biolinguistics

ARTICLES

Biolinguistics 14.SI: 59–101, 2020
www.biolinguistics.eu

Language in Language Evolution Research: In Defense of a Pluralistic View

Sławomir Wacewicz[1,*], Przemysław Żywiczyński[1], Stefan Hartmann[2], Michael Pleyer[3], & Antonio Benítez-Burraco[4]

[1] Center for Language Evolution Studies, Nicolaus Copernicus University, Toruń, Poland
[2] Faculty of Arts and Humanities, University of Düsseldorf, Germany
[3] Faculty of Arts and Humanities, University of Koblenz-Landau, Germany
[4] Faculty of Philology, University of Seville, Spain

[*] Corresponding author: wacewicz@umk.pl

SW: https://orcid.org/0000-0003-1488-6220
PŻ: https://orcid.org/0000-0002-0267-3845
SH: https://orcid.org/0000-0002-1186-7182
MP: https://orcid.org/0000-0001-6685-391X
ABB: https://orcid.org/0000-0003-4574-5666

Many controversies in language evolution research derive from the fact that *language* is itself a natural language word, which makes the underlying concept fuzzy and cumbersome, and a common perception is that progress in language evolution research is hindered because researchers do not 'talk about the same thing'. In this article, we claim that agreement on a single, top-down definition of language is not a *sine qua non* for good and productive research in the field of language evolution. First, we use the example of the notion FLN ('faculty of language in the narrow sense') to demonstrate how the specific wording of an important top-down definition of (the faculty of) language can—surprisingly—be inconsequential to actual research practice. We then review four approaches to language evolution that we estimate to be particularly influential in the last decade. We show how their breadth precludes a single common conceptualization of language but instead leads to a family resemblance pattern, which underwrites fruitful communication between these approaches, leading to cross-fertilisation and synergies.

Keywords: language; evolution of language; language faculty; language readiness

Editors: Lluís Barceló-Coblijn, Universitat de les Illes Balears, Spain
Evelina Leivada, Universitat Rovira i Virgili, Spain

Received: 10 May 2020
Accepted: 18 July 2020
Published: 23 March 2021

ISSN 1450–3417

1. Introduction

The emergence of linguistic behaviour undoubtedly counts among the few most defining developments in the history of our species. Darwin (1871) considered language to be the greatest invention of humankind, only equalled by fire, and Maynard Smith & Szathmáry (1995) include language on their list of only eight "major evolutionary transitions", alongside e.g. chromosomes or sexual reproduction. Christiansen & Kirby (2003) call the evolutionary emergence of language "the hardest problem in science", a label that is as bold as it is useful: Its last part underscores the progression of academic interest in language origins from the spheres of mythology, religion and philosophising to the domain of scientific investigation. Language evolution understood as a field of study (or "the Science of Language Evolution"; Żywiczyński 2018) is an interdisciplinary research field concerned with addressing this problem (see, e.g., the collection of papers in Tallerman & Gibson 2012 for an overview).

However, *language* is itself a natural language word that is fuzzy and polysemous, and as such eludes precise definitions. To a great extent, the same can be said of the diverse conceptualizations and technical uses of *language* promoted by the different branches of linguistics. Many have viewed the elusive nature of the term *language* as an inherent stumbling block to progress, lamenting the fact that researchers do not 'talk about the same thing' when discussing the evolution of language (e.g., Wescott 1991, Botha 2000, Hauser et al. 2002; see also Jackendoff 2010, Haspelmath 2016). In particular, many approaches use the term *language* to refer to a socially shared external code and see this as the explanandum of language evolution research, whereas other approaches are interested in language as a cognitive system, and still others as a biological entity, such as a genetically specified faculty or as a component part of the human brain (see Balari & Lorenzo 2016). Attempts to specify this explanatory target with more precise technical terminology have remained unsuccessful, and although the ontological complexity of language admittedly plays a role, to a large extent this is because different theoretical stances presuppose different conceptions of what constitutes language "proper".

2. Criticisms of the Conceptual Diversity of *Language*

Disagreements on the nature of language have always been present in the modern era of language evolution research, which to many starts in 1996 with the launch of the Evolang conference series, "the major meeting for researchers worldwide in the origins and evolution of language" (http://evolang.org). This first meeting resulted in a proceedings volume (Hurford et al. 1998) that was scrutinised by Rudolf Botha, himself one of the pioneers of the Evolang movement. Botha (2000) lists thirteen conceptualisations of *language* that can be found in this single volume: 'aspect of human behavior', 'process', 'gigantic meta-task', 'special human skill', 'activity', 'species-specific capacity', 'sort of contract signed by members of a community', 'hard-wired (individual) competence', group behaviour of social animals, application of Theory of Mind and social intelligence, 'mass phenomenon actualized by different agents interacting with each other', emergent property

that spontaneously forms itself, and 'complex system of labels for concepts and conceptual structures'. From this, Botha (2000) concludes that the

> […] profusion of ontologically distinct ways of characterizing language is symptomatic of a foundational flaw in work on language evolution: it indicates the absence of a shared, well-founded linguistic ontology. (2000: 152)

He goes on to suggest that "[t]he linguistic entity or entities whose evolution is at issue should be identified and characterized in a clear and non-arbitrary way", and that a lack of such a consensus is a fundamental stumbling block to progress:

> [I]n the absence of broad agreement about what language is as opposed to other linguistic entities, discussions [of language evolution] are bound to deal with questions of language evolution in ways that are inconclusive and internally disconnected. [1]
>
> (Botha 2000: 149, 152–153)

A similar critique—also catalysed by an Evolang meeting (2002 in Harvard)—was voiced in the consequential *Science* paper by Marc Hauser, Noam Chomsky, and Tecumseh Fitch. The founding premise of these authors was that "[t]he word 'language' has highly divergent meanings in different contexts and disciplines" (Hauser et al. 2002: 1570) and is simply too capacious to be productively used in scientific discourse without further specification.[3] Hauser et al. (2002) propose a remedy in the form of a more circumscribed and thus supposedly more scientifically productive notion of (the biological faculty of) language, which they term the faculty of language in the narrow sense, FLN, as separate from the faculty of language in the broad sense, FLB. We return to this distinction in section 3; here we underscore that the paper and the distinction essentially resulted from a deep dissatisfaction with the breadth and polysemy of the everyday word *language*, and the resulting diversity in its use in language evolution research. This is particularly evident in later commentaries, where the authors explain that their motivation behind proposing FLN was "to clarify misunderstandings and aid interdisciplinary rapprochement" (Fitch et al. 2005: 179), and

---

1 Similar and equally influential criticisms were also voiced before Evolang. For example, "[o]ne of the reasons for the extremely inconclusive outcome of scholarly debates on the origin and evolution of language is that so few glossogonists define language in the same way. Yet because their definitions, in most cases, remain implicit, definitional differences are rarely acknowledged. Before we can distinguish terminological disagreements from substantive disagreements, we must, I think, be as explicit as possible about what each of us means by the word 'language'" (Wescott 1991: 77).

3 Cf. Bolhuis et al. (2014: 1): "In our view, *for the purposes of scientific understanding*, language should be understood as a particular computational cognitive system, implemented neurally, that cannot be equated with an excessively expansive notion of 'language as communication'. […] In place of a complex rule system or accounts grounded on general notions of 'culture' or 'communication,' it appears that *human language syntax* can be defined in an extremely simple way that makes conventional evolutionary explanations much simpler."

