

Long-Distance Paradox and the Hybrid Nature of Language

Guillermo Lorenzo, ¹✉

Phone 00-34-985104653

Email glorenzo@uniovi.es

¹ Department of Spanish Philology / Linguistics, Faculty of Humanities, University of Oviedo, C/ Amparo Pedregal s/n, E-33011 Oviedo, Spain

Received: 7 February 2018 / Accepted: 18 August 2018

Abstract

Non-adjacent or long-distance dependencies (LDDs) are routinely considered to be a distinctive trait of language, which purportedly locates it higher than other sequentially organized signal systems in terms of structural complexity. This paper argues that particular languages display specific resources (e.g. non-interpretive morphological agreement paradigms) that help the brain system responsible for dealing with LDDs to develop the capacity of acquiring and processing expressions with such a human-typical degree of computational complexity. Independently obtained naturalistic data is discussed and put to the service of the idea that the above-mentioned resources exert their developmental role from the outside, but in compliance with other internal resources, ultimately compounding an integrated developmental system. Parallels with other human and nonhuman developmental phenomena are explored, which point to the conclusion that the developmental system of concern can be assimilated to cases currently been conceptualized as ‘cue-response systems’ or ‘developmental hybrids’ within the ecological-developmental paradigm in theoretical biology. Such a conclusion is used to support the idea that both current externalist and internalist concepts fall short of a correct characterization of language.

Keywords

Externalism

Internalism
 Eco-devo biology
 Cue-response mechanisms
 Developmental hybrids

Introduction

Long-distance—or non-adjacent—dependencies (LDDs) are pervasive in language, manifest in such diverse phenomena as discontinuous phrases across embedded constituents (1a), pronominal and anaphoric co-reference (1b), wh-word/gap or NP/gap connections in relative clauses (1c), interrogative (1d), and passive constructions (1e), etc. (Zushi 2013)¹:

(1)

- a. The man was saying [that his wife had been shot] with a desperate voice
- b. The woman₁ that you₂ visited is worried about herself₁/your₂ sons
- c. The man who₁ she visited t₁ yesterday is completely out of control
- d. Who₁ did you visited t₁ last night?
- e. She₁ was declared guilty t₁

In many accounts, LDDs are taken to be the hallmark of language uniqueness, as compared to the cognitive makeup of nonhuman species (Corballis 2011), or even in the context of human cognition (Hauser et al. 2002). The rationale for this tenet is that LDDs reveal that the powers the computational the powers of the computational resources of humans or language exceed anything hitherto known in other species or abilities. To begin with, the capacity for nested embedding—as in (1a), which illustrates the kind of recursion customarily associated to context-freeness, is widely believed to be lacking in nonhuman cognition—see below.

Among LDDs, special mention should be given to the family of cases routinely referred to as ‘agreement’ phenomena, in which an item adopts a morphological appearance that mirrors the feature composition of another designated one. Agreement covers, from an interlinguistic point of view, an extremely wide range of related phenomena, which may be epitomized by the relatively simple case of subject agreement in English. In this language, as in many others, it is common for the verb and the agreeing subject to appear as adjacent units in the linear

order of utterances—as in (2a); however, (2b) clearly illustrates the LDD character of the relation:

(2)

- a. The girl₁ wants₁ to know who is afraid of her
- b. The girl₁ who is afraid of the boy wants₁ to know who is afraid of her

Crucially, note that in (2b) there is a matching item that is closer (actually, adjacent) to the relevant verb—namely, the boy, which however does not qualify as a better candidate for agreement than the six-item-distant correct one—the girl. Note, also, that (2b) further illustrates that different agreement relations may concur and interfere in a single utterance—namely, the girl is also the semantic controller of the abstract subject of to know, as well as the antecedent of the pronoun her. In each of these different instances of LDD, the second relatum is increasingly more distant from the first one, without hindering the interpretation of the whole utterance.

These kinds of observations have served to rank natural language in a relatively high position within the standard Chomsky hierarchy of computational complexity (Chomsky 1956): Namely, an area above the context-free recursive languages—i.e. those that accept discontinuous dependencies created by embedding segments within larger segments, since natural language accepts dependencies that hold across segments belonging to different levels of embedding. (3a) offers a suitable illustration, which is shown as an abstract pattern in (3b)—different capital letters represent that units belong to different embedded segments:

(3)

- a. The girl_{1/3} who said that the boy₂ hated her₃ loves₁ him₂
- b. X_{1/3} ... Y ... Z₂ ... Z₃ ... X₁ X₂

The reason why context-freeness is customarily seen as critical in the literature on comparative cognition is the more or less consensual claim that no clear cases of bona fide structural embedding have to date been found in any example of nonhuman sequential behavior whatsoever, while human language appears to be situated even some steps ahead of such a level of computational complexity. Context freeness thus represents a kind of chasm between human and nonhuman cognition (see Fitch and Hauser 2004; van Heijningen et al. 2009; Beckers et al. 2012).

Agreement dependencies are worth special mention in this context because, unlike other kinds of LDDs, they exhibit the peculiar trait of not being interpretable, in the sense that they make no specific contribution to the compositional meaning of the sentence (Chomsky 1995). Rather, agreement merely involves associating some morphological material to an item, which mirrors the (interpretable) feature composition of another (potentially distant) item—e.g. –s in not required to contribute the meaning ‘subject, 3rd person, singular’ in (2a), which is already contributed by the constituent the girl itself. For this reason, it is also commonly said that agreement is an asymmetric relation, since it is the verb that agrees with the subject, but not the other way around; in other words, the subject is a meaningful constituent, but the subject-agreement bearer item affixed to the verb is meaningless. The fact that agreement features are in many cases uninterpretable replicas of interpretable ones is certainly confounding (Adger and Svenonius 2011). Some features function both within interpretable and uninterpretable bundles in many languages; moreover, homologues of certain features may be interpretable in some languages and uninterpretable in some others (Van Valin 2003). In any event, the kind of features and morphology that I am interested in here are those displayed in LDDs without any conceptual import.

Such property raises an important question, to which no clear, consensual answer exists to date: Why are agreement relations so pervasive, if they do not appear to make any clear meaningful contribution to utterances? This is one of the main issues on which this paper will focus. I shall specifically defend that neither cognitive/functionalist approaches (Barlow and Ferguson 1988) or formalist/generativist counterparts (Boeckx 2006) have hitherto provided an account of the ultimate *raison d’être* of agreement systems in natural language. In a nutshell, functionalist/cognitive approaches try to make sense of agreement by ascribing some communicative role to its bearers (e.g. fixing discourse referents, creating noise-preventing redundancies, etc.), while formal/generativist approaches conceptualize it as a kind of imperfection that the language system needs to get rid of when processing expressions. The alternative that I shall suggest is a developmental one, which is not incompatible with prior suggestions along the lines of those referred to above, but which is more basic and closer to a ‘first cause’ for the existence of such linguistic entities.

Another point of interest of this paper has to do with a further characteristic of agreement relations, which illustrates an aspect of language in a way that other kinds of LDDs do not, or, at least, not as straightforwardly. The ways in which agreement relations instantiate in different languages are manifold and sensitive to a very wide range of criteria. To cite a few, agreement may be sensitive to a

particular syntactic relation—e.g. subject agreement, as in (2) above; or sensitive to a particular morphological feature—e.g. number agreement, as in (4), where the affix may indistinctly refer to the subject or the object, which must be a plural one; or sensitive to complex combinations thereof—as in (5), where the affix responds to a person/number hierarchy (1st/2nd > 3rd pl. > 3rd sg.), but only when the agent is focused and the transitive verb turned into an intransitive:

(4) Georgian (from Nevins 2011)

g-xedav-t.

