “private languages” and against an interpretative conception of rule-following (see, e.g., [Wittgenstein, 1953](#)), we assume that normative constraints apply within a public domain of expression (even in cases where we privately rehearse responses to a simulation of such interactional challenges in the public domain).

Concomitantly, in the REEC affordance literature, a distinction is made between individual *abilities* which are non-representational capacities allowing individuals to perceive and pick up what opportunities and restrictions are available in their sociomaterial environment and *affordances* which are potential interactions with the environment available within practices or ‘forms of life’ independently of any particular individual agent.

Grammars, in our view, operate at the level of regulating actions within practices, both linguistic and non-verbal, and, for this reason, are part of the public landscape of affordances available to interacting agents in each particular instance of engagement with the environment. Whether agents interact solely with the physical environment or with other agents, their actions (which not only reveal but actually constitute their conceptualisations) are enabled and restricted by normative constraints. Such constraints which are non-representational ways regulating behavioral patterns (systems of habits) are imposed by the various cultural groups agents inhabit or wish to associate with. Due to membership in various such cultural groups, in each occasion of engagement, agents' actions in turn "enact" and hence modify or enrich the normative constraints available in the current practice.

Words, both as forms and meanings, and syntactic constructions are established patterns of actions that can be fitted in across various 'language games' to enact the nature of the current activity ('semantics') subject to normative judgements emanating from sources outside the acting agent (despite the fact that sophisticated agents can simulate internally such externally sourced judgements). This results in a system constantly in flux (Cooper, 2012) but with enough emergent stability in each particular occasion to underpin agent coordination in the service of various purposes.

In accordance with these assumptions, the DS syntactic engine, including the lexicon, is articulated in terms of goal-driven actions (see also van Benthem, 2011) accomplished either by giving rise to expectations of further action opportunities, by exploiting contextual resources, or by being abandoned as unviable in view of more competitive alternatives. Thus words, syntax, and morphology are all modelled as affordances, opportunities for (inter-)action, produced and recognised by interlocutors to perform step-by-step coordinated mappings from perceivable stimuli (phonological strings) to concept-constituting action patterns (routines, *macros*) or vice-versa (Gregoromichelaki et al., 2019).

The substance of the DS framework is given by a specialised dynamic modal logic (PDL, Propositional Dynamic Logic) whose state-transition language describes a process of gradual unfolding of a diagrammatically laid out relational struc-

ture modelling the landscape of salient affordances as a Directed Acyclic Graph (DAG), see Fig 1. DAG nodes (Interaction Control States, ICS) are in turn structured states modelling hubs leading further embedded paths of potential interpretations/productions. Interpretation or production relies on the process of actualising selected paths among the potential provided within local graph structures in the form of trees.

In formal terms, such path traversals at the intermediate level of DAG nodes are enacted and constrained by means of employing in the constraint language specialised modal operators that restrict the trajectories of parsing/generation to incrementally emerging tree-shaped routes. The emergent licensed trajectories are models of sets of constraints articulated within the logic of finite trees (LOFT, Blackburn and Meyer-Viol, 1994). Each node of a tree-structure model is in turn inhabited by graphs resembling feature structures whose attributes are modelled by modal operators bearing linguistically relevant labels like type, $\langle Ty \rangle$, treenode address, $\langle Tn \rangle$, and $\langle Fo \rangle$ for content. In current versions of DS, content values are Record Types (henceforth RT) originating in Type Theory with Records (TTR, Cooper and Ginzburg, 2015) under construction (Eshghi et al., 2013; Hough, 2015; Purver et al., 2011). RTs' inherent underspecification fits well with the requirement to model the potential for indefinite enrichment across various dimensions. RTs (standing for "concepts") can in turn be conceived as mini-grammars of the DS kind articulating affordances for engagement with aspects of the environment (Gregoromichelaki et al., 2019), or potential for interaction with others (Eshghi et al., 2017a; Eshghi and Lemon, 2014). In a parallel effort, Sadrzadeh et al. (2018) show how combining a Vector Space Semantics methodology with DS can model incremental construction of ad hoc concepts to resolve issues of ambiguity and underspecification, allowing for pervasive variability in NL interpretation.

In both these versions of DS, DAGs map out the potential transitions globally available for selection in a particular context of interaction while, locally, at each DAG node (ICS), the potential constructions and transitions are constrained by LOFT axioms that ground tree axiomatisation principles. As DS is a model-theoretic formalism (Pullum and Scholz, 2001), all the inferential

activity modelling incremental parsing/production is defined at the level of constraints articulated through the vocabulary of the relevant logics. The transformation of such PDL and LOFT descriptions is accomplished via general and lexicon-driven mappings. These model patterns of established and relatively stable sequences of basic actions (*macros*). Such macros effect the progressive projection of content, i.e., further opportunities for action, for emergent trajectories following the time-linear order of presentation of a linguistic string. The paths traversed by means of the LOFT modal operators, reflect the constrained structuring of physical and cognitive actions that take place during linguistic processing. Some of the outcomes of such actions can be reified as predicate-argument structures by assuming an external, God's eye, non-modal-logic point of view that abstracts from the process of inducing and traversing such structures. This is a useful and often needed perspective, both for theory construction and for practical human purposes like teaching, metalinguistic reflection and many others but, in our view, it should not obscure the most basic subpersonal level of processing from which it emerges.

2.1 Rethinking syntax as process

Bringing this framework to bear on current views within REEC, Gregoromichelaki et al. (2019) argue that DS is compatible with a view of NLs as activities, “*linguaging*”, rather than manipulation of internal knowledge structures that define arbitrary mappings from sound to propositional symbolic representations. This is because DS articulates state-transition mechanisms relative to an ever-evolving interaction context, meshing with nonverbal actions and, for this reason, without needing to only license fixed propositional contents associated with sentence strings. Crucial in this perspective is incorporating in the formal framework the social normativity expressed by the grammar. The native incrementality and action-orientation of DS, as demonstrated by Kempson et al. (2016) amongst others, is well-suited to implement the fine-grained modelling of conversational dialogue dynamics. In particular, we are able to capture the wholly fluent manner in which co-participants in a dialogue switch roles, share utterance responsibility, or effect repair, as the formalism permits the distribution of all syntactic and semantic de-

pendencies across more than one agent.

From this viewpoint, in dialogue, what can be considered as “complete sentences” may emerge through a sequence of fragmentary contributions, with each participant adding some fragment to a partial structure.

- (1) Alex: There. To the left.
 Hugh: Got it . . .
 Hugh: but it's heavy
 Eliot: without the lid?
 Hugh: if you fancy burning yourself.

In dialogue that is, employing “*fragments*” and switching of roles between speaking and hearing is widespread because not only the physical context but also the interlocutors' actions and attitudes contribute to the construction of meaning and the achievement of (joint) goals. In such circumstances, incomplete linguistic dependencies can act as prompts soliciting the interlocutor to provide not only physical objects but also utterance continuations:

- (2) A: The the ...
 B: Here.
 A: Thanks. And get a a
 B: mattock. For breaking up clods of earth
 [BNC (modified)]
- (3) Jack: I just returned
 Kathy: from
 Jack: Finland [Lerner 2004]

Each such contribution can add unproblematically to whatever partial structure has been set out so far, irrespective of whether or not what precedes it is a full sentence. The effect is one of a rich potential for interactivity between dialogue participants.