> […] to clarify discussion and avoid confusion, once we realized that researchers (including ourselves) had been using the same word, 'language', to talk about two different things (FLB and FLN) for many years, and thus had been talking past each other.
>
> (Fitch 2010: 22)

A related but more recent concern was a much-discussed terminological critique by Martin Haspelmath, commenting on the foundation of the *Journal of Language Evolution* (which, next to Evolang, was another institutional milestone to language evolution research; see Dediu & de Boer 2016). While the target of Haspelmath's (2016) commentary is the term *evolution* and its apparent semantic extension from 'language origins' to 'language emergence and language change', it also targets *language* by extension, since the types of processes that are included under the rubric "evolutionary" predetermine the range of entities they apply to. Haspelmath criticises the resulting lack of clarity and terminological rigour:

> As usual, this semantic change of 'evolution' has happened because different communities are interested in different concepts, and people (including scientists) are reluctant to coin new terms for new concepts, preferring to adopt old terms from neighbouring communities. Moreover, even linguists tend to be unaware of semantic changes and thus sloppy about terminological use.[4] (Haspelmath 2016)

In a recent paper, Haspelmath (2020) has coined the cover term *human linguisticality*, understood as a biological capacity which is best studied in a broadly comparative perspective and which imposes some constraints on possible language systems, but with most similarities between languages resulting from convergent cultural evolution. As he explains, this was done to avoid "confusing terminology ('language faculty', 'universal grammar') [that] has often clouded the substantive issues in the past."

All in all, these examples show that the terminological fuzziness surrounding the notion of language in language evolution research has been subject to much criticism. In the remainder of this paper, we will defend a position directly opposed to these influential voices, and to a degree contrary to standard intuitions. We will claim that agreement on the ontology of language is not a *sine qua non* for good and productive research in language evolution, and question not

[4] Following up on Haspelmath's comment, Mendívil-Giró (2019) argues that the concepts of language evolution and change should be kept apart. He defines language as "a historically modified mental organ" and argues that language change gives rise to language diversity, as new languages can develop out of another language. This process, however, is distinct from the emergence of language from non-language, which, he argues, is beyond the scope of historical linguistics and rather brought about by processes studied in evolutionary biology. However, he also acknowledges that whether or not language evolution and change are seen as a continuum depends on the underlying conceptualisation of language: While he defends the view of language as externalisation patterns of a species-specific Faculty of Language, he concedes that "the conflation of the process of linguistic change and the process of the evolution of FL is natural in those approaches that conceive of languages as social and cultural objects."

only whether such an agreement may be possible, but also whether it would be necessarily desirable.

As our main point, we question the value of a single top-down notion of language. Such a definition seems not to be possible for language evolution as a field of research both because of the nature of highly interdisciplinary scientific practice in the field and due to the special and multi-faceted ontology of language as an object of study. The only type of overarching definition of language is a bottom-up one, as a family-resemblance notion derived from the patterns of use of the word *language* in everyday language(s) and reflected in the patterns of actual research practice. This leads to definitions of language evolution that could be seen as disappointingly broad and possibly circular, as in "[l]anguage evolution researchers are interested in the processes that led to a qualitative change from a non-linguistic state to a linguistic one" or "[w]e can characterise the study of language evolution as being concerned with the emergence of language out of non-language" (Scott-Phillips & Kirby 2010: 412). However, a 'usage-based' understanding of *language* and *language evolution* has the virtue of actually capturing how these terms function in the scientific community of language evolution researchers.

In addition—and again largely as a consequence of the nature of scientific practice—we suggest that conceptual diversity may actually have beneficial consequences. In short, even though there are considerable differences in the understanding of *language* (resulting in differences in the understanding of the explanatory goal of the entire enterprise of language evolution research), this does not necessarily imply incommensurability (see especially Pleyer & Hartmann 2019 for a supporting argument). It also does not necessarily hinder local progress on individual phenomena relevant to the understanding of language evolution, and further, may even be conducive to progress more globally and thus benefit the entire field of research.

3. FLN/FLB

We return to the distinction mentioned in Section 2, between the faculty of language in the narrow versus broad sense (FLN/FLB; Hauser et al. 2002), which for several reasons provides a perfect case in point. Not least among these reasons is the central status of FLN/FLB to language evolution research, and even beyond: it is one of the very few terminological-conceptual exports from the field of language evolution to the study of language and cognition at large, included in important linguistic and interdisciplinary tertiary literature such as *The Stanford Encyclopedia of Philosophy* (Cowie 2008) or *The Encyclopedia of Language and Linguistics* (Tincoff & Hauser 2006). Likewise, the original source of the distinction, the paper by Hauser et al. (2002) already mentioned above is doubtlessly among the most influential works in the field, and probably its most widely cited article (5,716 Google Scholar cites as of 11 April 2020).[5]

[5] The rest of this section develops an argument originally stated in Wacewicz (2012).

In what follows we will not discuss the content and agenda of Hauser et al. (2002) and the ensuing debates in detail.[6] Instead, we focus on establishing two points that are central to the rest of our argument:

(1) The proponents of FLN have defined it twice, in fundamentally discrepant ways: the definition originally formulated in Hauser et al. (2002) and the one later formulated in Fitch et al. (2005) describe two distinct entities (not simply different versions of the same entity).

(2) This fundamental discrepancy has remained virtually completely unrecognised in the literature, which does not appear to have had major consequences for language evolution research, even in frameworks that explicitly adopt and rely on the FLN/FLB distinction.

In sum, the notion of FLN is a highest-profile case in point, serving to illustrate that even a very fundamental confusion about the explicit, top-down definition of a central notion—(the faculty of) language—can remain in the background of actual research practice, without readily perceptible detrimental effects.

3.1. The 2002 vs. 2005 Definitions of FLN

*3.1.1. The 2002 **D**efinition of FLN*

In their original paper, Hauser et al. (2002) define FLN as the "computational core" of the language faculty: The cognitive subsystem responsible for generating the discrete infinity of linguistic expressions. FLN so defined is a term internal to linguistic theory, and the distinction between FLN and FLB is one based solely on a particular theoretical account of language (and thus of 'the language faculty'). On this original definition, FLN is one part of the more general faculty of language in the broad sense (FLB), which also includes at least two other major components: the sensorimotor (SM) subsystem and the conceptual-intentional (CI) subsystem, which Hauser et al. (2002) illustrate with examples but do not explain in further detail.

Although later misinterpreted (see below), the definition of FLN as the "computational core" is quite unambiguous, and systematically recurs throughout the paper, for example:

> *Faculty of language–broad sense* (FLB). FLB includes an internal computational system (FLN, below) combined with at least two other organism-internal systems, which we call "sensory-motor" and "conceptual-intentional". (Hauser et al. 2002: 1569–1570 [italics in the original])

[6] Readers of *Biolinguistics* are likely familiar with the Hauser–Chomsky–Fitch vs. Pinker and Jackendoff debate (Hauser et al. 2002, Pinker & Jackendoff 2005, Jackendoff & Pinker 2005, Fitch et al. 2005), on which there is a large body of existing commentaries (see, e.g., Parker 2006, Wacewicz 2007, Lewandowska-Tomaszczyk 2008, Balari & Lorenzo 2012, Barceló-Coblijn 2012, Wróbel 2012, Boeckx 2012).

> FLB includes sensory-motor, conceptual-intentional, and other possible systems (which we leave open); FLN includes the core grammatical computations that we suggest are limited to recursion.
>
> (Hauser et al. 2002: 1570 [Figure 2, caption])

> *Faculty of language–narrow sense* (FLN). FLN is the abstract linguistic computational system alone, independent of the other systems with which it interacts and interfaces.
>
> (Hauser et al. 2002: 1571 [italics in the original])

It is important to observe that Hauser et al. (2002) discuss the property of 'uniqueness to humans' and tag it, clearly and repeatedly, as a *hypothesis* about FLN.