2obj-saw-pl

‘I saw you all; we saw you all; he saw you all; we saw you’ (indistinctly)

(5) Quichean (from Preminger 2014)

a. ja yīn x-**in**-ax-an ri achinfoc me asp-1sg.abs-hear-af the man

‘It was me that heard the man’

b. ja ri achin x-**in**-ax-an yīn

foc the man asp-1sg.abs-hear-af me

‘It was the man that heard me’

Such variable and convoluted patterns (Corbett 2009) are difficult to justify on conceptual or communicative grounds, which makes the *raison d’être* of agreement phenomena even more intriguing.

At first sight, the extreme language-particular specificity of agreement systems invites one to think of them as being closer to the historical, traditional, or cultural layer of languages than to the evolved linguistic phenotype (Balari and Lorenzo 2013; Bickerton 2014). In any event, the main claim to be made in this paper is that agreement systems actually stand in a middle ground between the historically constructed and the biologically evolved, somehow belonging to both, making them strong evidence to support the ultimately hybrid nature of language that will be defended in the pages to come. Note from the onset that by ‘hybrid’ I mean something that is not entirely biological or entirely cultural, nor a simply additive admixture of biology and culture—thus avoiding the conflictive stance that Susan Oyama refers to as ‘conventional interactionism’ (Oyama 2000).

In the next sections, I shall unfold the following (two step) argument. Firstly, I shall try to show that agreement systems have a primarily developmental motivation, and, with this aim in mind, I shall draw upon certain intriguing parallelisms with the functioning of other, non-language devoted sensorimotor systems—namely, those responsible for gait control. After that, I shall focus on the fact that agreement-bearer units are self-subserved by the language mechanism, contrary to what can be observed in these other cognitive domains, where resources with a comparable role are provided by the environment and recruited on the fly by the system. This contrast will give grounds to the idea of conceptualizing agreement systems and the language-devoted computational mechanism as being part and parcel of a single, integrated organic unit, made up of both of external and internal contributors. The take home message will thus be that both current externalist (e.g. Christiansen and Chater 2008) and internalist (e.g. Chomsky 1986) approaches fall short of a correct characterization of the ultimately hybrid character of language.

From Kinesia to Long-Distance Paradox

Kinesia paradox (KP) is the name given to a condition that may affect people with motoric disorders like Parkinson's disease (PD), who on occasion appear to recover their capacity for executing complex motor sequences. This mostly happens when subjects enter surfaces with traverse lines drawn on the ground—e.g. a tiled floor. A cogent explanation of KP has been provided, based on the fact that while PD patients are highly dependent on visual feedback for their motoric performance, they also suffer from visual distortions that continuously put them at the risk of stumbling and falling. Consequently, when salient signals within the visual field become more markedly perceivable, PD patients can make use of them to counterbalance their otherwise deficient visuo-motoric feedback. Such an explanation has paved the way for the implementation of devices that, by projecting virtual marks on the floor while neurological patients walk, have proved very effective in the improvement of gait (Prothero 1993; Baram 1999, 2013, 2017; Palacios-Navarro et al. 2016). Interestingly, similar effects have also been obtained by means of auditory stimuli (Baram and Miller 2007).

Within this context, the main point that I would like to raise in this section is that KP provides a good model for making sense of a condition observed in cases of atypical language development, which I shall refer to as 'long-distance paradox' (LDP). The population potentially exposed to this condition are children with Specific Language Impairment (SLI), actually a family of rare developmental disorders that mostly affect the language capacity, in the absence of other non-linguistic deficits that could explain them (Bishop 1997; Leonard 1998).

Difficulties with the productive morphology of the linguistic system that the

child is acquiring—including agreement morphology, and with the establishment of LDDs, are among the most common traits of the linguistic symptomatology associated to SLI. What I refer to as LDP here specifically refers to the observation that after supervised training with agreement relations, children with SLI appear to significantly improve their performance with LDDs generally. Obviously, agreement relations are themselves LDDs. However, as I shall explain in the following paragraphs, the former exhibit the distinctive property, as compared to the more common forms of LDDs, of not being interpretable (see above), which in my opinion makes them apt to exercise the kind of causal impact that I am interested in.

The main source of evidence of LDP to date is to be found in Clahsen (1986), Clahsen et al. (1997), and Clahsen and Hansen (1997). In this series of papers, Clahsen and coworkers report that German speaking children with SLI experience difficulties both with the correct matching of person and number features for verb-subject agreement, and with some characteristic LDDs of German, such as the so-called V2 phenomenon. V2 refers to the fact that in German inflected verbs are systematically placed in second position in the sentence, irrespective of which constituent appears at the beginning, be it the subject, an emphatic object or adverb, a wh-phrase, etc. Grammarians agree on the verb-final character of German, which is witnessed by sentences in which the inflected element is not the verb proper, but an auxiliary, as then the former remains in situ and the latter shows up in the second position—(6a). Thus, according to the standard interpretation, in sentences like (6b), where no auxiliary exists, the verb is displaced to the second position, leaving a trace in the final one:

(6) In the printed proof, there is no upper space

a. Fußball haben die Kinder den ganzen Tag gespielt

football have-3pl the kids the whole day played.

‘Football, the children have played all day’

b. Fußball spielten_i die Kinder den ganzen Tag t_i

football play-3pl the kids the whole day.

‘Football, the children played all day’.

Children with SLI have problems to choose the correct inflected verb from in sentences like (6b)—incorrectly using e.g. spielte ‘3rd person, singular’, or

spieltet ‘2nd person, plural’, and locating the verb in second or final position at random—thus frequently producing incorrect sentences like * Fußball der Kinder den ganzen Tag spielte (* Football, the children-pl all day played-sing), in which both errors concur. Clahsen and coworkers’ most intriguing finding is that by improving their performance with the agreement system after intensive training, they also spontaneously improve their performance with the V2 phenomenon, but not the other way around. Such a conclusion gives support to the idea that there exists a causal link that runs from agreement to V2.

It is important to note that V2, like most LDDs, is a meaningful phenomenon, since it is the formal reflex of displacement as a means for emphasizing constituents or creating operator-variable chains in general; however, as already established, agreement is meaningless. Consequently, the idea immediately comes to mind that the basic role to be ascribed to agreement is not of a conceptual, informative or communicative kind, but one that relates with the kind of facilitating effect that is observed in rehabilitation contexts like the one described by Clahsen’s team. My specific claim is that agreement markers have a primordially developmental role, which consist of exciting the system in charge of the identification and acceptance of LDDs—in all likelihood, one associated with working-memory (see below), by means of a kind of self-training ancillary mechanism (see [Balari and Lorenzo](#) add 2015 ... 2018, where the idea is tentatively suggested). Maybe due to their well known difficulties in perceiving items with weak or light phonological body—affixes, clitics, etc. (Leonard et al. 1992), the mechanism is not fully functional in the case of children with SLI. As a result, LDDs become a very challenging and enduring problem for them. The observation that unimpaired children appear to fully master typical LDD constructions, like relative clauses, only after they have fully acquired the agreement paradigm—see, for the case of German, Mills (1985), may be read as further evidence of the suggested developmental role of agreement systems.