Moreover, fine-grained morphosyntactic and semantic constraints, with seemingly no functional motivation, apply across the exchange of turns and sharing of utterances so that the initiation of a turn by one participant constrains the options available for the next. For example, in the Romance languages, Modern Greek, and other language families, a well-known morphosyntactic restriction on clitic clusters¹ is the Person Case Constraint (PCC; Perlmutter, 1970; Bonet i Alsina, 1991). In its most general standard variant, the PCC states that

¹Clitic clusterings are cases in which weakened pronouns occur in a fixed sequence characteristically immediately preceding or following the verb.

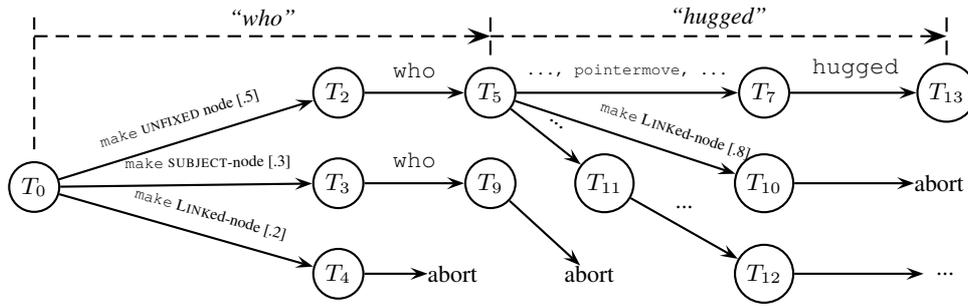


Figure 1: DAG: Who hugged



Figure 2: Introducing an 'unfixed' node



Figure 3: who-induced actions

a dative clitic cannot co-occur with a 1st/2nd person accusative clitic:

(4) A: *Irthe xtes o Giorgos ke*
 A: came yesterday the George and
tis to edose
 her_{CL-DAT} it_{CL-ACC} gave
 'A: Yesterday George came and he gave it to her'

(5) A: **O Giorgos tu se*
 the George him_{CL-DAT} you_{CL-ACC}
edixe.
 showed
 'A: George showed you to him' [Greek]

(6) A: **Le me ha dado*
 him_{CL-DAT} me_{CL-ACC} has given
 'A: S/He has given me to him.' [Spanish]

Notably, such morphosyntactic restrictions also apply across turns:

(7) A: *Irthe xtes o Giorgos ke*
 came yesterday the George and
tis...
 her_{CL-DAT}
 A: 'Yesterday George came and to her ...

B: *to edixe?*
 it_{CL-ACC} show
 B: 'he showed it to her ?'

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

The attempted continuation of the clitic cluster in the completion + confirmation request in (8) above is unacceptable because the PCC is violated. This shows that the grammar coordinates simultaneously and seamlessly the complementary and multifunctional actions of multiple agents while they are involved in interaction without imposing abstract licensing restrictions that concern single individuals acting on their own. For example, there is no requirement imposed by the grammar that turns need to consist of (overt or covert) complete sentences/propositions to be licensed.

However, as can be seen here, e.g. in (2) where the appropriate word is sought, the set of affordances in a context are perceived differentially by each agent or group of agents depending on their level of attunement to the relevant practices that constitute the source of such affordances, their skills, their attentional state, and previous experiences. Available affordances will

also be partitioned and probabilistically ranked depending on the current concerns and purposes of the agents involved. Agent-relevant affordances are called ‘solicitations’, i.e., a subset of the available affordances that stand out with respect to particular agents (Rietveld et al., 2018; Bruineberg et al., 2019). We extend this notion to apply to groups of agents. Such demarcated group solicitations within so-called ‘fields of affordances’ are modelled in DS by the time-linear unfolding of the DAG during conversational interaction, see Fig. 1 (see also Sato, 2011; Eshghi et al., 2013; Hough and Purver, 2012).²

Focussing attention at the local level of DAG nodes (ICSs), the first step is to consider the mapping between a string of words, in DS, a sequence of triggers for macros, and the DAG-transformations that it induces. The specialised PDL backbone of DS operates by means of prediction of available open paths of development constrained by so-called *requirements* (these are indicated by ? appearing in front of any available grammatical action). The constraints thus induced by the grammar (and other contextual factors) define a range of goals for the next steps of processing. Agents attuned to the grammatical practices available in a particular context can pursue the most relevant goal defined in this context guided by a process of affordance competition (Cisek and Thura, 2019). For example, a potential starting point of such a process can be a node state as displayed in Fig 2. Here we find the indicator of the current focus of attention, the pointer \diamond , and a prediction reflecting the expectation that a proposition (*?Ty(t)*) can be developed. This expectation will lead to many alternative predicted paths of achieving this development that can be displayed as shown in Fig 1. In Fig 2 we pursue one of those paths, the topmost path in the DAG, which is an option made available by the English grammar: a macro called **Adjunction* can introduce a radically structurally underspecified (‘unfixed’) node predicted to be needed to accommodate the processing of a content-underspecified element like *who* as seen in Fig 3. The latter processed first in a sentence can end up in multiple

²In order to simplify presentation, the available macros have been significantly condensed and schematically mentioned through the more central effects they induce; number values indicate toy probability rankings for particular paths; ellipsis (...) indicates that multiple steps have been omitted as they have been judged as irrelevant to the point we wish to make.

argument positions eventually so its contribution, a metavariable notated as **WH**, needs to be held in memory until a suitable place is found for it later in the process. Therefore, the identification of the node that accommodates the macro associated with *who* does not specify immediate dominance relations with respect to the root node, as shown in the illustration, Fig. 2-3. This is indicated by the dashed line. The underspecified relation is expressed via the modal operator $\langle \uparrow_* \rangle$ appearing in front of the root node identifier, $Tn(a)$ specifying that the relation between the two nodes is one of (non-immediate) dominance³.

As can be seen in the topmost path of the DAG, eventually the verb *hug* will be processed and its subject requirement will enforce the contribution of *who* to be specified as such. In DS this means that the unfixed node will now become fully specified with respect to all dominance relations. Once such a DAG path has been successfully traversed, the state reached will record a tree-structured topology of information with no words left to be parsed/generated any more (see Kempson et al., 2001, 2016; Hough, 2015, for details).