> By this *hypothesis*, FLB contains a wide variety of cognitive and perceptual mechanisms shared with other species, but *only those mechanisms underlying FLN—particularly its capacity for discrete infinity—are uniquely human*. (Hauser et al. 2002: 1573 [emphasis added])

> Second, although we have argued that most if not all of FLB is shared with other species, whereas *FLN may be unique to humans*, this represents a tentative, testable *hypothesis* in need of further empirical investigation. (Hauser et al. 2002: 1576 [emphasis added])

> *Hypothesis 3: Only FLN is uniquely human*.
>
> (Hauser et al. 2002: 1573 [italics in the original])

*3.1.2. The 2005 **D**efinition of FLN*

As stated above, the 2005 article by Fitch, Hauser, and Chomsky (published as a reply to Pinker & Jackendoff 2005) provides a different definition of the FLN/FLB distinction. Fitch et al. (2005) claim there that FLN is defined as being unique to both humans and language:

> […] given that language as a whole is unique to our species, it seems likely that some subset of the mechanisms of FLB is both unique to humans, and to language itself. We dubbed this subset of mechanisms the faculty of language in the narrow sense (FLN).
>
> (Fitch et al. 2005: 180–181)

> We thus made the further, and independent, terminological proposal to denote that subset of FLB that is both specific to language and to humans as FLN. To repeat a central point in our paper: FLN is composed of those components of the overall faculty of language (FLB) that are both unique to humans and unique to or clearly specialized for language. (Fitch et al. 2005: 182)

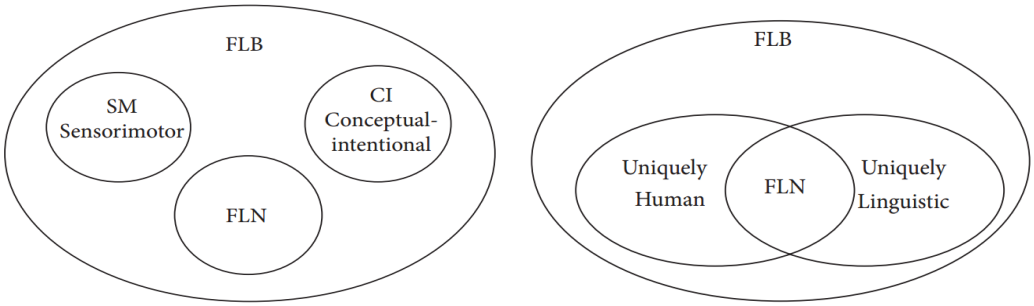

Figure 1: The FLB/FLN distinction in Hauser et al. 2002 (left) and in Fitch et al. 2005 (right).

3.1.3. Summary

In sum, the two papers by Hauser, Chomsky and Fitch reverse the relation between the essential-*cum*-definitional and the accidental-*cum*-hypothetical properties of FLN (Figure 1). The 2002 paper defines FLN as a computational core and hypothesises its human uniqueness; conversely, the 2005 paper defines FLN as uniquely human and hypothesises that it includes a computational core (Table 1). As one example of the consequences, if we accept the 2005 *definition* of FLN as that which is 'both unique to humans and unique to or clearly specialized for language', then an unpacked Hypothesis 3 from the 2002 text effectively becomes *only that which is uniquely human (and linguistic) is uniquely human*. This is why—emphatically—the 2002 and 2005 definitions are not just different variants of the same definition, but two different definitions that are discrepant in a strong sense. Unsurprisingly, they produce different answers to consequential questions, such as 'Can homologous traits be part of FLN?' Further, they cause a number of literal contradictions, for example:

> The contents of FLN are to be empirically determined, and could possibly be empty, if empirical findings showed that none of the mechanisms involved are uniquely human or unique to language, and that only the way they are integrated is specific to human language. The distinction itself is intended as a terminological aid to interdisciplinary discussion and rapprochement, *and obviously does not constitute a testable hypothesis*. (Fitch et al. 2005: 180–181 [emphasis added])

> Second, although we have argued that most if not all of FLB is shared with other species, whereas FLN may be unique to humans, this represents a tentative, testable hypothesis in need of further empirical investigation. (Hauser et al. 2002: 1576 [emphasis added])

3.2. Reception and Takeaway

Interestingly, Hauser, Chomsky and Fitch themselves have never addressed the inconsistency and may even remain unaware of it. In Fitch et al. (2005: 181–183), the authors maintain that the later, 2005, definition, and hence the 'uniqueness to humans' criterion, was in place in the original paper (which, as demonstrated

	FLN in Hauser et al. (2002)	**FLN in Fitch et al. (2005)**
Definition	FLN is the core computational mechanism of the FLB.	FLN is the part of FLB that is unique to humans and unique to language.
Hypotheses	• FLN is the part of FLB that is unique to humans (i.e. FLN is unique to humans, and no other part of FLB is). • FLN can be equated with recursion.	Only recursion is unique to humans.

Table 1: Definitions of, and hypotheses about, FLN in Hauser et al. 2002 vs. in Fitch et al. 2005.

above, is incorrect). In later work, they continue to use the term FLN in these two incompatible senses, for example:

> FLN—and especially the mechanism of recursion—was defined by Hauser et al. (2002) as a computational process that is responsible for the generative and hierarchical properties of narrow syntax.
>
> (Tincoff & Hauser 2006: 536)

> HCF proposed a distinction between the faculty of language in the broad [FLB] and narrow sense [FLN]. FLB is simply those processes of the mind that are both necessary and sufficient to support language. Thus, for example, attention is involved in language processing but is neither unique to language nor unique to humans. FLN includes those processes that are both uniquely human and unique to language... HCF hypothesized that FLN, though potentially an empty set, may only include the computational resources subserving recursion and their interface or mapping to the conceptual-intentional [semantics] and sensory-motor [phonetic] systems. (Hauser et al. 2007: 105)

As for the larger community, the discrepancy seems to have been essentially overlooked in the rather sizable body of commentaries that followed in the wake of the original FLN paper and the debate of Hauser et al. with Pinker and Jackendoff. Most interestingly, the rejoinder by Jackendoff & Pinker (2005) does not clearly expose the discrepancy, but instead proceeds to address the revised definition by Fitch et al. (2005). Other commentators tacitly assume FLN to only have a single definition, sticking to either the 2002 'computational core' definition (e.g. Armstrong & Wilcox 2007, Johansson 2005, Kurcz 2004, Lewandowska-Tomaszczyk 2008) or to the 2005 'uniquely human + uniquely linguistic' definition (e.g., Okanoya 2007, Parker 2006, Számadó & Szathmáry 2006). It is not uncommon to see the 2005 definition incorrectly attributed to the 2002 paper, such as in Samuels (2009: 356): "Hauser et al. (2002) define FLN as those aspects of the language faculty that are unique both to humans and to language". Occasionally,

commentators equivocate between the two interpretations, conflate them or use them interchangeably without noting their mutual incompatibility (e.g., Kinsella 2009).

In summing up this part of our argument, we wish to underscore that the definitional problem itself, that is the existence of two parallel definitions of FLN, is not unusual in science (since competing definitions of technical terms are commonplace) and is tangential to our present interests. Again, the point we make here is not that the definitional discrepancy exists; rather, our point is how the discrepancy has continued to go *essentially unnoticed*. However surprising—and contrary to an earlier analysis by one of us (Wacewicz 2012)—this provides a striking demonstration that the specific wording of the top-down definitions of language was inconsequential to the research practice of the field.

In the next section, we will focus on a number of approaches that have led to significant progress in the field regardless of the fact that they do not directly map onto the different definitions of FLN/FLB. Instead, they represent 'multicomponent' approaches to language evolution (cf. Fitch 2017, Benítez-Burraco & Progovac 2020) that outline important aspects of 'language' and 'language evolution' and eschew the kind of restrictive definitions outlined in Section 3. As such, these approaches are further evidence of our view that the definitional discrepancies discussed above did not hinder progress in the field. Instead, these approaches show the importance not of top-down definitions of language, but of focusing on particular aspects of and hypotheses about language and investigating their relation to other factors relevant to language and its evolution (see also Roberts et al. 2020).

4. *Language* as a Family Resemblance Category in Language Evolution Research

In this section, we flesh out our argument with a brief survey of presently influential lines of language evolution research. Our main goal here is to illustrate the current breadth of the field and show how this breadth brings with it conceptual diversity as an inevitable consequence. We propose a categorisation into four general approaches that constitutively differ in how they conceptualise language, where 'constitutiv' means such differences that preclude a neat grouping under a single common definition. Nevertheless, we wish to show that these conceptualisations are not entirely disjunct but are in fact characterised by patterns of overlapping similarities—in other words, these uses of *language* form a family resemblance category. As an additional point, we also mention some benefits of this conceptual diversity, i.e. ways in which these four approaches have been mutually valuable and invigorating. We will first discuss conceptualisations of language as a multimodal phenomenon (Section 4.1), before turning to approaches that treat language as a complex adaptive system (Section 4.2). We will then discuss approaches that see language as a form of social interaction (Section 4.3), and finally we will explore approaches that look at language from the perspective of the language-ready brain (Section 4.4).