AQ1

Before proceeding, let me clarify that my interpretation of the facts above is not fully coincidental with that of Clahsen and coworkers. According to them, the locus of the impairment of the population that they studied is agreement itself; consequently, they speak of a ‘grammatical agreement deficit’ as the core affectation of the whole family of SLI syndromes. Note that, in doing this, they are positing a very strong, direct determination between having an unimpaired capacity for agreeing units, on the one hand, and being capable of developing a full competence with LDDs at large. My own view is that things are not so cleanly, directly, and linearly connected, which is by no means unusual in the case of interacting cognitive (or, for that matter, biological) agents. According to my alternative reading, the ultimate locus of the impairment is to be found in the

brain system in charge of keeping units active for symbol-manipulation operations—LDDs among them, which in all likelihood is a general-purpose one—i.e. not language-specific, as stressed by several recent sources (Marton et al. 2011; Lukács and Kemény 2014; Hsu and Bishop 2014). Therefore, in line with this view, the developmental impact of agreement morphology would be, primarily, on a brain circuitry in charge of capturing LDDs generally, and, derivatively, language-particular LDDs (see Section “From Cue-Response to Hybrid Systems”, for some anatomical details).

Within this developmental interpretation, both endogenous and exogenous factors impact on the growth of the aforementioned circuitry. The system thus may become impaired—in comparison to typical outputs—due to internal causes, external causes, or both. According to this reading, difficulties with agreement morphology is one such exogenous cause—but see below for some qualifications, in line with the well-known observation that most speakers with SLI tend to omit or to use at random such items (Rice and Oetting 1993; Marchman et al. 2004). Note that morphological relations are commonly based on adjacent—i.e. not-long—dependencies (NLDDs), the prototype of which is the root-affix relation in words like *sings* ($\text{sing}_{\text{root}}\text{-s}_{\text{affix}}$). Consequently, problems with productive morphology—predominantly of a NLDD kind—and problems with LDDs are, on mere formal grounds, of a different character. If my hypothesis is correct, their connection is contingent and of a developmental character, which is unveiled when limited perception of agreement-bearing affixes has a type of cascading effect that puts children at risk to an abnormal development of the system in charge of dealing with LDDs.

AQ2

The parallelisms between KP and LDP should be clear at this point. Traverse lines, tiles, and so on, have a triggering effect on the execution of motor programs in the case of neurological patients—and, arguably, also in development (Thelen and Smith 1994: Ch.4); in like manner, during development, agreement connections have a triggering effect on the establishment of complex LDDs—conspicuously, in specifically impaired speakers. Neurological patients freeze, stumble or fall, when deprived of such environmental cues; in like manner, children with SLI have lifelong difficulties with LDDs, which may be lessened by early training sessions specifically focused on agreement. It would seem reasonable to suggest that non-interpretable agreement relations relate to the uppermost layer of linguistic complexity—namely, the one that makes speakers competent in syntactic connections among distantly located constituents, similarly to what certain distinctive environmental cues do in relation to other areas of non-linguistic cognitive specialization.

Before closing this section, let me briefly introduce an extra, more speculative parallelism. Baram (2013, 2017) observes that the motor performance of people with PD improves when virtual cues (e.g. tiles) projected on the floor are not continuously shot, but adjusted via feed-back to the the patients' pace. According to Baram's explanation, the reason is that cues lose their effect when they become habitual and predictable, a fact that he relates with the dopamine rewarding system of basal ganglia, which privileges saliency and novelty (Baram 2017: 5). Note, as regards the case of language, that the point at which a particular agreement connection is solved is not predictable, since the distance between relata is unconstrained. Consequently, it seems reasonable that agreement relations should qualify as good candidates for feeding the same rewarding system as the inputs subserved by the devices envisioned by Baram's team. More to the point, basal structures like the striatum and the thalamus, which are known to be involved in the rewarding system of concern (Ballard 2015), are also well-known for their participation in the language-related brain circuitry (Booth et al. 2007), which I shall introduce with some detail in the next section.

From Cue-Response to Hybrid Systems

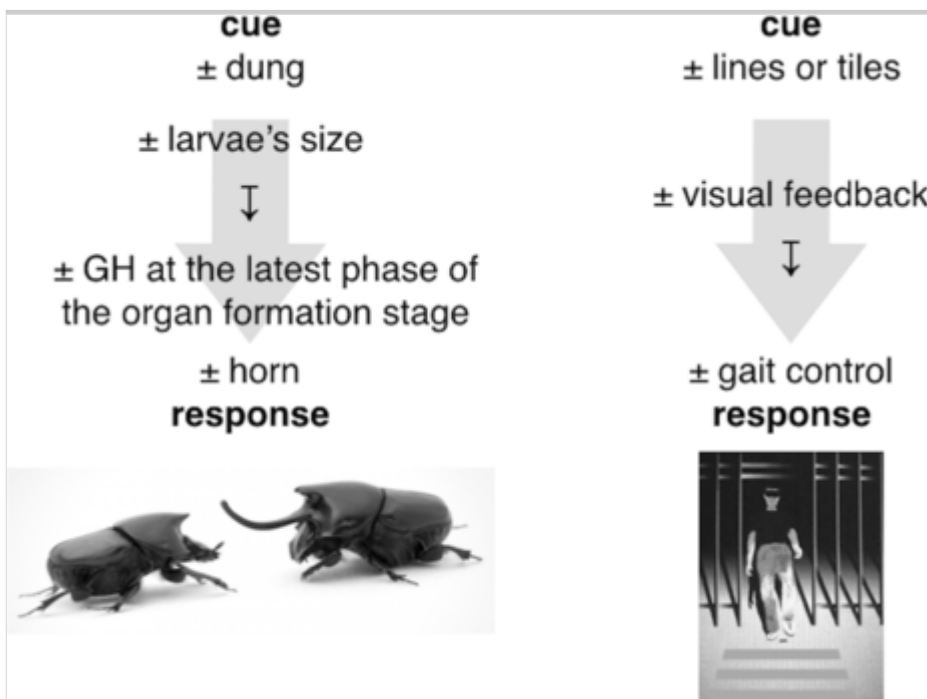
Obviously, the parallelism between PD and SLI is not a complete one. In particular, it is worth noticing that while cues apt to lessen the motoric problems associated to PD are of an environmental provenance—sometimes artificially created by technicians, the kinds of cues apt to enhance language development and lessen the linguistic difficulties associated to SLI are provided by the language system itself. The difference is non-trivial and has far-reaching consequences in the delimitation of language as an evolved organic system.

The connection between visual (and auditory) cues and the motoric and rewarding brain systems is not difficult to accommodate to the current understanding of many cognitive and other biological phenomena. Two conceptual paradigms come immediately to mind. Firstly, the phenomenon fits well with the 'external mind' one, as put forward by Clark (1997) and subsequent works. According to this, visual stimuli, like lines drawn on the floor or tiles, are part and parcel of the cognitive system in charge of walking, as witnessed when their participation in the programing and execution of gait become critical in pathological contexts. These kinds of cases, according to the paradigm, show that mind and brain are different entities, since the latter is an organism-internal one, while the former also spreads to the environment, recruiting all kinds of organism-external entities capable of playing relevant roles in cognitive tasks. In other words, the mind is partly internal and partly external.