2.1.1 A universal constraint?

The constraint on development of such structures which is then proposed for all stages of every DS process is that there be only one copy of any node type at each individual stage of a particular DAG path. Indeed, it has been argued in major support of the framework that the abandonment of universal constraints on structure urged by some (Christiansen and Chater, 2008; Evans and Levinson, 2009; Bybee, 2010; Haspelmath, 2020) can be reversed by the shift to a dynamic system that operates with partial trees and their incremental introduction. This is because universal structural constraints can be shown to hold due to the inferential principles of LOFT, the constraint articulation basis of describing tree graph transitions in DS. So the proposed restriction is a general constraint which follows from the logic of partial tree construction: each node in a partial structure at any stage of the development

³ $\langle \uparrow_* \rangle$ is the modal operator expressing an immediate dominance relation, $\langle \uparrow \rangle$ is the operator which, using the Kleene * operator, expresses the weaker dominance relation. A yet more restricted variant is also defined requiring resolution of the under-specified dominance relation within a single sequence of functor relations, hence within a local predicate-argument structure

process has a unique identity with respect to other nodes in that structure; and this restriction applies equally to *structurally underspecified* nodes whose precise position in the emergent structure is yet to be established. Conceptually, this is a desirable property of the system because we assume that the information accumulated on a node is always underspecified and can be enriched at various stages of processing, even long after a node has been initiated. The restriction on a single uniquely identifying node address ensures that information encountered later in the processing of a string can have access to some particular already existing node, even if such a node is radically underspecified in terms of either position or other properties. Consequently, once a node 'unfixed' relative to particular domain has been introduced no other node can claim a matching underspecified identification. The effect is a "no-copy" principle for any node type. However, as we are going to see presently sometimes this restriction gets in the way, especially with free word-order languages that can exploit word order for information structuring purposes. In attempting to circumvent the effects of this restriction, linguistic patterns might appear that might give the impression of functionally unmotivated arbitrariness and taken as evidence for the "autonomy of syntax" and the particularity of the "language faculty". However, as we are going to argue such patterns do not provide any such evidence. Instead, the confluence of various other affordances and constraints explain the observed patterns.

2.2 Syntactic and morphosyntactic puzzles resolved

This dynamic of processing incorporating parsing and generation under the same formalism has been applied to a range of puzzles concerning syntactic and semantic/pragmatic phenomena as well as their interaction. For example, in verb-final languages, despite very free ordering of all constituents other than the verb, long-distance dependencies are subject to the strict constraint: more than one long-distance "movement" relation is licensed only if the items "dislocated" to the left periphery are co-arguments of some embedded verb. However, locally such co-arguments retain flexibility of ordering with respect to each other as in simple clauses (Pritchett, 1992; Koizumi, 2000):

(9) *Masami-ni prezento-o Kiyomi-ga*
 Masami_{DAT} prezento_{ACC} Kiyomi_{NOM}
katta
 bought [Japanese]
 Kiyomi bought a present for Masami

(10) *Masami-ni prezento-o John-ga*
 Masami_{DAT} prezento_{ACC} John_{NOM}
Kiyomi-ga katta-to omotteiru
 Kiyomi_{NOM} bought-COMP believed
 John thought Kiyomi bought Masami a present

This flexibility of ordering and the left periphery restriction is characteristic of the verb-final language pattern, by report the most common language pattern among natural languages (Dryer, 2013). But it also occurs in languages with more flexible word orders like Modern Greek where the verb along with the arguments can occur in any order. At first sight, it might appear that this freedom of ordering and piling up of coarguments in the left periphery will be as problematic for DS as it is for other frameworks (see e.g. Koizumi (2000)). If sequences of noun phrases may occur in any order, or, in the embedded case, the left periphery can host any number of arguments, there would appear to be, in either case, a violation of the "no-more-than-one" underspecified relation constraint.

However, as argued in (Kempson and Kiaer, 2010), within DS, there is a simple explanation of this puzzle. The DS system is able to distinguish degrees of underspecification across all grammatical notions, from syntactic to semantic domains. In the case of node position, fine-grained variants of 'unfixed' relations are defined depending on the domain within which the underspecification is confined. One such variant is the case of a node specified solely as dominated by the root node, so an underspecified relation within that global domain. Another variant that naturally arises due to the locality defined by a verb and its arguments is underspecification relative to a more restricted co-argumental domain. And naturally there are transitive implicative relations among such definitions with the global less specified transitions being able to be updated into the more locally-constrained ones (Cann et al., 2005). Interaction with other simultaneous grammatical constraints now solves the puzzle of free ordering.

The solution is provided by considering the role of case morphology. Case in DS has a crucial processing and semantic purpose, namely, the introduction and restriction of further possibilities

of action, instead of just being seen as a feature-matching reflex of some overarching syntactic structure. What this means for free constituent order in languages such as Japanese and Modern Greek is that the simple monoclausal sequence involves the iterated process of introducing an unfixed relation followed by an immediate update step to a specific fixed position provided by processing the case-marking. Once this latter step has been taken, it immediately enables the licensing again of a new underspecified relation this time concerning a newly introduced node. Moreover, this process can also apply within the province of a currently unfixed node with the more embedded underspecification variant concerning a more local domain than the one defined by the overarching underspecified node. However, what is not possible in such a case is that the processing of arguments from various clauses will intermingle. This is because the only underspecification that is available in such circumstances is the more local variety which can only be updated within a local domain. Thus the composite effect of excluding long-distance dependencies from multiple sources but with local ordering flexibility is then directly predicted through the licensed feeding relation between the two differently constrained types of underspecified relation. This modelling of local word order freedom with immediate update despite the lack of a predicate in verb-final languages is notably confirmed by experimental work establishing the incremental nature of Japanese sentence processing by [Witzel and Witzel \(2016\)](#) who demonstrate experimentally that both structural and semantic reference judgements are incrementally made before the verb is processed.⁴

Moreover, surprising confirmation of the restriction on underspecified dependencies and its potential evasion comes from its applicability to the superficially wholly unrelated morphosyntactic phenomenon of clitic clustering. As we saw earlier in (4)-(8) due to the PCC, arbitrary lexical gaps in clustering paradigms are attested. The restriction on unfixed nodes, namely the fact that no more than one treenode with the same address is possible, gives us a natural explanation of the PCC effects because it forces us to look at the observable properties, the morphology, and the interpretation of clitic morphemes within each particular

⁴See also ([Kiaer, 2007, to appear](#)) for extensive discussion and experimental evidence from Korean.

language with no presumption of any allegedly universal syntactic structuring that has to be imposed contrary to the observable facts.

The effects of case are involved again. Dative clitics are in general notoriously underspecified. In Romance, this underspecified nature of the dative can be traced back to Latin, where dative marking is characterised as ambiguous ([van Hoeske, 1996](#)) in the sense that, in and of itself, dative does not determine a fixed hierarchical position in the structure. Similar diachronic considerations also apply for other languages that show the same PCC effects, e.g. Greek ([Chatzikyriakidis, 2010](#); [Chatzikyriakidis and Kempson, 2011](#)). In DS terms now, given this underspecification of the contribution of the dative, the plausible assumption to make is that dative clitics are processed on an unfixed node, so as to allow variable interpretations of the dative-marked argument depending on the clausal context. Furthermore, 1st and 2nd person accusative clitics in Spanish and many other Romance languages are syncretised with the dative, i.e. the same morphological form is used for both the 1st and 2nd person accusative, as well as the 1st and 2nd person dative clitics. For this reason, 1st and 2nd person accusative clitics can also be considered as underspecified and, under DS assumptions, as processed on an unfixed node too.