Since, as we emphasise in Section 5, language evolution is a fast-changing field, we focus on approaches that we consider as particularly prolific and

impactful in the last decade. As such, these approaches also represent trends likely to further gain in importance in the near future (see also Nölle et al. 2020). However, such a selection can never aspire to being fully objective, and in particular our temporal perspective means leaving out foundational work by highly prominent but early language evolution scholars, such as Bickerton (1990), Dunbar (1996), Deacon (1997) or Jackendoff (2002). Such works were central to the inception of language evolution as a science, and are still used as reference points for the discussion of specific topics in the modern day science of language evolution—for instance, Bickerton and Jackendoff often feature in debates about the nature of protolanguage (e.g., Fitch 2010), and Dunbar's views are referenced with regard to the problem of the social preconditions of language emergence (e.g., Dor et al. 2014, Zlatev 2014).

4.1. Language as a Multimodal Phenomenon

One approach that has become ever more popular in the last decade sees language as a multimodal phenomenon. As Vigliocco et al. (2014) forcefully argue, "speech signals are invariably accompanied by visual information on the face and in manual gesture" (Vigliocco et al. 2014: 1). Such a view differs considerably from more traditional conceptualisations, on which non-verbal behaviour (e.g., as defined and taxonomised in the seminal paper by Ekman & Friesen 1969), and primarily gesture, supports but is definitionally separate from linguistic communication. This definitional framework has very profound consequences for language evolution: Since gesture and occasionally other forms of non-verbal communication are inseparable from (spoken) language, it follows that even though the nonverbal component and the verbal component may be analytically distinct, their evolutionary origins constitute an indivisible explanatory target. In other words, at least for the purposes of explaining its evolutionary origins, gesture must be considered as an integral part of language. In the remainder of this section, we illustrate the view of language as a multimodal phenomenon by discussing three exemplary frameworks, focusing on the question of how language is conceptualised in each of these approaches: Adam Kendon's idea of languaging, David McNeill's growth point, and Jordan Zlatev's mimesis hierarchy.

4.1.1. Adam Kendon: Languaging

Adam Kendon rejects the traditional idea that a language forms an abstract system of rules (as proposed by e.g. de Saussure 1916), and opposes the view that language has systemic properties. Instead, Kendon (e.g. 1990, 2004) takes a dynamic, usage-oriented view, whose roots can be traced back to Humboldt's *energeia* and more recently to Goffman's interactionism. Kendon (e.g. 2014a, 2017) sometimes uses the term *languaging*, to underline the dynamic character of language, or *gesture-speech ensemble* (Kendon 2004: 108), to underline its multimodal character. He argues that language involves "the mobilization of several different semiotic systems in different modalities and deployed in an orchestrated relationship with one another" (Kendon 2014a). The most linguistic element of this orchestration is speech, which has linear structure and is organised by the morpho-

syntactic component (Kendon 2014a). However, as Kendon insists, "the 'natural' state of spoken language" (Kendon 2014b: 76) is the context of physical "co-presence", in which the transmission of meaning, both propositional and non-propositional, depends on speech being coordinated with "extra-oral bodily action"—hand and arm movements, postural shifts, eye contact or facial expressions (Kendon 2004, 2011). The traditional focus on the systemic properties of language results from abstracting it from this "natural" state, and language so construed is of secondary importance both in ontogenetic and evolutionary terms (Kendon 2014b: 72).

The basic unit of language (or alternatively, 'languaging') is the utterance, which is the coming-together of speech and extra-oral visible action to translate ideas into "observable behavior, which may be read by others as reportive of those ideas" (Kendon 1980: 208; see also Kendon 2004). The meaning, including propositional meaning, of an utterance results from an interplay of speech and extra-oral bodily visual actions. Importantly, Kendon sees extra-oral visible bodily action as so closely coordinated with speaking that it has to be understood as an integral component of language. The idea of language as the interaction and co-expression of speech and body movement is of crucial importance to Kendon's view on language origins. He subscribes to a uniformitarian hypothesis, according to which "the early steps of language evolution also consisted of multi-modal signals, instead of being predominantly hand-based or vocalization based" (Kendon 2014b: 69). Hence, he looks for such an evolutionary context that could explain an early integration between vocal-auditory and visual-bodily semiotic resources and argues for a praxic origin of language—in other words, he argues that language is rooted in concrete actions. Accordingly, there was one, albeit complex, executive system for oral-laryngeal and manual action, which served such purposes as mastication and food-handling (cf. MacNeilage 2008). Later, this system was rededicated "in the service of communicative action" (Kendon 2014b: 72): Articulated vocalisation developed early in the hominin line to manage and maintain complex social relations; gesture, understood broadly as deliberate and expressive-communicative movement, also emerged early from the primary practical, manipulatory function of the hand and forelimb. On Kendon's view, the common origin of speech and gesture (as defined above), as well as the same evolutionary trajectory (from the praxic to the communicative function), explains what he refers to as "the 'natural' state of language" (see above):

> […] gestures that are so often a part of speaking are neither supplements nor add-ons. They are integral to speaking. They are so because they are derived from practical manipulatory actions from which speaking itself is also derived. Looked at in this way, we can better understand why it is that visible bodily action is mobilized when speakers speak and why, more generally, when language is used in co-present interaction it always involves *poly-modalic* forms of action.
>
> (Kendon 2014b: 75)

To support his scenario, Kendon extensively appeals to research on the integration of speech and gesture in the communication of modern humans, but also to developmental and neurocognitive evidence, such as the coordination of syllabic babbling with hand movements (Kendon 2014a, cf. Ejiri & Masataka 2001) or the role of Broca's area in controlling hand movements as well as movements of the expressive muscles of the face (Kendon 2014b: 69, cf. Willems et al. 2007, Aboitiz 2012). Kendon argues against the view that asserts continuity between ape gestural communication and modern human gestures, which has been emphasised by gestural accounts of language origin (e.g., Hewes 1977, Arbib 2012, Corballis 2013).

4.1.2. David McNeill: Growth Points

In McNeill's model, speech and gesture synergistically express the same overall meanings while remaining semiotically distinct and responsible for the transmission of different aspects of the message: speech for propositional content and gestures for imagistic content. According to McNeill, the *stroke* (i.e. the most pronounced phase) of a gesture accompanies the semantically most prominent element of the utterance. In this way, the Growth Point, the basic unit of thinking, becomes externalised. Here, McNeill departs from Kendon's account of modern human communication, which does not posit a categorical division of labour between speech and gesture, but rather argues for their functional interplay, for example, gesture can transmit propositional aspects of meaning (see above), while speech includes vocal means of expressing emotional-imagistic content, as in the case of paralinguistic features (e.g., emotional prosody) or iconic vocal phenomena, as in ideophones, phonesthemes, reduplication or word lengthening (Kendon 2008). They also disagree about the definition of gesture. McNeill (1992, 2012) would further limit (prototypical) gestures (i.e. co-speech gestures) to spontaneous and idiosyncratic hand and arm movements that are functionally integrated with speech. As we have seen, Kendon's understanding of gesture extends beyond the category of co-speech gestures and embraces any deliberately communicative bodily movement (hence, the use of the term 'kinesic'), including postural shifts, eye contact or facial expressions (Kendon 2004, 2011).