Secondly, the phenomenon may be more generally accommodated to the ‘cue-response’ concept, coming from ecological developmental biology (eco-devo) (Gilbert and Epel 2015; Sultan 2015). The idea refers to the very specific interactions that commonly hold between designated environmental inputs and certain aspects of the outcomes of organic development, in the sense that the latter may attain one or another steady state depending on the former, even in the case of individuals with an otherwise common genetic background. A suitable illustration of the phenomenon is the case of the two different phenotypes of male dung beetles, horned and hornless, which are attained depending on the amount of dung they ingest when larvae (Moczek 2005). Cases like these illustrate that very specific aspects of the environment may be component parts of systems of organic development, to the same extent as other organism-internal resources (see Oyama et al. 2001, for a comprehensive treatment of this issue) (Fig. 1).

Fig. 1

The attainment of the beetle’s horned and hornless phenotypes (left) and the improvement of gait **gait** in KP (right), conceptualized as cases of environmental cue/organic response mechanisms



The connection between agreement morphology and LDDs in the case of language adds something important to the cases above, since, unlike these, agreement morphology is not an environmental ‘given,’ as can be said of lines drawn on a floor, irrespectively of their being conceived as a means of assessing walking. In other words, agreement morphological systems are ‘creations’ of the

same language system that they (under hypothesis) help to unfold, in a sense in which e.g. tiles are not in relation to the motoric system. The bottom line is that the former require a partially different conceptualization.

A suitable model for doing this may be based, as a first step, on the concept of ‘developmental hybrid,’ as recently put forward by Griesemer (2014a, b; see also Chiu and Gilbert 2015; Giorgi and Bruni 2015). Griesemer refers as cases of developmental hybridization to those situations in which some biotic or abiotic materials become a substantive part of a given organic system, providing a facilitating effect on “maintenance, growth, development, or construction” tasks that “would otherwise be more difficult or costly without” (Griesemer 2014a: 26). Said materials are originally exogenous, but, when incorporated into the system, they temporarily behave as ‘scaffolds,’ a Vygotskian inspired category (Vygotsky 1986) for referring to environmental entities with which developing agents interact commonly, making possible or more easy the achievement of certain outcomes or effects. Griesemer’s favorite examples are taken from the molecular level, like the HIV-1 virus’ reproductive cycle, which takes advantage of the molecular machinery within the host cell approximately in the following way. After crossing the cell’s membrane, the virus spreads its contents through the cell’s cytoplasm, where they encounter all the required means, firstly, to reconstruct the virus’ DNA template; then, to transplant it into the cell’s nucleus; and, finally, to translate it into a new virus, ready to abandon the host’s membrane. The virus is thus a hybrid, since its replication critically relies on biotic materials (scaffolds, from the virus’ point of view) that it finds in the outer environment (the host cell), which ultimately become part and parcel of its own constitutive elements. Griesemer’s central claim is that this model is an apt one for understanding many cases of interaction between biotic and abiotic materials as examples of ‘developmental hybridization,’ instead of declaring them problematic from the point of view of cultural or biological reproduction alone. This way, cultural items, without being self-reproducing materials, may be part and parcel of a ‘collective’ kind of reproductive dynamics, namely, that of the whole ‘hybrid’ entity capable of (re)producing the former and, in doing so, also benefitting its own developmental and reproductive dynamics. The same pattern can be applied to humans’ interaction with all kind of artifacts in evolutionary time, from prehistoric lithic industries to current mobile technologies (Griesemer 2017).

A version of the same ‘scaffold’ concept has also been incorporated into the external mind paradigm by Clark (1997). However, differently to this latter version, Griesemer’s hybrids are scaffolds that behave as such when (or because) they become internal or endogenous to the system. If applied to cognition, developmental hybrids do not extend the mind outside the brain—as

Vygotskian/Clarkian scaffolds purportedly do, but are rather embodied into the brain functioning, ultimately preserving the brain's centrality in cognition (Rupert 2009). As for the ultimate fate of developmental scaffolds, Griesemer suggests that they may be “removed” or become “invisible” (Griesemer 2014a: 26)—or, I dare to suggest, perhaps ‘recycled’ to other purposes.

However, Griesemer's concept does not provide a completely accurate image of the case under investigation. For one thing, agreement markers are not materials exogenous to the language system, but rather outcomes of its own activity and endogenous in a sense not completely different to e.g. hormones in relation to the endocrine system. Thus, in a certain way, the phenomenon mirrors, rather than mimics, Griesemer's hybrids. I shall refer to it as ‘retro-hybridization’—or, for the sake of simplicity, ‘r-hybridization,’ which I nevertheless conceive of as a particularly special instance of the general category of ‘developmental hybridization.’ My suggestion is that agreement systems illustrate a particular case of r-hybridization in relation to the language system, in turn conceived of as a ‘developmental hybrid’ as a whole. Let me explain.

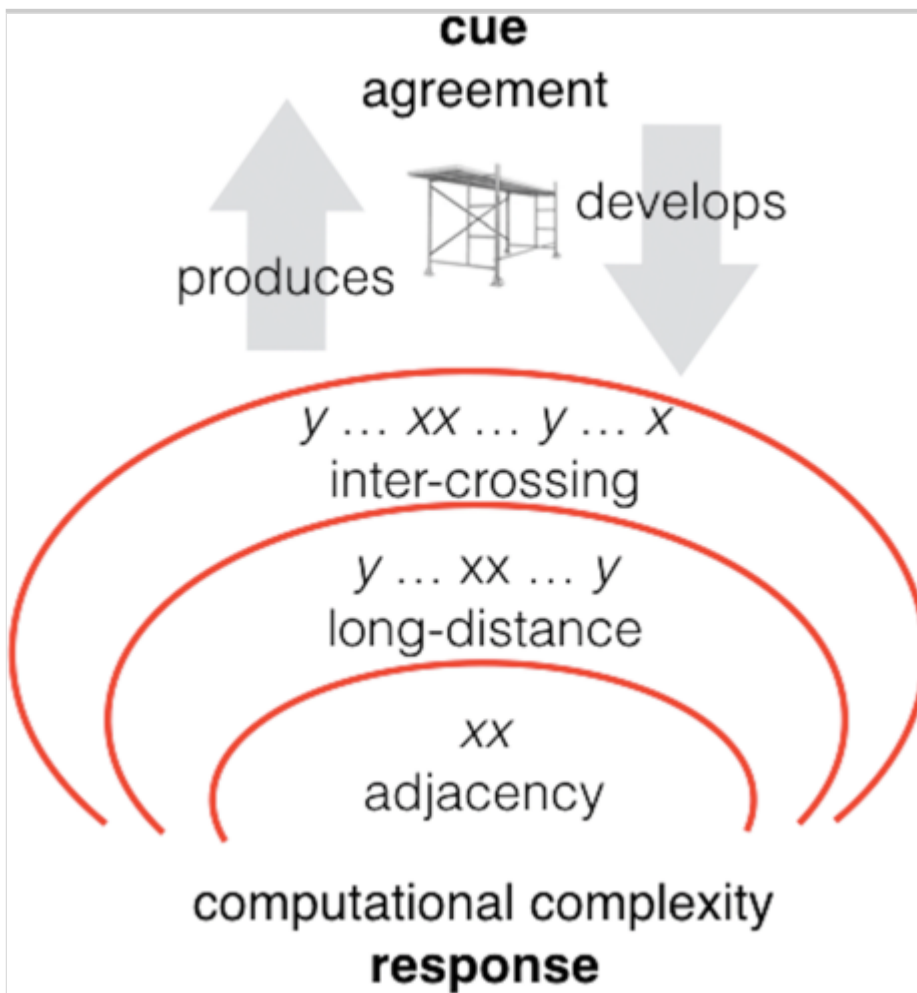
Agreement markers clearly obey to Griesemer's general characterization of ‘scaffolds’ in that, as argued in the previous section, they facilitate “maintenance, growth, development, or construction” of a system of computation ultimately capable of dealing with a (moderately complex) Type 1 language—i.e. one which contains sequences with symbols connected at indeterminate lengths and intercrossing connections. This is a ‘developmental role’ not unknown in other organic realms. For example, Minelli (2003) credits such a role to cuticles and flagella, which first constrain the form of the developing organisms and only later on serve as protection or motility, respectively. Similarly, agreement markers first provide means for exciting and exercising the brain circuits in charge of processing LDDs, and then may putatively play other communicative or informative roles, like identifying discourse referents or providing redundancy—as suggested by most functionalists; or, alternatively, no other role at all, being simply identified and ignored by the system—as suggested within the Chomskyan framework. In all this, agreement markers are scaffolds.