Under these underspecification of form assumptions and given the “no copy” constraint, we can predict that any combination of a dative clitic with a 1st or 2nd person accusative clitic will be disallowed. This is because, as we said, two distinct unfixed nodes of the same type cannot be introduced in a single domain. As a result, the ordering Dative-1st/2nd person in, e.g., Modern Greek, will appear as unacceptable due to incompatible information coming from the two types of clitics accumulating on a single node, see (8). On the other hand, 3rd person accusative clitics are not syncretized with the dative and, furthermore, they are always interpreted as direct objects. In DS terms, this means that they are processed within a fully specified structure, rather than on an unfixed node. Thus, the “no copy” constraint does not interfere with the processing of combinations of a dative and a 3rd person accusative clitic, see (7).⁵

⁵There are a number of variants of the PCC, e.g. the weak PCC version, which allows combinations of 1st and 2nd person pronoun clitics, but this variability has been shown to be afforded by the formal machinery of the system as well ([Chatzikyriakidis and Kempson, 2011](#)).

The significance of these results concerning syntactic and morphosyntactic restrictions is that these are observable low-level facts which are usually taken as *sui-generis* and arbitrary without any grounding in human interaction, meaning, or processing, hence apparently warranting complex and unavoidable stipulation in distinct components of the grammar. Current individualistic linguistic theories justify this split on the basis of notions like competence-performance, modularity (Fodor, 1983), and Marrian computational vs algorithmic level distinctions (Marr, 1982; Steedman, 2000; Koble, 2012). Within the DS framework, in contrast, these patterns along with their interactional effects (see e.g. (8)) are seen to fall out in virtue of modelling syntax as the progressive shaping of the landscape of affordances for interpretation and production in the everyday coordination of action. Such normative morphosyntactic constraints are explainable within such a grammar because they can be seen as historically sedimented practices routinising the most frequently taken up processing paths (macros), whether as a parser or a producer. Such routinisations are not qualitatively distinct from the other grammatical or lexical constraints given the flexibility afforded by the grammar (as modelled by the various paths within the DAG) even within single instances of interaction. Conceptualised as normative constraints, such macros are necessarily independent from individual NL users and constitute affordances within the processing environment (the ‘form of life’). However, in order to be taken up appropriately, the processing agent needs to be attuned to their potential, i.e., they need to appear as ‘solicitations’ to the agents involved. But the possibility exists that such affordances are potentially inadequately grasped either because the agents do not possess the skills (attunement, ‘abilities’) required or because the concerns of individual agents and groups prioritise other normative considerations in the competition among which affordances to pursue. Hence the constant potential for innovation and change as well as flexibility and adaptability resulting from novel combinations of affordances pursued. Nevertheless, the availability of such macros in a practice-sharing ‘form of life’ allows agent coordination in particular interactional episodes due to their operation as manifestly joint relevant affordances (joint ‘solicitations’).

3 NL within the REEC perspective

There is a range of approaches within REEC usually grounded within dynamical systems models of socio-cognitive phenomena (Chemero, 2009). Most of those espouse non-representational accounts of perception and action, so-called “lower cognition” or “basic minds”. However, even advocates of radical enactive perspectives stop short from extending this approach to “higher order” cognition, especially language (see e.g. Clark, 2016; Hutto and Myin, 2012). Moreover, even accounts that aim to develop NL models compatible with dynamical and complex systems approaches like connectionism, due to the individualistic perspective they adopt, have suggested occasionally that neural network implementations strengthen the competence/performance distinction and support the assumption of symbolic representations, albeit of an emergent nature (Prince and Smolensky, 1997).

Against this view, a lot of current REEC research (Rietveld et al., 2018; Bruineberg et al., 2019; Paolo et al., 2018) aims to integrate all levels of cognitive activity within the enactive approach. And DS sides with this approach in that the constraints defined through the logic underlying the framework constitute additional affordances available in the sociomaterial environment of human interactions and are not confined within individual brain structures (Gregoromichelaki, 2018).

DS constraints concern the *process* of comprehension/production in a social context as aspects of general perception/action mechanisms implemented via sensorimotor feedback loops with the environment that do not necessarily engage representational constructs (even though reification of such subprocesses and representational abstraction is also considered possible).

Under such a view, all grammatical dependencies are able to function as word-by-word incremental coordinating devices, i.e., affordances, for either interlocutor. They can do that either by providing forward momentum in the conversation. For example, see (2) earlier for a determiner-noun dependency whose satisfaction completes an already initiated statement but also provides a platform for an enrichment and further specification of the expressed content. Similarly in (3) a preposition-prepositional complement dependency functions as a query speech

act indicator to induce the interlocutor to provide the completion as an answer to the question (Gregoromichelaki et al., 2011, 2013a). Moreover, grammatical constraints can disrupt what is expressible and in what manner by the interlocutor as in (8). In all these instances, grammatical dependencies can function coordinatively sufficiently and efficiently irrespective of whether or not coherent units like the standard representational constructs of ‘sentences’ or ‘propositions’ are ever derived or not:

- (11) Hester: It’s for me.
 Mrs Elton, the landlady: And Mr Page?
 Hester: is not my husband. But I would rather you continue to think of me as Mrs. Page. [The Deep Blue Sea (film)]⁶
- (12) Jem: Mary, whatever it is you think you know you mustn’t speak of it, not if you want to stay safe.
 Mary: says the horse-thief [Jamaica Inn, BBC Transcripts]
- (13) A: SOMEONE is keen.
 B: says the man who slept here all night [The A-Word, BBC Transcripts]

Under DS assumptions, in (11), we don’t need to invoke pro-drop in English to explain the missing subject of Hester’s utterance. Instead, we can implement the intuitive explanation that the subject is provided due to the fact that the utterance is a continuation of Mrs Elton’s interrogative non-sentential query in the previous turn. For both utterances, there is no reason to assume that speech acts are licensed only if interlocutors can derive fully specified propositional contents and/or sentential structures embedded under speech act predicates. Neither us, as a third party reader, nor Hester, nor Mrs Elton herself, have particular predicates in mind when encountering and understanding “And Mr Page?” in (11) since its affordances mesh with the rest of the affordances in our respective contexts and allow us to go on interacting with the utterance or its utterer. And it is precisely this lack of predicate, not only in the surface structure but also in the conceptual structure, that allows Hester to make use of the previous utterance to build her own which is thus fulfilling the

⁶Along with natural data, constructed data from literature, film scripts etc. are particularly relevant in this context as they show that such constructions are not “speech errors” that can be easily dispensed with.

pending affordance for specification that has been introduced by the verbless utterance.