The idea of a tight integration between spoken messages and co-speech gesture is also central to McNeill's theory of language evolution, the critical moment of which is the integration of gestural and vocal communication, both at the level of cognition and expression (McNeill 2012). The claim is that language originated from the coming together of vocalisation and gesture to form a propositional-imagistic dialectic. Like Kendon, McNeill submits a uniformitarian explanation as the rationale of his hypothesis: language in its beginnings was qualitatively similar to what it is now; but it should be remembered that he proposes a more limited view of what language is than Kendon. The critical element in the formation of the propositional-imagistic dialectic was the 'twisting' of mirror neurons, whereby they began "to respond to one's own gestures, as if they were from someone else" (McNeill 2012: 65). To support this idea, McNeill paraphrases Mead (1974): "[A] gesture is a meaningful symbol to the extent that it arouses in the one making it the same response it arouses in someone witnessing it" (2012: 180; cf. Arbib's

parity requirement in Section 4.4.1 below). As this gestural system was co-orchestrated with vocalisation, the Growth Point emerged.

It should be noted that McNeill does not provide any evolutionarily grounded pressures that could have been responsible for these changes. In fact, he ventures two rather different accounts of how speech started, deriving it either from ingestion, which assumed vocal properties and was subsequently orchestrated with gesture (2012: 180–181), or from the type of communication that is found in extant non-human apes, such as "chimp gestures with vocalization" (2012: 195). Although McNeill refers to the 'twisting' of mirror neurons and the voice-gesture integration as adaptations, he actually describes them as saltational leaps, not unlike Chomsky's idea of a lucky mutation giving rise to the operation of Merge, which first endowed humans with a language of thought and then with the communicative use of it (Berwick & Chomsky 2016).

*4.1.3. Jordan Zlatev: **T**he **M**imesis **H**ierarchy*

A different account of language and language evolution is put forward by Zlatev (2008, among others). Zlatev objects to the very term 'multimodality' as used by Kendon and McNeill (but also many other researchers, see e.g., Wacewicz & Żywiczyński 2017); for him, language and gesture are two distinct semiotic systems, i.e. systems of signs and relations between them (Zlatev et al. 2020), which are characterised by different design features. For instance, linguistic signs are mainly conventional, gestural signs mainly iconic; the syntagmatic relations between linguistic signs are compositional, between gestural signs, they are linear; language uses double articulation, gesture does not. Next, language can utilise different modalities: Vocal in the case of speech, material in the case of writing, bodily in the case of signed languages, and so on (Żywiczyński & Zlatev, in press). Hence, face-to-face communication is typically both polysemiotic, that is it makes use of different semiotic systems (most importantly, language and gesture), and multimodal, that is it makes use of different communication channels (most importantly, vocal for speech and bodily for gesture; Zlatev 2019).

Zlatev's key theoretical concept is mimesis, adapted from Donald (1991, 2001). His most recent definition of bodily mimesis is the following:

> […] [A]n act of cognition or communication is an act of bodily mimesis if: (1) it involves a *cross-modal mapping* between exteroception (e.g. vision) and proprioception (e.g. kinesthesia); (2) it is *under conscious control* and *is perceived by the subject to be similar to* some other action, object or event, (3) the subject *intends* the act to stand for some action, object or event *for an addressee,* and for the addressee to recognize this intention; (4) it is *not fully conventional and normative*, and (5) it does *not divide (semi)compositionally* into meaningful sub-acts that *systematically relate* to other similar acts, as in grammar. (Zlatev 2014: 206)

On this basis, Zlatev proposes an evolutionary and developmental model known as the *mimesis hierarchy* (Zlatev, 2008). The rudimentary form of *proto-mimesis,* based on requirement (1), is found in activities like emotional and

attentional contagion, and is common for all primates. The more advanced form of *dyadic mimesis* (based on 1 and 2) involves volition and imitation, but not true representation or sign-function; it is common for all great apes. Only at the next level (based on 1, 2 and 3), referred to as *triadic mimesis*, do mimetic acts gain a clear sign-function, as well as Gricean communicative intentions (i.e. that the addressee should understand that a communicative act is being performed for their benefit). Further, point (4) distinguishes mimesis from a conventionalised protolanguage and point (5) from language proper.

This provides a useful conceptual apparatus, but does not answer key questions such as what drove the evolutionary process, as well as more specific aspects of how the transition from triadic mimesis (i.e. pantomime) to protolanguage and language took place, including the shift from a dominance of gesture to a dominance of vocalisation. Zlatev (2016) addresses these gaps, but in a somewhat schematic matter. With respect to evolutionary pressures, Zlatev appeals to an increase of pro-sociality in hominins (cf. Tomasello 2008), which might in turn have been ecologically driven by the reproductive strategy of cooperative breeding, where the biological parents receive help in rearing their young from the wider group (Hrdy 2009). Concerning the gradual transition to vocalisation, this is sought in the nature of pantomime itself: a hybrid system that is polysemiotic (i.e. combines various sign and signal systems) and multimodal (i.e. involves different sensory channels). The dominant semiotic system in pantomime is claimed to have been *robustly iconic gesture* (cf. the notion of primary iconicity; Sonesson 1997). The transition towards language entailed a gradual loss in iconicity along various parameters (see Zlatev et al. 2020 for details). Zlatev (2016) attempts to motivate the gradual transition from gesture to vocalisation when the need for less iconicity and more 'arbitrariness' arose.

But while language (realised as speech, writing or signing) may be the dominant system in modern human communication when it comes to expressing propositions and narratives, it is rarely used alone, but alongside other semiotic systems such as gesture and depiction (e.g. Green 2014): Polysemiotic communication. An advantage of the mimesis/pantomime approach is that it can help explain this, as pantomime consisted of gesture, vocalisations as well as 'protodrawing', when gestures left marks on surfaces such as sand (Zlatev 2019, Zlatev et al. 2020).

*4.1.4. Language as a **M**ultimodal **P**henomenon: Taking **S**tock*

The defining feature of the views on language discussed above is its multimodal character. However, the term 'multimodality', especially as used by Kendon and McNeill, conflates multimodality itself, i.e. the use of different sensory modalities, with polysemioticity, that is the use of different sign systems, most importantly speech and gesture. Beyond the general consensus that language is multimodal, there are differences in the way these approaches account for language and its separability from other semiotic systems. For McNeill and Kendon (cf. the latter's idea of languaging), language and gesture are two manifestations of the same system—importantly, this system is at its core both communicative and cognitive. Zlatev enumerates criterial attributes of language, in contradistinction to the

criterial attributes of gesture, but emphasises that human-specific communication is nevertheless inherently multimodal and polysemiotic; hence, language should be seen as cognitively distinct from the semiotic system of gesture and other semiotic resources but is inseparable from them in actual communicative behaviour. In this regard, his position is similar to that of Levinson (see Section 4.3 below).

All of these authors underline that language is species-specific, but they also posit its continuity with ape cognition and communication, albeit with various degrees of emphasis. They also agree about the watershed in the evolution of language, which was of semiotic nature and consisted in the emergence of iconic gestures, although again, they differ in their account of how abrupt the semiotic breakthrough was. Related to that point is the division of labour between biological and cultural evolution. In multimodal approaches, the bulk of biological preadaptations for language, mainly related to the organisation of the neural infrastructure, happened prior to the semiotic breakthrough and facilitated it. The later course of language evolution was almost exclusively the domain of cultural evolution, which led to the emergence of arbitrary symbols and grammar. In accounting for both protolinguistic beginnings of language and its later phases, the multimodal approaches emphasise the importance of cognitive and social factors, for example, the development of complex forms of Theory of Mind, intentionality or cooperation, and treat modern language as integrated in the human socio-cognitive niche. In this regard, they are highly compatible with theories that treat language both as a complex adaptive system and as a form of social interaction.

4.2. Language as a Complex Adaptive System

Another influential perspective on language in language evolution is constituted by approaches that view language as complex adaptive systems (CAS) that emerges from social interaction across the timescales of biological evolution, cultural evolution and ontogenetic development (e.g., Steels 2000; Beckner et al. 2009; Kirby 2012). These approaches place different points of emphasis on particular aspects of complex-adaptive processes—some stress their direct relevance to language emergence; others focus on the (socio-)cognitive mechanisms that underlie them. However, we discuss them together as they share the underlying view of language as being multifactorial and dynamic, and whose evolution is channelled by cognitive, interactive-communicational and cultural-historical contexts.