Besides, also like Griesemer's scaffolds—but unlike Vygotsky's or Clark's external counterparts, agreement markers are somehow exogenous and endogenous to the system at the same time. For different reasons, however: Namely, while Griesemer's scaffolds are originally exogenous and then turn into an endogenous condition in order to exert their developmental role, agreement markers are originally endogenous (i.e. self-subserved by the language system itself), and then turn into an exogenous condition (being prominently among the most distinguishing traits of particular grammatical traditions; Corbett 2009) that

enables them to exert their developmental role. All in all, they are, according to my terminological suggestion, ‘r-scaffolds.’ The net effect of all this is that language is ultimately neither a totally external or a totally internal entity; nor one wherein the external and the internal merely pile up. Language is a hybrid, a particular kind of hybrid—an r-hybrid (Fig. 2).

Fig. 2

The developmental dynamics between agreement and the system of computation underlying language acquisition and production, conceptualized as a cue/response system. The resulting scaffolding effects pave the way to a further conceptualization of the overall system as a developmental hybrid



Such a ‘hybrid’ concept of language may serve to gain a more integrative naturalistic view of this human capacity without forcing the limits of current biological conceptions regarding the boundaries of the organism and the environment. An organic system needs, of necessity, to have an inside; however, current biological thinking also teaches us that such a truism does not exclude the possibility of a given organic system also existing beyond its holder’s boundaries (Pradeu 2011; Gilbert et al. 2012). Boundaries are to be thought of in

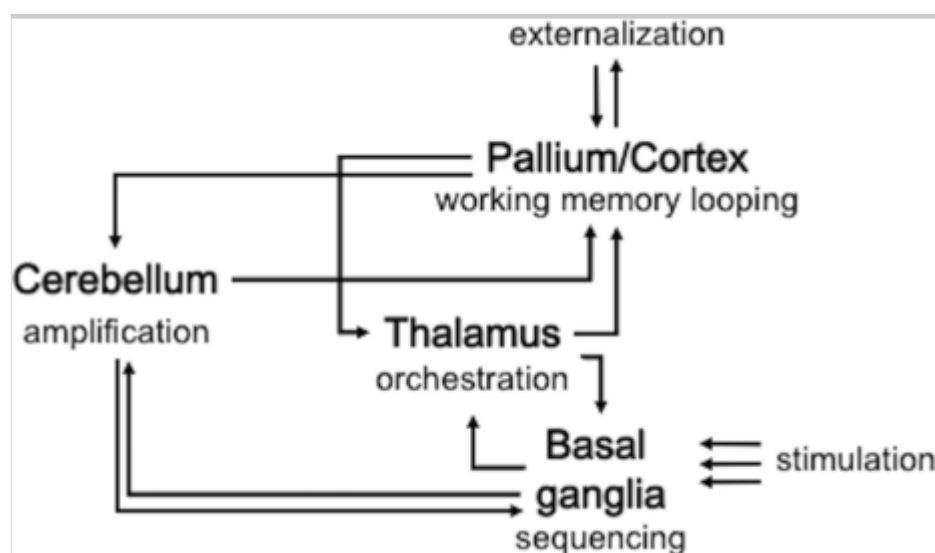
biology more as protective or interfacing structures than as limiting ones. ‘Limit’ is maybe not even a bona fide biological category.

The ‘hybrid’ concept may also serve to offer a new solution to the debate on the brain specialization of language—or lack thereof. In recent years, a convergent amount of neuro-anatomical and genetic evidence has led to a consensual view on the cerebral basis of language, which locates it in a widely extended cortical-basal-cortical network (Aboitiz et al. 2006; Balari and Lorenzo 2013; Friederici 2017; Lieberman 2006; Ullman 2006). Within this anatomical system, parts of the basal ganglia, exercising a repetitive sequencer role, and left frontal-temporal-frontal circuits, in all likelihood exercising the role of a reverberating memory loop, appear to be crucially involved in the language function.

Moreover, the thalamus, with a monitoring or controlling role (Booth et al. 2007; Hebb and Ojemann 2013), and the cerebellum, with that of amplifying and protecting the ongoing loop from other incoming similar stimuli (Booth et al. 2007; Bostan et al. 2013; Middletona and Strick 2000), have been more recently added to the whole picture (Fig. 3).

Fig. 3

A schematic overview of the brain system in charge of sequential processing tasks. Under hypothesis, this is the system on which agreement units have a developmental impact, ultimately leading to their mutual hybridization and specialization for the linguistic function



AQ3

These brain circuits certainly display the capacities and potentials that make them apt for dealing with all kinds of sequentially organized stimuli/behaviors, including linguistic sequences. However, they appear to be crucially involved in tasks other than just language (motor planning and execution, abstract reasoning,

etc.; see Lieberman 2006, for a synthesis). Based on this observation, the most reasonable conclusion is that the relevant circuits lack any fine-grained specialization at the onset, yet they become language-devoted ones in the early experience of children, through a process that may be equated to other episodes of developmental plasticity (West-Eberhard 2003). Moreover, in as much as these circuits are characteristically sensitive to particular kinds of indisputably language-specific environmental inputs (e.g. agreement items and connections), which have been argued here to have a developmental impact on the circuits of concern, the idea is strengthened that the human brain ends up incorporating an organic system which is not strictly language-specific, but in any event language-specialized. Such a fate is not blueprinted on a language genotype (Chomsky 1980), but it is nonetheless robustly and reliably attainable as a function of the epigenetic conspiracy of cross-generationally relentless resources of both internal and external provenance (Bateson and Gluckman 2011; Gottlieb 2007).

Admittedly, the model that I have expounded in the previous pages may inspire a chicken or the egg kind of dilemma, as agreement connections are said to partake in the development of the very system that produces and processes them. However, I believe that such an impression is deceptive. Note that the model does not ascribe to this particular kind of LDD the role of constructing said system from scratch. Actually, it has been experimentally shown that children are endowed, almost from birth, with the capacity of identifying sequential stimuli (e.g. meaningless CV syllables) if they contain repeated units, if repetitions are located at one or another edge of segments, and so on (Gervain et al. 2008, 2012; Marcus et al. 1999). Similar stimuli, but in naturalistic contexts, may in all likelihood facilitate the strengthening of such early computational endowment, paving the way for the capacity of generating and processing connections more complex than mere repetitions, and instigating a sort of developmental dialectic of sorts leading to increasingly high levels of structural complexity of inter-sequential connections. Similarly, but in evolutionary time, this kind of dialectic might also have paved the way for the brain basis for the capacity of creating and dealing with intricate traditional agreement systems, aimed at benefiting the developmental dialectic itself.