In contrast, in (12)-(13), the previous utterance, which appears to be a fully-specified independent sentence and proposition, is taken over by the interlocutor at the next processing stage. As a result, it becomes embedded under a *verbum dicendi* thus losing its supposed independence as a “complete thought” and becoming instead part of a report that uses the original speaker themselves as the ‘animator’ (Goffman, 1981) of the message now subsumed under the authorial control of the second speaker but remaining, nevertheless, a joint construction. Hence the effectiveness as a hostile move of this construction in English. But to achieve the modelling of this effect it is essential that we do *not* assume that there is a *copy* of the original utterance embedded under the reporting verb and under the second speaker’s voice. Instead, the second speaker’s utterance employs the original speaker’s own words against them, and that is the whole point of the construction (Gregoromichelaki, 2018)

So, instead of propositional knowledge of rules and manipulation of representations, the abilities of individual agents for interacting efficiently with others under the guidance of grammatical constraints could be characterised as subpersonal mechanisms which allow access to a normative landscape of affordances. Access to this landscape is then mediated by inducing a range of predictive goals (solicitations) to be fulfilled by either interlocutor in the very next steps which they will be taking. The skills required to take advantage of solicitations in such a context do not presuppose any ‘rational’ high-order individualistic inference, standardly taken to be the basis of all successful human communication (Clark (1996); Sperber and Wilson (1995) and many others).

Instead, the task of selecting appropriate actions is taken over by affordance competition, which operates at the much lower level of sensorimotor contingencies (Cisek and Thura, 2019; Anderson, 2014; Rietveld et al., 2018). However, this presupposes that not only does the grammar incorporate processing features like incrementality and predictivity but also that the grammar provides a shared “workspace” (Kempen, 2014) for both production and comprehension to operate and interact (Gregoromichelaki et al., 2013b).

In addition, non-linguistic practices need to

be integrated with NL actions to contribute their qualitatively identical types of constraints (Gregoromichelaki, 2018). Evidence that morphology directly interacts with embodied situational affordances of every day action comes from elliptical constructions. For example, in case-rich languages such as German and Greek, elliptical fragments necessarily display the form suitable to what might have been a complete sentence formulated in response, even when there is no obvious antecedent of a verbal sentential form— in the German example in (14) the accusative form *den Arzt* seamlessly blends with other situational affordances to constrain future action to the effect that someone should call a doctor:⁷

- (14) *A and B see a woman lying on the floor:*
 A to B: Schnell, den Arzt/#der Arzt
*Quick, the doctor*_{ACC}/#*the doctor*_{NOM}

This shows that what is called an ‘antecedent’ in linguistic accounts of ellipsis need not have linguistic form, and can in fact be an action concurrent with the utterance, or even in the future of it as in commands like (14). Moreover, physical actions can combine with linguistic actions to seamlessly contribute constituents like subjects and objects to sentences (see, e.g., Slama-Cazacu, 1976):

- (15) while sitting on the piano: He doesn’t play
 [PLAYS TUNE]. But rather he plays –he does
 it better than I do– [PLAYS TUNE - SINGS]
 [Clark, 2016]

It seems then that NLS are not underpinned by distinctive mechanisms residing within individual minds encapsulated from the demands and contingencies of interaction with the sociomaterial environment.

3.1 Rethinking NL Autonomy

The question that then remains is how to think of alleged sui-generis properties of NL morphology that have been adduced as arguments for claims like “the autonomy of syntax” hypothesis. Within DS, this question is pertinent regarding the status of claimed universal NL processing constraints like the “no-copy” restriction. While this remains an issue for much further development, we suggest that, within a domain-general

framework like DS, there are grounds for seeing this restriction as a general control property of socio-cognitive coordination systems (control in the cybernetic sense of ‘regulation’, e.g. Bickhard (2009); Carver and Scheier (2012)) and, therefore, indeed a general property of physical systems involving the interaction of multiple simultaneous processes.

In this connection, Anderson (2014) attributes to individual brain mechanisms (‘abilities’) the function of action control via perception/action feedback loops with the environment without necessary representational mediation. Like DS, he sees “higher” cognitive processes like NL as developments of already existing brain capacities, rather than as species-specific cognitive adaptations. The brain is modelled as a connectionist network, superimposed on the physical brain structure, inspired by Smolensky (1986)’s architecture but without the individualistic representational interpretation originally assigned to this model. The state transitions and attractor landscapes of this network control the interaction of the individual and the environment without presuming that the brain constructs a model of the world it interacts with. As in REEC’s view of skills, abilities, and dispositions, the contribution of the individual brain is seen as complementary to the equally significant roles of the whole body and the environment. Brain activity on its own is in no way to be considered as enacting conceptually-transparent reactions to the environment. Some sequences of states can be assigned interpretations as individuated action goals but it is only at the end of a sequence of preparatory stages during which multiple partially activated goals coexist and interact with environmental inputs that such an abstract interpretation becomes possible. Accordingly, the states that the network passes through are not assigned representational contents but are individuated by their responsiveness to particular inputs and the behaviours that they enable. This control of behavior architecture also dispenses with a central decision module that processes abstract goals prior to action. Instead, different patterns of activation at each state track the gradual convergence of the totality of the control mechanism towards a particular action, that is, they track the “decision” process to pursue a selected subset of the available affordances that compete for realisation within an agent’s environment. What is im-

⁷For a critique of views that purport to reduce all fragments to a sententialist form of explanation, see Kempson et al. (2018).

portant for us here is that, at preliminary stages, affordance competition in Anderson's interpretation is implemented by the fact that neural patterns of activation might reflect simultaneously multiple partially-activated but incompatible options. This is only possible before final convergence to the selected action goal which can only consist of the pursuance of multiple but compatible affordances. Architecturally, the potential for the presence of incompatible options arises because various incompatible affordances can engage the activation of the same network nodes as long as they are not fully activated. At the final update stage of convergence (i.e. when a "decision" has been reached), the partial specifications of multiple action goals will be necessarily eliminated if their full activation patterns are incompatible, i.e., if they require the same network nodes for their realisation. The nonlinear update function ensures this. Only affordances with the strongest activation, having been gradually reinforced through environmental inputs, will survive. Their prevalence means that they will assume the role of controlling action as immediate agent goals while disrupting and inhibiting the activation patterns of any competitors.

We suggest that the 'no-copy' constraint on LOFT tree-path traversals is of the same type and aetiology. The graph structure imposed by LOFT instantiates an action control mechanism within the DS conception of NLS as grounded in goal-directed action. Restrictions such as the 'no-copy constraint' then provide a restricted architectural bottleneck to facilitate action selection in view of multiple competing action opportunities in the environment. To take a concrete example, we assume that an 'unfixed node' state can accommodate multiple affordances as goals. In the case of the Greek clitic clusters (see (4)-(8)), during production, multiple gender, person, and number specifications, compete for realisation at the sequential cluster positions. Each clitic morpheme at a particular cluster position when processed will trigger a combination of affordances (macros) appearing on a single treenode. But given the multiplicity of competing specifications, the eventual combination of affordance annotations on any treenode corresponding to a single cluster position will be the outcome of a convergence after affordance competition. In this sense, each treenode's specification of action goals is the equivalent of fully-specified activation patterns which

cannot be realised by neurons serving overlapping action specifications. Similarly, in DS, a node can only accommodate compatible selections of multiple macros as requirements for further update, for example, a macro associated with morphological marking for gender 'female' can combine with a macro for '3rd person singular' but not with a macro for gender 'masculine'. If inconsistent macros come to appear as goals at the same node the result will be anomaly (technically implemented in PDL as reaching a state where absurdity, \perp , holds or, equivalently, that the `ABORT` action is implemented). The reason is that there will be no possibility of pursuing the action paths indicated as subgoals by those macros and reaching a consistent state, hence the action control mechanism provided by the grammar will have failed its purpose.