Complex adaptive systems are defined as "processes involving a number of interacting parts which give rise to emergent processes that show the appearance of design." (Kirby 2012: 590). This idea has gained momentum in both theoretical and empirical approaches in language evolution. On a more theoretical plane, it has been adopted by many practitioners of usage-based approaches, for example, those resonating with Construction Grammar and Cognitive Linguistics (see, e.g., Pleyer & Winters 2014). But it has also been adopted as a framework for computational modelling and behavioural experiments. Both of these domains of empirical research adhere to the concept of Iterated Learning: "[A] particular kind of cultural transmission" whereby "a behaviour arises in one individual through induction on the basis of observations of behaviour in another individual *who acquired that behaviour in the same way*" (Kirby et al. 2014: 108, emphasis in original).

Iterated Learning has informed a number of computational models of language evolution (e.g., Smith et al. 2003, Smith & Wonnacott 2010), and has extensively been used as the paradigm for lab experiments on the emergence of novel communication (e.g., Kirby, Cornish & Smith 2008, Garrod et al. 2010, Tinits et al. 2017, among many others). Iterated Learning crucially depends on the concept of a *transmission bottleneck*: The number of possible utterances is larger than an agent can observe in their lifetime, which is why language adapts to the agents' learning biases (see e.g., van Trijp 2011). Thus, "language is adapting in such a way as to ensure its own survival through the transmission process" (Kirby 2012: 595). Importantly, it is not only the users of a language but also languages themselves that undergo adaptation (see also Deacon 1997, Christiansen & Chater 2008).

The Iterated Learning model is, in principle, not only applicable to language but also to other cultural artefacts, traditions or communicative codes such as writing systems (e.g., Garrod et al. 2010). However, most interestingly for our present concerns, it operationalizes a certain general concept of language: *Signals from a finite signal space are mapped to meanings from a finite meaning space* (see, e.g., Kirby et al. 2008, Cornish 2010). While this is of course a deliberate simplification for modelling purposes, it bears many similarities with the widespread view of languages as inventories of form–meaning pairs. It can be traced back at least to Saussure's (1916) sign concept and has been adopted explicitly in Construction Grammar and other usage-based perspectives, where constructions, that is pairings of form and meaning/function, are understood as the basic units of linguistic description. Of course, the notion of the centrality of the Saussurean sign and the evolution of its components is shared by a variety of approaches, including Bouchard (2013) and Hurford (1989, 2007, 2012), who himself mentions Construction Grammar as a suitable framework for investigating the evolution of language (see Hurford 2012: 348–362).

A major advantage of the CAS perspective on language is its generality: The evolution of language and of other cultural traits can be investigated in a shared theoretical framework. However, this generality also entails that its explanatory value is limited. This is why, for example, Larsen-Freeman (2017) characterises Complexity Theory as a 'metatheory' that warrants different object theories. Among the more widespread object theories are a number of usage-based approaches to language, some of which adopt a generalized theory of evolution, or adapt ideas from evolutionary biology (see, e.g., Croft 2000, 2011, Ritt 2004). One consequence is that many approaches within this framework do not take an explicit stance on the issue of unimodality vs. multimodality as signs can potentially be constituted through multiple modalities. The Iterated Learning framework programme might not explicitly label language as multimodal, but this approach agrees with multimodal approaches in stressing that other modalities than the vocal-auditory modality are subject to social-interactional, communicative, and learning pressures and played an important role in the evolution of language (Verhoef et al. 2014, Little et al. 2017, Motamedi et al. 2019).

Given the wide range of factors discussed in the emergence of language, language and its development are clearly not conceived of as domain-specific developments but as being part of a broader suite of cognitive and interactional processes, although this distinction is generally becoming increasingly blurred in a

number of approaches (cf., e.g., Pleyer & Hartmann 2019). From a Complex Adaptive Systems perspective, language involves a multitude of different cognitive and physiological capabilities, including but not limited to embodied cognition (e.g., Bergen 2012), intention reading and pattern finding (e.g., Tomasello 2009), associative memory (e.g., Divjak 2019), and 'massive storage' (Hurford 2012: 261). There is therefore no agreement on components that are criterial for language, although social cognitive abilities often take centre stage. This idea is of course not unique to the CAS approach, but it is a stance that has arguably been taken more forcefully in CAS approaches than in most other frameworks.

An important consequence of viewing language from a CAS perspective is that the boundary between cultural and biological factors gets blurred. As Pleyer & Hartmann (2019) have pointed out, this is in line with recent developments in biolinguistics that increasingly adopt an evo-devo perspective (e.g., Benítez-Burraco & Boeckx 2014, Martins et al. 2016, Bowling 2017, see also Section 4.4.2). This is also one important aspect in which the conceptualisation of language as CAS has influenced, and continues to influence, research on language (evolution): Language is investigated on a par with other phenomena that can be seen as results of cumulative evolution—for example, in the framework of cultural evolution theory, which has become increasingly influential in recent years (see, e.g., Richerson & Boyd 2005, Mesoudi 2011). As a consequence, the challenge that language cannot be easily delineated from other phenomena becomes part of a research programme that aims at taking the continuous nature of the phenomena it investigates into account. Even though most approaches that can be seen as belonging to the CAS framework aim at overcoming the strict divide between biology and culture, it seems fair to say that most of them view language, in the first place, as a cultural and communicative phenomenon.

Given that the feedback loop between individual actions and emergent phenomena on a population level is part and parcel of the CAS model, its proponents see language both as an individual and as a supra-individual/social phenomenon, even though different approaches may emphasise one of these two aspects more than the other. In this regard, it is also quite instructive to take a look at the brief history of Construction Grammar, which originally took "a synchronic and mentalist perspective" (Hilpert 2013: 1) by trying to describe the linguistic knowledge of individuals. However, Construction Grammar increasingly widened its scope to models of language variation and change, which makes a population-level perspective necessary. As the emergence of structure is a dynamic, cultural process, there are in principle no categorical distinctions between language and non-language. So, although CAS approaches assume continuity between language and other forms of communication, language is usually seen as species-specific in the sense that the different components that make up the 'mosaic' of language may also be found in other animals, but they are only fully in place in modern humans (Elman 1999).

CAS approaches more generally have arguably had a substantial impact on the research landscape in the field of language evolution and the CAS approach in general has been widely adopted (from Steels' pioneering 2000 article, to Beckner et al.'s widely-cited 2009 paper, to Kirby's 2012 handbook article), probably partly due to its compatibility with a very broad spectrum of approaches: While

its foundational assumptions have been widely shared across various approaches in language evolution research for a long time, the CAS framework offers a convenient terminological toolkit for making these assumptions explicit, and it also invites researchers to broaden the scope of their research by putting the phenomena they investigate into perspective, which, as we have seen, also has consequences for the question of how a notion like *language* is conceptualized.

A second reason why CAS approaches are gaining popularity is not theoretical by nature, but empirical, as this framework can accommodate more easily than others increasing evidence of complex feedback effects among a variety of factors—physical, cognitive, behavioural, environmental—involved in language evolution, acquisition, and use. Among others, we can mention the constraints imposed on language form (sounds, morphology) by physical (temperature, humidity) and cultural factors (population size, topology of social networks; Lupyan & Dale 2010); or the differential impact of language features on cognitive abilities (such as working memory in Amici et al. 2019). Ultimately, CAS approaches fit better than others with views of human evolution that see the emergence of modern cognition and culture as the result of a complex feedback loop between our biological endowment and our cultural practices, instead of as the outcome of a linear evolutionary process, with modern cognition appearing first and modern culture/behaviour happening later. In the case of language evolution, it is now viewed as the outcome of a feedback loop (seemingly ongoing) between our biological language-readiness, faculty of language, or linguisticality, and our language-supported cultural practices. This ultimately entails that the boundaries between language evolution—understood as the processes that give rise to fully-fledged human language—and diachronic language change become blurred as the biological processes that lead to the emergence (and further development) of language cannot be neatly separated from the cultural ones. Smith (2018) mentions two areas where the assumption of a close interaction between culture and biology seems plausible: On the one hand, he summarizes de Boer's (2000) modelling work on phonological niche construction in the evolution of vowel systems, in which a selection pressure for individuals with more fine-grained articulatory or perceptual capabilities interacts with cultural-evolutionary pressures that continuously push the vowel system to the limits of the available articulatory or perceptual space. On the other hand, Smith (2018) argues that "process of gene–culture co-evolution might also act to constrain cultural evolution, by imposing biological constraints on the kinds of systems which can be learned", which may be particularly relevant for the evolution of syntax.