Concluding Remark

Linguistics is nowadays strongly divided regarding the issue the ontology of its subject matter: What kind of object is it? Where should one locate it? Ultimately, what is linguistics about? On the one side is the position that conceives language as a diversified cultural object, subject to traditional transmission, ultimately belonging to historical speaking communities. This position, which can be traced back to Ferdinand de Saussure's *Course de Linguistique Générale* (Saussure

1916), is currently defended by the likes of Terrence Deacon, Michael Tomasello, Simon Kirby, Morten Christiansen, and Nick Chatter, among others (Deacon 1997; Tomasello 2001; Kirby et al. 2007; Labov 2007; Christiansen and Chater 2008, 2016; Koster 2009). The most recurrent theme of this externalist position is that the human brain does not exhibit any evolved language-exclusive specialization, so languages have an entirely brain-external kind of existence, yet exhibiting definitional traits that reflect general constraining conditions related with how brains learn and process information, as well as how they are intergenerationally transmitted. This is also the position embraced by prominent defenders of the ‘external mind’ framework (Clark and Chalmers 1998), who, as a matter of fact, point to language as the epitome of the kinds of objects external to the brain, by means of which the mind becomes capable of spreading out the body—Rupert (2009) (critically) refers to this argument as the ‘Language-based Inference,’ which is developed by Clark (1998, 2004, 2005, 2006). A by no means minor risk of all these kinds of approaches is that they locate language very close to (if not within) the unclear ontological territory of the Fregean/Popperian ‘objective’ thought (Frege 1918; Popper 1972; Carr 1990), this in turn in close vicinity to the directly problematic realm of Platonic objects, where the likes of Jerrold Katz and Paul Postal have actually argued that language belongs (Katz 1981, 1984; Soames 1984; Katz and Postal 1991; for a more recent, sympathetic approach, see Devitt 2006).

AQ4

On the other side, an alternative position exists that claims that language is but one of the domain-specialized components of the human brain—namely, one in charge of articulating (a special kind of) internal thought, thus a strictly organism-internal object. It is the position commonly held by Chomskyans, explicitly and extensively propounded by Noam Chomsky himself in Chomsky (1986). Chomsky’s main anti-externalist argument is that devoid of any connection with some mind-internal mechanism, there is no way of deciding the correct (or true) description (grammar) of a given (external) language among different extensionally equivalent such descriptions—i.e. different grammars equally capable of giving a formal account to every well-formed expression of the (external) language. According to Chomsky’s conclusion, language conceived of as an external object cannot, therefore, be the subject matter of bona fide scientific inquiry—see Lohndal and Narita (2009), for a methodological, rather than ontological, reading of this conclusion. More recently (e.g. Chomsky 2013, 2016), Chomsky’s internalism has become even more radical. Now Chomsky excludes the whole ‘externalization’ component (including Linearization, Morphology, Phonology, and Phonetics) from the Faculty of Language proper, somehow as predicted by Burton-Roberts (2011), and equates it with a ‘language of thought’ (LOT), “quite probably the only such

LOT,” in his own words (Chomsky 2016: 13). In any event, even if Phonology (in a broad sense) is brought into the picture, Chomsky has famously argued that its units (broadly speaking, phonemes) lack any external correlates (Chomsky 2012). Despite Chomsky’s purportedly biological conceptualization of the language faculty, such radically deep, environmental-refractory kinds of internal organic systems are not (to say the least) a common currency in nature (Sultan 2015).

In recent years, some proposals have been made which try to make sense of the fact that language appears to have a dual kind of existence, in part cultural and in part biological (Balari and Lorenzo 2013; Bickerton 2014). However, proposals like these are not completely satisfactory, since they are based on a kind of ‘divide and conquer’ principle of sorts, distinguishing two layers of linguistic reality and referring distinct aspects of language to each—e.g. lexical inventories, or language-particular rules belong to the cultural layer, the structure-building generative engine belong to the biological one, and so on and so forth. However, this kind of ‘conventional dualism’ has been previously criticized (Oyama 2000), for the fact that it relies on a dualist stance, not very different from the familiar mind/body divide, without a clear proposal about which kinds of interactions make the resulting continuum ontologically coherent. As for certain self-publicized integrative projects—like the one of Mondal (2012), while certainly worth serious consideration, it is not however clear whether they really fit the bill. In the case of Mondal’s otherwise insightful approach, for example, a lot of theoretical weight is put on an ‘(elastic) undulation’ concept, used to describe the interactive nature of the external and internal sides of language. According to this concept, there exists an interactive trade-off of knowledge from the community to the individual (outside-in) and from the individual to the community (inside-out); but while it is easy to conceive of individuals as cognitive agents, it is unclear how communities can be (literarily) endowed with the same kind of agency. Mondal tries to connect this to the ‘external mind’ paradigm (Clark 1997), which in turn tries to show that the mind spreads into the external world, so the former cannot be equated with the (internal) brain. However, the same paradigm does not show (nor does it aim to show) that the external world is the site of a kind of cognitive agency complementary to that of brains. Moreover, the concept of ‘undulation,’ which Mondal acknowledges to be at a complex ontological, epistemological and phenomenological crossroads, has no recognizable homologue in the life sciences.

In this paper, I have referred to certain ‘new wave’ concepts of the emerging eco-devo paradigm as offering the means for a particularly promising reframing of the issue; above all, the ‘developmental hybrid’ concept, conceived as a strong

variant of the overarching ‘cue-response system’ category, which appears to be particularly apt in order to capture the kinds of developmental dynamics that represent the ins and outs of the matter. According to this frame, every single component part of language is at the same time internal and external. Such a kind of promiscuity between the external and the internal appears to be a broadly extended developmental strategy, to which wide-ranging effects of a biosemiotic character are in all likelihood associated, from multi-cellularity assemblage or symbiosis associated phenomena (Chiu and Gilbert 2015; Griesemer 2014a, b, 2017) to the construction of animal traditions of varying degrees of richness (Jablonka and Avital 2000).

Acknowledgements

This paper has benefitted from a grant of the Spanish Government (Ministry of Economy, Industry and Competitiveness) (Ref. FFI2017-87699-P). I would like to express my gratitude to the editors and two anonymous reviewers for their valuable help and comments. All remaining errors are of my own responsibility.

Compliance with Ethical Standards

Conflict of Interest The author declares that he has no conflict of interest.

References

Aboitiz, F., García, R., Brinetti, E., & Bosman, C. (2006). The origins of Broca’s Area and its connection from an ancestral working-memory network. In Y. Grodzinsky & K. Amunts (Eds.), *Broca’s region* (pp. 3–16). Oxford: Oxford University Press.

Adger, D., & Svenonius, P. (2011). Features in minimalist syntax. In C. Boeckx (Ed.), *The Oxford handbook of linguistic minimalism* (pp. 27–51). Oxford: Oxford University Press.