As a result, in a language like Modern Greek with its otherwise generally free word order we do nevertheless find phenomena of strict ordering and epiphenomenally arbitrary positional exclusions. The processing reason for these phenomena though is that any underspecified clitic morpheme will need to be processed on an unfixed node whose specifications will be a particular combination of compatible affordances. Any other similarly underspecified clitic cannot then follow because its morphological specifications will annotate the same node leading to inconsistency. For this reason, whichever clitic morpheme is initially more strongly favoured by the contextual expressive needs of the interlocutor will win the competition among the various candidate combinations and will get to occupy the relevant cluster position thus excluding seemingly unaccountably the appearance of other morphemes not only in the current position but also in sequentially later ones. The only resort remaining for the expression of such a defeated and thus excluded combination of specifications will then be to appear not as a clitic but as a full NP, pronominal or nominal.

So, affordance competition is the underlying cause for action selection, action control, both facilitation and obstruction, with NL surface structural constraints being subsumed under an overall behaviour-guidance and general systems-control architecture.

In closing, we note that, besides capturing non-determinism and its structural effects in the updates achieved by morpho-syntactically induced

actions, a further integral task in this program is to capture the enormous range of interpretation possibilities which words display, a challenge often treated as peripheral for grammar frameworks. However, in our view, given the domain-generality of DS, the same mechanisms set out here can account for contextual modulation of word meaning in terms of updates selecting, through affordance competition, the relevant aspects of meaning of words in context. Words can be seen as triggering packages of sets of potential actions (macros) formalised as TTR types (Gregoromichelaki et al., 2019).

Accounting for contextual shift in meaning, current DS work on the incorporation of a vector-based semantics for the affordances words provide is ongoing, with the goal of modelling the pervasive variability of word meaning and its incremental resolution, without invariably invoking lexical ambiguity. This approach builds on previous work on compositional distributional semantics for pregroup grammars and the Lambek Calculus (Coecke et al., 2010, 2013), and for Combinatory Categorical Grammar (CCG) (Maillard et al., 2014; Moot, 2018). There is current work in this area, where distributional meanings of words are combined with their grammatical types, albeit more informally (e.g. Baroni and Zamparelli (2010)). But these do not consider an incremental setting with underspecified nodes in the grammatical construction, a challenge which our program addresses. In this vein, Purver et al. (forthcoming) suggest two distributional counterparts to underspecificational notions, and present experimental results regarding affordance competition that results in optimal choice in the disambiguation task

4 Conclusion

The shift of emphasis from static structural generalisations to actions is central here. The goal of NL processing and NL interaction is not to generate a sentence structure or proposition as such: it is to promote progressive context-updates via pursuing selected action opportunities (affordances) yielding further affordances guided by the various practices constraining and enabling the actions of the interlocutors. Given the DS architecture's domain-generality, constraints arising from non-verbal practices blend seamlessly with morpho-syntactic and semantic constraints. Context-updates ensuing from NL sequential sig-

nals follow constrained trajectories modelled as graph-transitions of restricted shapes, with an intermediate level of tree-structured paths. This structuring accounts for apparent idiosyncratic morphosyntactic constraints because it channels processing towards accomplishing interpretation and production in an incremental manner by restricting the available predictions. Considerable but not unconstrained flexibility is achieved due to various types of licensed underspecification and updates and action selection requirements are implemented despite the multiplicity of action opportunities in the environment. During the process of formulating or comprehending an utterance, the initial availability of various options for update is gradually restricted via incremental culling of options. This is carried out as constraints accumulate from various sources and at various stages, with a process of affordance competition serving as the necessary action selection mechanism. This mechanism replaces "high order" cognitive individualistic and internalist notions like intention recognition, mind-reading, and planning (see also Mills and Gregoromichelaki, 2010; Silberstein and Chemero, 2011; Froese, 2018) and implements relevance considerations at the low-level of subpersonal mechanisms of action control. Within this perspective, it is notable how this action control mechanism applies across sources of constraints bridging the levels of both "lower-" and "high-level" cognition.

We believe that, by abandoning representational individualism and grasping the potential that rethinking language as an action system offers, we can explain in an integrated way the dynamics of dialogue interaction and the seamless interweaving of NL utterances with nonverbal actions. Additionally, given that NL constraints are here subsumed under a theory of general action, the practices guiding NL users are a source of normativity independent of themselves, hence avoiding "private language" Wittgensteinian objections. On the other hand, the established practices guiding action need to be enacted afresh every time in new contexts, and either reinforced if successful or modified if ineffective (see, e.g., Eshghi et al., 2017b). We can then get a handle on both the apparent stability of a language for its users and yet its endless potential for ongoing variability and change. Furthermore, combining these perspectives, the ontogenetic trajectory of acquiring a first

language can then be seen not as maturational development executing a genetic program but as a grounded process crucially involving interaction with the social environment and scaffolded by non-verbal expressive and communicative joint actions like gesture and turn-taking (Clark and Kelly, to appear).

Acknowledgements

We wish to thank three anonymous reviewers for their careful reading and thought-provoking comments which led, in our minds, to significant improvements to this paper. We would also like to thank the Technical Programme Chairs for important practical assistance.

Chatzikyriakidis and Howes were supported by grant 2014-39 from the Swedish Research Council, (VR), which funds the Centre for Linguistic Theory and Studies in Probability (CLASP) in the Department of Philosophy, Linguistics, and Theory of Science at the University of Gothenburg. Howes was also supported by the VR grant 2016-0116 – Incremental Reasoning in Dialogue (IncReD). Purver is partially supported by the EPSRC under grant EP/S033564/1, and by the European Union’s Horizon 2020 programme under grant agreements 769661 (SAAM, Supporting Active Ageing through Multimodal coaching) and 825153 (EMBEDDIA, Cross-Lingual Embeddings for Less-Represented Languages in European News Media). The results of this publication reflect only the authors’ views and the Commission is not responsible for any use that may be made of the information it contains.