A recent promising spin-off of CAS approaches is the self-domestication hypothesis of language evolution, also adopted by Kirby (Thomas & Kirby 2018). In a nutshell, the existence in humans of features of domesticated mammals compared to wild extant primates is claimed to account for both the emergence of a modern language-ready brain, mostly via a biological mechanism, and of modern languages, endowed with all the features that are familiar to linguists, mostly via a cultural mechanism. The cornerstone of the hypothesis is the reduced reactive aggression and the increased social tolerance brought about by self-domestication, which favoured language teaching and learning, and ultimately, the complexification of linguistic forms. However, brain and cognitive changes are also

expected, either directly, as a consequence of domestication processes, or indirectly, via a feedback effect triggered by the new social environment and the new language forms and uses (see Benítez-Burraco 2020, Benítez-Burraco & Progovac 2020 for details).

In sum, the CAS approach thus provides a convenient metatheory that is compatible with a variety of different approaches to language and its evolution, including the view of language as a form of social interaction, which will be discussed in the next section. As it is applicable both to cultural and to biological systems, it is also very much in line with recent approaches that reject a categorical divide between culture and biology.

4.3. Language as a Form of Social Interaction

In explaining the emergence of human language, many scholars stress the social embedding of language, and its resultant importance for language evolution. The centrality of social cognition has a long history in the research on the evolution of language and mind (e.g., Byrne & Whiten 1988, Cheney & Seyfarth 2007, Dor et al. 2014, Dunbar 1993, Tomasello et al. 1993). Here, we single out two conceptions that have proved seminal within this perspective on language and its evolution, namely the shared intentionality framework by Michael Tomasello and colleagues (e.g., Tomasello et al. 2005; Tomasello 2008), and Stephen Levinson's (2006) 'human interaction engine' hypothesis. Although Tomasello and Levinson do not expressly formulate a theory of language—the former concentrates on the prerequisites of language; the other, on the problem of human-specific interaction—they definitely subscribe to the vision of language as a form of social cognition and social action, providing the empirical evidence and theoretical scaffolding for this vision.

*4.3.1. Tomasello's **S**hared **I**ntentionality and Levinson's Interaction Engine*

Tomasello's shared intentionality framework (e.g., Tomasello et al. 2005) certainly ranks among the most important approaches in language evolution, even though one might wonder whether the framework actually seeks to explain the emergence of language or rather the evolution of the cognitive prerequisites for language. The answer to this question depends, again, on how exactly we define *language*. It therefore makes sense to first take a look at Tomasello's conceptualisation of language. While he does not provide a formal definition of *language*, he makes it clear that he sees language as a form of social action (Tomasello 2008: 342–345): "What is language if not a set of coordination devices for directing the attention of others?" (Tomasello et al. 2005: 690). On this view, language can be described as a way not only to coordinate attention, but also to construe objects and events from a particular perspective. But although language itself might aid in the cognitive development of perspective-taking (e.g., Lohmann & Tomasello 2003), the cognitive and interactional machinery it is built on must have emerged prior to language. As Tomasello puts it:

> If we want to understand human communication, [...] we cannot begin with language. Rather, we must begin with unconventionalized, uncoded communication, and other forms of mental attunement, as foundational. (Tomasello 2008: 59)

Tomasello (2008) sees language as a human-specific form of coded communication that uses conventionalized ('codified') signs, in contrast to uncoded communication making use of spontaneous, ad-hoc signs (cf. Arbib's pantomime). But importantly, there is no clear dividing line between these two modes of communication, which is why they cannot be discussed in isolation. In this regard, Tomasello's conception of language is similar to the ones espoused by the proponents of the multimodal view. The main difference between them lies in that while Tomasello emphasises the social grounding of language, Kendon, McNeill, and Zlatev tend to focus on the cognitive-interactional dynamics of linguistic communication and assess language and its evolution from this vantage point.

Tomasello and his collaborators see the cognitive infrastructure supporting shared intentionality as being central to human cultural cognition. Shared intentionality can be described as the motivation and ability to engage with others in collaborative activities with joint goals, plans and intentions and to share attention, experiences and other psychological states with others (cf. Tomasello et al. 2005, Tomasello & Carpenter 2007, Tomasello 2008). The shared intentionality infrastructure is hypothesised to be the foundation of uniquely human sociality and cumulative culture. This 'we-perspective' (Tuomela 2007) is what enables humans to have not only an understanding of shared goals in the way that other animals do not but represents the foundation of distinctive human cultural artefacts. These include institutional realities such as money, as well as conventions and norms more generally. They are seen as crucial for the emergence and acquisition of language, understood as a cultural artefact that is both conventional and subject to cumulative cultural evolution as evidenced in language change (Tomasello 1999, 2008, 2019; cf. the view of language as a cumulative technology, Dor 2015). In this way, the shared intentionality infrastructure is seen as the *sine qua non* for the evolution, acquisition, and use of language. Tomasello et al. (2005) acknowledge that language "must play a central role in all discussions of the evolution of human cognition." However, as they stress,

> […] saying that only humans have language is like saying that only humans build skyscrapers, when the fact is that only humans (among primates) build freestanding shelters at all. Language is not basic; it is derived. (Tomasello et al. 2005: 690)

Accordingly, language is seen as part and parcel of other human-specific social and cognitive skills and motivations that are already evident in the nonverbal communicative and cooperative behaviour of prelinguistic infants: the capacities that lead infants to communicate informatively and declaratively.

With regard to the evolutionary trajectory of language emergence, Tomasello agrees with proponents of multimodal approaches (see Section 4.1, but also Section 4.3.2 below) that gesture played a pivotal role in language evolution.

Specifically, he argues that the socio-cognitive infrastructure of shared intentionality was directly responsible for two original forms of human communication – pointing and pantomiming. Tomasello (2008) proposes declarative pointing, and especially informative-declarative pointing (i.e. pointing performed with the intention of providing the recipient with new information) to be the first step distinguishing human ancestors from the generalized baseline of the last common ancestor (LCA) that humans shared with chimpanzees. Pantomiming, the other rudiment of the original human communication system, is understood as an iconic and action-based representation of an event (Tomasello 2008). In this respect, Tomasello's account ties in with those by Zlatev and Arbib (cf. Sections 4.1. and 4.4) but differs especially from the former in that Tomasello does not assume that pantomime must have been multimodal.

A similar idea, partly drawing on Tomasello's account, can be found in Stephen Levinson's 'interaction engine' hypothesis (e.g., Levinson 2006), according to which what evolved in our ancestors was a socio-cognitive adaptation allowing "joint attention, common ground, collaboration and the reasoning about communicative intent" (Levinson & Holler 2014: 369). Levinson argues for a model of language evolution in which elements characteristic of modern language incrementally grew upon each other. Levinson designates a package of these elements as the Human Interaction Engine (Levinson 2006). Its key properties include intention-attribution, i.e. responses are to intentions, not to behaviours; cooperation, understood along the Gricean lines; turn-taking; predetermined sequential structures, such as adjacency pairs (cf. Pomerantz 1984); and multimodality, whereby

> […] face-to-face interaction is characterized by multimodal signal streams—visual, auditory, and haptic at the receiving end, and kinesic, vocal, and motor at the producing end. (Levinson 2006: 46)

The Interaction Engine represents "a human interactional specialization", which is universal across cultures, evolved prior to language and played a key role in its emergence (Levinson 2006: 42).