Balari, S., & Lorenzo, G. (2013). *Computational phenotypes. Towards an evolutionary developmental biolinguistics*. Oxford: Oxford University Press.

Balari, S., & Lorenzo, G. (2018). The internal, the external and the hybrid: The state of the art and a new characterization of language as a natural object. *Glossa. A Journal of General Linguistics*, 3(1), 22.
<https://doi.org/10.5334/gjgl.330> .

Ballard, D. H. (2015). *Brain computation and hierarchical abstraction*. Cambridge: The MIT Press.

Baram, Y. (1999). Walking on tiles. *Neural Processing Letters*, 10(2), 81–87.

Baram, Y. (2013). Virtual sensory feedback for gait improvement in neurological patients. *Frontiers in Neurology*, 4.
<https://doi.org/10.3389/fneur.2013.00138> .

Baram, Y. (2017). Transformative autonomous entrainment of gait in neurological patients. *Journal of Neurology and Neuroscience*, 8.
<https://doi.org/10.21767/2171-6625.1000177> .

Baram, Y., & Miller, A. (2007). Auditory feedback control for improvement of gait patients with multiple sclerosis. *Journal of the Neurological Sciences*, 254(1–2), 90–94.

Barlow, M., & Ferguson, C. (Eds.). (1988). *Agreement in natural language: Approaches, theories, descriptions*. Stanford: CSLI Publications.

Bateson, P., & Gluckman, P. (2011). *Plasticity, robustness, and evolution*. New York: Cambridge University Press.

Beckers, G. J. L., Bolhuis, J. J., Okanoya, K., & Berwick, R. C. (2012). Birdsong neurolinguistics: Songbird context-free grammar claim is premature. *NeuroReport*, 23(3), 139–145.

Bickerton, D. (2014). *More than nature needs: Language, mind, and evolution*. Cambridge: Harvard University Press.

Bishop, D. V. (1997). *Uncommon understanding. Development and disorders of language comprehension in children*. Hove: Psychology Press.

Boeckx, C. (Ed.). (2006). *Agreement systems*. Amsterdam: John Benjamins.

Booth, J. R., Wood, L., Lu, D., Houk, J. C., & Bitan, T. (2007). The role of the basal ganglia and cerebellum in language processing. *Brain Research*, 1113(1), 136–144.

Bostan, A. C., Dum, R. P., & Strick, P. L. (2013). Cerebellar networks with the cerebral cortex and basal ganglia. *Trends in Cognitive Science*, 17(5), 241–254.

Burton-Roberts, N. (2011). On the grounding of syntax and the role of phonology in human cognition. *Lingua*, 121(14), 2089–2102.

Carr, P. (1990). *Linguistic realities. An autonomist metatheory for the generative enterprise*. Cambridge: Cambridge University Press.

Chiu, L., & Gilbert, S. F. (2015). The birth of the holobiont: Multi-species birthing through mutual scaffolding and niche construction. *Biosemiotics*, 8(2), 191–210.

Chomsky, N. (1956). Three models for the description of language. *IRE Transactions on Information Theory*, 2(3), 113–124.

Chomsky, N. (1980). *Rules and representations*. New York: Columbia University Press.

Chomsky, N. (1986). *Knowledge of language: Its nature, origin, and use*. New York: Praeger.

Chomsky, N. (1995). *The minimalist program*. Cambridge: The MIT Press.

Chomsky, N. (2012). *The science of language. Interviews with James McGilvray*. Cambridge: Cambridge University Press.

Chomsky, N. (2013). Problems of projection. *Lingua*, 130(June), 33–49.

Chomsky, N. (2016). *What kind of creatures are we?* New York: Columbia University Press.

Christiansen, M. H., & Chater, N. (2008). Language as shaped by the brain. *Behavioral and Brain Sciences*, 31(5), 489–509.

Christiansen, M. H., & Chater, N. (2016). *Creating language: Integrating development, acquisition, and processing*. Cambridge: MIT Press.

Clahsen, H. (1986). Verb inflections in German child language: acquisition of agreement markings and the functions they encode. *Linguistics*, 24(1), 79–121.

Clahsen, H., & Hansen, D. (1997). The grammatical agreement deficit in specific language impairment: Evidence from therapy experiments. In M. Gopnik (Ed.), *The inheritance and innateness of grammars* (pp. 141–160). Oxford: Oxford University Press.

Clahsen, H., Bartke, S., & Göllner, S. (1997). Formal features in impaired grammars: a comparison of English and German SLI children. *Journal of Neurolinguistics*, 10(2–3), 151–171.

Clark, A. (1997). *Being there: Putting body, brain, and world together again*. Cambridge: The MIT Press.

Clark, A. (1998). Magic words: how language augments human computation. In P. Carruthers & J. Boucher (Eds.), *Language and thought: Interdisciplinary themes* (pp. 162–183). Cambridge: Cambridge University Press.

Clark, A. (2004). Is language special? Some remarks on control, coding, and co-ordination. *Language Sciences*, 26(6), 717–726.

Clark, A. (2005). Word, niche, and super-niche. How language makes minds matter more. *Theoria*, 54, 255–268.

Clark, A. (2006). Language, embodiment, and the cognitive niche. *Trends in Cognitive Sciences*, 10(8), 370–374.

Clark, A., & Chalmers, D. J. (1998). The extended mind. *Analysis*, 58(1), 7–19.

Corballis, M. C. (2011). *The recursive mind. The origins of human language, thought, and civilization*. Princeton: Princeton University Press.

Corbett, G. G. (2009). *Agreement*. Cambridge: Cambridge University Press.

de Saussure, F. (1916). *Cours de linguistique générale*. Paris: Payot.

Deacon, T. W. (1997). *The symbolic species. The co-evolution of language and the brain*. New York: W.W. Norton.

Devitt, M. (2006). *Ignorance of language*. Oxford: Oxford University Press.

Fitch, T., & Hauser, M. D. (2004). Computational constraints on syntactic processing in a nonhuman primate. *Science*, 303(5656), 377–380.

Frege, G. (1918/1977). *Thoughts: A logical inquiry* (translated from German by P. Geach & R. Stoothoff). In P. Geach (Ed.), *Logical Investigations* (pp. 1–30). Oxford: Blackwell.

Friederici, A. D. (2017). *Language in our brain. The origins of a unique human capacity*. Cambridge: The MIT Press.

Gervain, J., Macagno, F., Cogoi, S., Pena, M., & Mehler, J. (2008). The neonate brain detects speech structure. *Proceedings of the National Academy of Sciences USA*, 105(37), 14222–14227.

Gervain, J., Berent, I., & Werker, J. F. (2012). Binding at birth: The newborn brain detects identity relations and sequential position in speech. *Journal of Cognitive Neuroscience*, 24(3), 564–574.

Gilbert, S. F., & Epel, D. (2015). *Ecological developmental biology: The environmental regulation of development, health, and evolution*. Sunderland: Sinauer.

Gilbert, S. F., Sapp, J., & Tauber, A. I. (2012). A symbiotic view of life: we have never been individuals. *The Quarterly Review of Biology*, 87(4), 325–341.

Giorgi, F., & Bruni, L. E. (2015). Developmental scaffolding. *Biosemiotics*, 8(2), 173–189.

Gottlieb, G. (2007). Probabilistic epigenesis. *Developmental Science*, 10(1), 1–11.