References

- M Eulalia Bonet i Alsina. 1991. *Morphology after syntax-pronominal clitics in romance*. Ph.D. thesis, Massachusetts Institute of Technology.
- Michael L. Anderson. 2014. *After Phrenology*. Cambridge University Press.
- Marco Baroni and Roberto Zamparelli. 2010. Nouns are vectors, adjectives are matrices: Representing adjective noun constructions in semantic space. In *Proceedings of the 2010 Conference on Empirical Methods in Natural Language Processing*, pages 1183–1193, Cambridge MA. Association for Computational Linguistics.
- Johan van Benthem. 2011. *Logical dynamics of information and interaction*. Cambridge University Press.
- Mark H Bickhard. 2009. The interactivist model. *Synthese*, 166(3):547–591.
- Patrick Blackburn and Wilfried Meyer-Viol. 1994. Linguistics, logic and finite trees. *Logic Journal of the Interest Group of Pure and Applied Logics*, 2(1):3–29.
- J. Briglia, P. Servajean, A. H. Michalland, L. Brunel, and D. Brouillet. 2018. Modeling an enactivist multiple-trace memory. ATHENA: A fractal model of human memory. *Journal of Mathematical Psychology*, 82:97–110.
- Jelle Bruineberg, Anthony Chemero, and Erik Rietveld. 2019. General ecological information supports engagement with affordances for ‘higher’ cognition. *Synthese*, 196(12):5231–5251.
- Joan Bybee. 2010. *Language, Usage and Cognition*. Cambridge University Press.
- Ronnie Cann, Ruth Kempson, and Lutz Marten. 2005. *The Dynamics of Language*. Elsevier, Oxford.
- Charles S Carver and Michael F Scheier. 2012. Cybernetic control processes and the self-regulation of behavior. In *The Oxford handbook of human motivation*, pages 28–42. Oxford University Press New York, NY.
- Stergios Chatzikyriakidis. 2010. *Clitics in four dialects of Modern Greek: A dynamic account*. Ph.D. thesis, King’s College, London.
- Stergios Chatzikyriakidis and Ruth Kempson. 2011. Standard Modern and Pontic Greek person restrictions: A feature-free dynamic account. *Journal of Greek Linguistics*, pages 127–166.
- Anthony Chemero. 2009. *Radical Embodied Cognitive Science*. MIT Press.
- Morten Christiansen and Nick Chater. 2008. Language as shaped by the brain. *Brain and Behavioral Sciences*, 31:489–509.
- Paul Cisek and David Thura. 2019. *Reach-to-Grasp Behavior*, chapter Neural circuits for action selection. Routledge.
- Andy Clark. 2016. *Surfing Uncertainty: Prediction, Action, and the Embodied Mind*. Oxford University Press, Oxford.
- Eve Clark and Barbara Kelly. to appear. Constructing a system of communication with gestures and words. In Goldin-Meadow and Morgenstern Aliyah, editors, *Gesture in language*. Mouton de Gruyter.
- Herbert H. Clark. 1996. *Using Language*. Cambridge University Press.
- Bob Coecke, Edward Grefenstette, and Mehrnoosh Sadrzadeh. 2013. Lambek vs. lambek: Functorial vector space semantics and string diagrams for lambek calculus. *Annals of Pure and Applied Logic*, 164(11):1079 – 1100. Special issue on Seventh Workshop on Games for Logic and Programming Languages (GaLoP VII).
- Bob Coecke, Mehrnoosh Sadrzadeh, and Stephen Clark. 2010. Mathematical foundations for a compositional distributional model of meaning. *Linguistic Analysis*, 36(1–4):345–384. Special Issue: A Festschrift for Joachim Lambek.

- Robin Cooper. 2012. Type theory and semantics in flux. In Ruth Kempson, Nicholas Asher, and Tim Fernando, editors, *Handbook of the Philosophy of Science*, volume 14: Philosophy of Linguistics, pages 271–323. North Holland.
- Robin Cooper and Jonathan Ginzburg. 2015. Type theory with records for natural language semantics. *The Handbook of Contemporary Semantic Theory*, pages 375–407.
- Matthew S. Dryer. 2013. Order of subject, object and verb. In Matthew S Dryer and Martin Haspelmath, editors, *The World Atlas of Language Structures Online*. Max Planck Institute for Evolutionary Anthropology, Leipzig.
- Arash Eshghi, Julian Hough, and Matthew Purver. 2013. Incremental grammar induction from child-directed dialogue utterances. In *Proceedings of the 4th Annual Workshop on Cognitive Modeling and Computational Linguistics (CMCL)*, pages 94–103, Sofia, Bulgaria. Association for Computational Linguistics.
- Arash Eshghi and Oliver Lemon. 2014. How domain-general can we be? Learning incremental dialogue systems without dialogue acts. In *Proceedings of SemDial 2014 (Dial-Watt)*.
- Arash Eshghi, Igor Shalymov, and Oliver Lemon. 2017a. Bootstrapping incremental dialogue systems from minimal data: the generalisation power of dialogue grammars? In *Proceedings of the 2017 Conference on Empirical Methods in Natural Language Processing (EMNLP)*.
- Arash Eshghi, Igor Shalymov, and Oliver Lemon. 2017b. Interactional Dynamics and the Emergence of Language Games. In *Proceedings of ESSLLI 2017*.
- Nicholas Evans and Stephen Levinson. 2009. The myth of language universals: language diversity and its importance for cognitive science. *Brain and Behavioral Sciences*, 32:429–492.
- Jerry A. Fodor. 1983. *The Modularity of Mind*. MIT Press.
- Tom Froese. 2018. Searching for the conditions of genuine intersubjectivity: from agent-based models to perceptual crossing experiments. In Albert Newen, Leon De Bruin, and Shaun Gallagher, editors, *The Oxford handbook of 4e cognition*, pages 163–186. Oxford University Press, Oxford, UK.
- Tom Froese and E. J. Izquierdo. 2018. A Dynamical Approach to the Phenomenology of Body Memory: Past Interactions Can Shape Present Capacities Without Neuroplasticity. *Journal of Consciousness Studies*, 25(7-8):20–46(27).
- C. Randy Gallistel and Adam Philip King. 2009. *Memory and the Computational Brain*. Wiley-Blackwell, Oxford, UK.
- E. Goffman. 1981. *Forms of Talk*. University of Pennsylvania Press.
- Eleni Gregoromichelaki. 2018. Quotation in dialogue. In *The Semantics and Pragmatics of Quotation*, pages 195–255. Springer.
- Eleni Gregoromichelaki, Ronnie Cann, and Ruth Kempson. 2013a. On coordination in dialogue: subsentential talk and its implications. In Laurence Goldstein, editor, *On Brevity*. OUP.
- Eleni Gregoromichelaki, Christine Howes, and Ruth Kempson. 2019. Actionism in syntax and semantics. In *Gothenburg Proceedings of Conference on Dialogue and Perception*. CLASP.
- Eleni Gregoromichelaki, Ruth Kempson, Christine Howes, and Arash Eshghi. 2013b. On making syntax dynamic: The challenge of compound utterances and the architecture of the grammar. In Ipke Wachsmuth, Jan de Ruiter, Petra Jaecks, and Stefan Kopp, editors, *Alignment in Communication: Towards a New Theory of Communication*, Advances in Interaction Studies (series editors: Kerstin Dautenhahn, Angelo Cangelosi), pages 57–85. John Benjamins.
- Eleni Gregoromichelaki, Ruth Kempson, Matthew Purver, Greg J. Mills, Ronnie Cann, Wilfried Meyer-Viol, and Patrick G. T. Healey. 2011. Incrementality and intention-recognition in utterance processing. *Dialogue and Discourse*, 2(1):199–233.
- Martin Haspelmath. 2020. Human linguisticity and the building blocks of languages. *Frontiers in Psychology*, 10:1–10.
- Willy van Hoeke. 1996. The Latin dative. In William van Belle and Willy Langendonck, editors, *The Dative Volume 1: Descriptive Studies*, pages 3–38. John Benjamins.
- Julian Hough. 2015. *Modelling Incremental Self-Repair Processing in Dialogue*. Ph.D. thesis, Queen Mary University of London.
- Julian Hough and Matthew Purver. 2012. Processing self-repairs in an incremental type-theoretic dialogue system. In *Proceedings of the 16th SemDial Workshop on the Semantics and Pragmatics of Dialogue (SeineDial)*, pages 136–144, Paris, France.
- Daniel D. Hutto and Erik Myin. 2012. *Radicalizing Enactivism: Basic Minds without Content*. MIT Press.
- Daniel D. Hutto and Anco Peeters. 2018. The Roots of Remembering: Radically Enactive Recollecting. In Kourken Michaelian, Dorothea Debus, and Denis Perrin, editors, *New Directions in the Philosophy of Memory*, pages 97–118. New York: Routledge.
- Gerard Kempen. 2014. Prolegomena to a neurocomputational architecture for human grammatical encoding and decoding. *Neuroinformatics*, 12(1):111–142.
- R. Kempson and J. Kiaer. 2010. Multiple long-distance scrambling: Syntax as reflections of processing. *Journal of Linguistics*, 46(01):127–192.
- Ruth Kempson, Ronnie Cann, Eleni Gregoromichelaki, and Stergios Chatzikiriakidis. 2016. Language as mechanisms for interaction. *Theoretical Linguistics*, 42(3-4):203–275.
- Ruth Kempson, Eleni Gregoromichelaki, Arash Eshghi, and Julian Hough. 2018. Ellipsis in Dynamic Syntax. In Jeroen van Craenbroeck and Tanya Temmerman, editors, *Oxford Handbook of Ellipsis*. Oxford University Press.
- Ruth Kempson, Wilfried Meyer-Viol, and Dov Gabbay. 2001. *Dynamic Syntax*. Blackwell.
- Jieun Kiaer. 2007. *Processes and Interfaces in Syntactic Theory: the case of Korean*. Ph.D. thesis, King’s College London.