Griesemer, J. R. (2014a). Reproduction and the scaffolded development of hybrids. In L. Caporael, J. R. Griesemer, & W. C. Wimsatt (Eds.), *Developing scaffolds in evolution, culture, and cognition* (pp. 23–55). Cambridge: The MIT Press.

Griesemer, J. R. (2014b). Reproduction and scaffolded development processes: An integrated evolutionary perspective. In A. Minelli & T. Pradeu (Eds.), *Towards a theory of development* (pp. 183–202). Oxford: Oxford University Press.

Griesemer, J. R. (2017). Landscapes of developmental collectivity. In S. B. Gissis, E. Lamm, & A. Shavit (Eds.), *Landscapes of collectivity in the life sciences* (pp. 25–48). Cambridge: The MIT Press.

Hauser, M. D., Chomsky, N., & Fitch, W. T. (2002). The faculty of language: What is it? Who has it? How did it evolve? *Science*, 298(5598), 1569–1579.

Hebb, A., & Ojemann, G. A. (2013). The thalamus and language revisited. *Brain and Language*, 126(1), 99–108.

Hsu, H. J., & Bishop, D. V. M. (2014). Sequence-specific procedural learning deficits in children with specific language impairment. *Developmental Science*, 17(3), 352–365.

Jablonka, E., & Avital, E. (2000). *Animal traditions. Behavioral inheritance in evolution*. Cambridge: Cambridge University Press.

Katz, J. J. (1981). *Language and other abstract objects*. Totowa: Rowman and Littlefield.

Katz, J. J. (1984). An outline of platonist grammar. In T. Bever, J. M. Carroll, & L. A. Miller (Eds.), *Taking minds: The study of language in cognitive sciences* (pp. 1–33). Cambridge: The MIT Press.

Katz, J. J., & Postal, P. M. (1991). Realism vs. conceptualism in linguistics. *Linguistics and Philosophy*, 14(5), 515–554.

Kirby, S., Dowman, M., & Griffiths, T. L. (2007). Innateness and culture in the evolution of language. *Proceedings of The National Academy of Science USA*, 104(12), 5241–5245.

Koster, J. (2009). Ceaseless, unpredictable creativity: language as technology. *Biolinguistics*, 3, 321–331.

Labov, W. (2007). Transmission and diffusion. *Language*, 83(2), 344–387.

Leonard, L. B. (1998). *Children with specific language impairment*. Cambridge: The MIT Press.

Leonard, L. B., McGregor, K. K., & Allen, G. D. (1992). Grammatical morphology and speech production in children with specific language impairment. *Journal of Speech, Language and Hearing Research*, 35, 1076–1085.

Lieberman, P. (2006). *Toward an evolutionary biology of language*. Cambridge: Harvard University Press.

Lohndal, T., & Narita, H. (2009). Internalism as a methodology. *Biolinguistics*, 3, 321–331.

Lukács, Á., & Kemény, F. (2014). Domain-general sequence learning deficit in specific language impairment. *Neuropsychology*, 28(3), 472–483.

Marchman, V. A., Saccuman, C., & Wulfeck, B. (2004). Productive use of the English past tense in children with focal brain injury and specific language impairment. *Brain and Language*, 88(2), 202–214.

Marcus, G. F., Vijayan, S., Bandi Rao, S., & Vishton, P. M. (1999). Rule learning by seven-month-old infants. *Science*, 283(5398), 77–80.

Marton, K., Campelli, L., & Farkas, L. (2011). Grammatical sensitivity and working memory in children with language impairment. *Acta Linguistica Hungarica*, 58(4), 448–466.

Middleton, F. A., & Strick, P. L. (2000). Basal ganglia and cerebellar loops: motor and Cognitive Circuits. *Brain Research Reviews*, 31(2–3), 236–250.

Mills, A. E. (1985). The acquisition of German. In I. Slobin (Ed.), *The crosslinguistic study of language acquisition* (Vol. 1: The data, pp. 141–254). Hillsdale: Lawrence Erlbaum.

Minelli, A. (2003). *The development of animal form*. Cambridge: Cambridge University Press.

Moczek, A. P. (2005). The evolution of development of novel traits, or how beetles got their horns. *Bio-Science*, 55(11), 937–951.

Mondal, P. (2012). Can internalism and externalism be reconciled in a biological epistemology of language? *Biosemiotics*, 5(1), 61–82.

Nevins, A. (2011). Multiple agree with clitics: Person complementarity vs. omnivorous number. *Natural Language and Linguistic Theory*, 29(4), 939–971.

Oyama, S. (2000). *Evolution's eye: A systems view of the biology-culture divide*. Durham, and London: Duke University Press.

Oyama, S., Girffiths, P. E., & Grey, R. D. (Eds.). (2001). *Cycles of contingency: Developmental systems and evolution*. Cambridge: The MIT Press.

Palacios-Navarro, G., Albiol-Pérez, S., & García-Mariño García, I. (2016). Effects of sensory cueing in virtual motor rehabilitation. *Journal of Biomedical Informatics*, 60, 49–57.

Popper, K. (1972). *Objective knowledge. An evolutionary approach*. Oxford: Clarendon Press.

Pradeu, T. (2011). A mixed self: the role of symbiosis in development. *Biological Theory*, 6(1), 80–88.

Preminger, O. (2014). *Agreement and its failures*. Cambridge: The MIT Press.

Prothero, J. D. (1993) *The treatment of akinesia using visual images*. University of Washington, M.A. Thesis.

Rice, M. L., & Oetting, J. B. (1993). Morphological deficits of children with SLI: Evaluation of number marking and agreement. *Journal of Speech, Language and Hearing Research*, 36(6), 1249–1257.

Rupert, R. D. (2009). *Cognitive systems and the extended mind*. Cambridge: Cambridge University Press.

Soames, S. (1984). Linguistics and psychology. *Linguistics and Philosophy*, 7(2), 155–179.

Sultan, S. E. (2015). *Organism and environment: Ecological development, niche construction, and adaptation*. Cambridge: Cambridge University Press.

Thelen, E., & Smith, L. B. (1994). *A dynamic systems approach to the development of cognition and action*. Cambridge: The MIT Press.

Tomasello, M. (2001). *The cultural origins of human cognition*. Cambridge: Harvard University Press.

Ullman, M. T. (2006). Is Broca's are part of a basal ganglia thalamocortical circuit? *Cortex*, 42(4), 480–485.

van Heijningen, C. A. A., de Visser, J., & ter Cate, C. (2009). Simple rules can explain discrimination of putative recursive syntactic structures by a songbird species. *Proceedings of the National Academy of Sciences USA*, 106(48), 20538–20543.

Van Valin, R. D. (2003). Minimalism and explanation. In J. Moore & M. Polinsky (Eds.), *The nature of explanation in linguistic theory* (pp. 281–297). Stanford: CSLI Publications.

Vygotsky, L. (1986). *Thought and language*. Cambridge, MA: The MIT Press.

West-Eberhard, M. J. (2003). *Developmental plasticity and evolution*. Oxford: Oxford University Press.

Zushi, M. (2013). *Long distance dependencies*. London: Routledge.

¹ In the examples, embedding is represented by square brackets; dependences are marked by means of subscripts; ‘t’ is for trace (or gap) in the case of displaced constituents. The following abbreviations are used in examples (4) to (6): abs = ‘absolute,’ af = ‘agent focus,’ asp = ‘aspect,’ foc = ‘focus,’ obj = object, pl = plural, sg = singular.