- Jieun Kiaer. to appear. *Pragmatic Syntax*. Bloomsbury.
- Gregory M Kobele. 2012. Eliding the derivation: a minimalist formalization of ellipsis. In *Proceedings of the 19th International Conference on HPSG*.
- Masatoshi Koizumi. 2000. String vacuous overt verb raising. *Journal of East Asian Linguistics*, 9:227–85.
- P. Linell. 2005. *The Written Language Bias in Linguistics: Its Nature, Origins and Transformations*. Psychology Press.
- Jean Maillard, Stephen Clark, and Edward Grefenstette. 2014. A type-driven tensor-based semantics for CCG. In *Proceedings of the EACL 2014 Workshop on Type Theory and Natural Language Semantics (TTNLS)*, pages 46–54, Gothenburg, Sweden. Association for Computational Linguistics.
- David Marr. 1982. *Vision: A Computational Approach*. Freeman & Co, San Francisco.
- G. Mills and E. Gregoromichelaki. 2010. Establishing coherence in dialogue: sequentiality, intentions and negotiation. In *Proceedings of SemDial (PozDial)*.
- Richard Moot. 2018. Combining logical and distributional methods in type-logical grammars. *Journal of Language Modelling*, 6:288–317.
- Ezequiel A. Di Paolo, Elena Clare Cuffari, and Hanne De Jaegher. 2018. *Linguistic Bodies: The Continuity between Life and Language*. MIT Press.
- David M Perlmutter. 1970. Surface structure constraints in syntax. *Linguistic inquiry*, 1(2):187–255.
- Alan Prince and Paul Smolensky. 1997. Optimality: From neural networks to universal grammar. *Science*, 275(5306):1604–1610.
- Bradley L. Pritchett. 1992. *Parsing with Grammar: Islands, Heads, and Garden Paths*, pages 321–349. Springer Netherlands, Dordrecht.
- Geoffrey Pullum and Barbara Scholz. 2001. On the distinction between model-theoretic and generative-enumerative syntactic frameworks. In *Proceedings of the 4th International Conference on Logical Aspects of Computational Linguistics (LACL)*, pages 17–43. Springer.
- Matthew Purver, Arash Eshghi, and Julian Hough. 2011. Incremental semantic construction in a dialogue system. In *Proceedings of the 9th International Conference on Computational Semantics*, pages 365–369, Oxford, UK.
- Matthew Purver, Mehrnoosh Sadrzadeh, Ruth Kempson, Gijs Wijnholds, and Julian Hough. forthcoming. Incremental composition in Distributional Semantics. *Journal of Logic, Language and Information*.
- Erik Rietveld, Damiaan Denys, and Maarten Van Westen. 2018. Ecological-Enactive Cognition as Engaging with a Field of Relevant Affordances. In *The Oxford Handbook of 4E Cognition*, page 41. Oxford University Press.
- Erik Rietveld and Julian Kiverstein. 2014. A rich landscape of affordances. *Ecological Psychology*, 46.
- Mehrnoosh Sadrzadeh, Matthew Purver, Julian Hough, and Ruth Kempson. 2018. Exploring semantic incrementality with dynamic syntax and vector space semantics. In *Proceedings of the 22nd Workshop on the Semantics and Pragmatics of Dialogue (SemDial 2018 - AixDial)*, Aix-en-Provence.
- Yo Sato. 2011. Local ambiguity, search strategies and parsing in Dynamic Syntax. In E. Gregoromichelaki, R. Kempson, and C. Howes, editors, *The Dynamics of Lexical Interfaces*. CSLI Publications.
- Michael Silberstein and Anthony Chemero. 2011. Dynamics, Agency and Intentional Action. *Humana Mente*, 4(15):1–19.
- Michael Silberstein and Anthony Chemero. 2012. Complexity and Extended Phenomenological-Cognitive Systems. *Topics in Cognitive Science*, 4(1):35–50.
- T. Slama-Cazacu. 1976. Nonverbal components in message sequence: ‘mixed syntax’. In W.C. McCormack and S.A. Wurm, editors, *Language and Man: Anthropological Issues*. Mouton, The Hague, pages 217–227.
- Paul Smolensky. 1986. *Parallel distributed processing*, volume 2, chapter Neural and conceptual interpretation on PDP models. MIT Press, Cambridge.
- Dan Sperber and Deirdre Wilson. 1995. *Relevance: Communication and Cognition*, second edition. Blackwell.
- Mark Steedman. 2000. *The Syntactic Process*. MIT Press, Cambridge, MA.
- Markus Werning. 2020. Predicting the past from minimal traces: Episodic memory and its distinction from imagination and preservation. *Review of Philosophy and Psychology*, 11(2):301–333.
- L. Wittgenstein. 1953. *Philosophical Investigations, Trans. G.E.M. Anscombe*. Oxford: Blackwell.
- Jeffrey Witzel and Naoko Witzel. 2016. Incremental sentence processing in Japanese: a maze investigation into scrambled and control sentences. *Journal of Psycholinguistic Research*, 45:475–505.