In contrast to the iterated learning framework and the research programme of Tomasello and colleagues, but in line with the multimodal approaches discussed in Section 4.1, Levinson explicitly stresses that human language is multimodal. For Levinson, language is one part of "human multi-modal communication" (Levinson & Holler 2014), which constitutes one integrated multimodal communication system, not in the sense of a special module or a 'language organ', but rather an assemblage of interrelated socio-cognitive abilities and communicative behaviours (Levinson 2006: 54 and passim). The assorted nature of human communication is reflected in its evolutionary history—"human communication is evolutionarily stratified, composed of layers of abilities of different types and different antiquity" (Levinson & Holler 2014). The bottom layer is constituted by what he refers to as the 'ethological elements', such as mutual gaze or turn-taking, including vocal turn-taking, which can be found in many primate clades (Levinson & Holler 2014). Levinson argues that cooperation may have been bootstrapped by these ethological elements in the context of novel ecological pressures (e.g., favouring bigger groups), which in turn laid the foundation for the

inferential background of human communication, exemplified by the presumption of cooperation (Grice 1975) and the presumption of relevance (Sperber & Wilson 1986).

Levinson and Holler forcefully argue for the continuity of ape and human gesture (in contrast, e.g., to Kendon, see Section 4.1.). Levinson assumes that ape gesture formed the platform for the development of indexical gestures, and most importantly declarative pointing. Here, Levinson largely accepts Tomasello's (2008) argument of how the evolution of prosociality in the hominin line, encapsulated by "joint attention, common ground, collaboration and the reasoning about communicative intent" (Levinson & Holler 2014), enabled the appearance of gestural indexes. In the next stage, iconic gesture emerged, accompanied by simple referential vocalisations, which gradually assumed the dominant role in the transfer of meaning (Levinson & Holler, 2014).

*4.3.2. Language as **S**ocial **I**nteraction: Taking **S**tock*

Tomasello and Levinson see language through the lens of Theory of Mind and social intelligence. Similar to the perspective of complex adaptive systems (Section 4.2. above), they conceive of language evolution as a multifactorial and emergent process but emphasise the role of socio-cognitive preconditions installed in humans via biological evolution. Tomasello (2003) suggests that the emergence of joint attention and joint action paved the way for the emergence of symbols, while grammaticalization led to the development of complex grammar. He points out that "different aspects of language—for example, symbols and grammar—may have involved different processes at different evolutionary times." (Tomasello 2003: 109). Levinson stresses the canalization of language through multifactorial constraints, with some attractors being "cognitive, some functional (communicational), some cultural-historical in nature" (Evans & Levinson 2009: 446).

Both Tomasello and Levinson see language as a layered 'mosaic' of different features, to use Hurford's (2003) metaphor (cf. Boeckx 2012). On this view, language (evolution) is neither strictly biological nor cultural, but instead characterized by an interplay of both evolutionary and cultural-historical processes (Tomasello et al. 2005, Tomasello 2008). They also converge on two other general points. First, language is first and foremost a communicative device—this tenet leads both Tomasello and Levinson to the appreciation of non-linguistic forms of communication, such as gesture, which forms a bridge between their positions and multimodal hypotheses (see Section 4.1). In particular, they both agree on the key role of iconic gesturing on the early, bootstrapping stages of language emergence.

Second, communication is rooted in social action, which itself is ramified by general-purpose cognitive mechanisms. For instance, Tomasello insists that language is not an object in any meaningful sense of that word, but rather one of the forms of social action. Tomasello does make a distinction between conventional, or 'coded', communication on the one hand and unconventionalised, uncoded communication on the other. However, he also reminds us that much of linguistic communication that makes use of conventionalised codes relies on uncoded aspects of meaning—as a case in point, consider pragmatic phenomena such as deixis and anaphora resolution (see Tomasello 2008: 57–59). Similarly, Levinson

resists the idea of language as abstracted from a rich interactional context constituted of social, cognitive and communicative factors.

4.4. Language in the Language-Ready Brain

As a fourth and final group of approaches to language evolution, we discuss those highlighting the neural implementation of the human ability to acquire and use language. We begin with Michael Arbib's notion of the *language-ready brain*, which we see as a promising candidate for integrating many lines of interdisciplinary evidence characteristic of more recent language evolution research. As we will show, the concept of language-readiness has been adopted well beyond the specific framework suggested by Arbib. As an example, we mention one specific theory of how the brain became language-ready, namely Offline Brain Systems proposed by Bouchard (2013). We then move on to a research avenue that adopts this notion but complements this perspective with a relatively greater reliance on genetic evidence.

4.4.1. Michael Arbib: The Language-Ready Brain

Michael Arbib's account of language origins holds considerable significance for current research on language evolution, not by being any less controversial than its alternatives, but through its remarkable theoretical completeness and the wide range of interdisciplinary data on which it is based. It began as the Mirror System Hypothesis (MSH; Arbib 2005, 2012, 2016) and now continues as Cognitive Neuroprimatology (CNP; Arbib 2018). In his work, Arbib and his collaborators (see especially Arbib et al. 2018) rely on a broad range of interdisciplinary data, which particularly prominently includes comparative data from extant primates (unlike in many other language evolution accounts, not limited to great apes but extending to macaques and other monkey species) as well as results of research on human visual-bodily communication, including both co-speech gesture and sign languages.

In line with this breadth of the evidential basis of his account, Arbib has an encompassing view of language as an explanatory target in language evolution research, making it compatible with other approaches discussed here: for example, it underscores the immanent multimodality of language (cf. Section 4.1.), the importance of both its formal-structural and social-interactional dimension (cf. Sections 4.1 and 4.3), as well as the division of labour between biological evolution in establishing the cognitive infrastructure for (proto-)language (see Sections 4.3 and 4.4.2 below) and cultural evolution in accomplishing the subsequent transition from protolanguage to full human languages (cf. Sections 4.2 and 4.3). Still, like other approaches, Arbib's MSH-CNP also has its specific focus, which in this case is on the cerebral implementation of language, as is evident in the name of the hypothesis as well as the title of his book-length manifesto, *How the brain got language* (Arbib 2012). Consistent with this focus is Arbib's notion of the 'language-ready brain'. This term is particularly useful in organising the discussion, since it does not inherently prioritise biological or cultural-evolutionary processes

but manages to capture human uniqueness (since "only the human brain is language ready", Arbib 2012: ix).

MSH-CNP espouses an "Evo-Devo-Socio"-perspective on language evolution:

> What evolved (*Evo*) was a language-*ready* brain—not a brain with an innate mechanism encoding a universal grammar (Arbib 2007) but rather one enabling a child to acquire language (*Devo*), but only if raised in a milieu in which language is already present, something which, it is claimed, required tens of millennia of cultural evolution after the emergence of *Homo sapiens* (*Socio*). (Arbib 2018: 7).

It highlights the building blocks that are not themselves (traditionally seen as) linguistic but are necessary for language, most importantly the cognitive infrastructure supporting the sharing of meaning: *Parity* and *imitation*. Parity is the ability to 'translate' between production and comprehension, whereby the same signal counts for more or less the same meaning to both the producer and the receiver of this signal (at least on a basic level, which neglects the complexities of pragmatic inference, e.g., Scott-Phillips 2015). Whereas a great majority of accounts of language evolution simply take this fundamental requirement for communication for granted, MSH-CNP offers a detailed account of the neuronal implementation of parity, based on the mirror neuron system (e.g., Arbib 2005, 2012). While imitation is an important component of other language evolution accounts, including Tomasello (2008; see Section 4.3 above), Arbib (2012) stands out by describing a succession of steps in its development: from a mirror-neuron system for grasping and manual praxic actions, through simple imitation, then complex action recognition and complex imitation (CAR&IM), ultimately leading to pantomime—initially of grasping and manual praxic actions, then of actions outside of own repertoire.

Pantomime is a characteristic feature of MSH-CNP. It is pantomime that is responsible for bringing about perhaps the most important qualitative breakthrough, i.e. that of *open-endedness* in communication: "freedom to create novel associations" (2012: 261). Arbib (2012: 219) observes that pantomime has "the ability to create an open-ended set of complex messages exploiting the primates' open-ended manual dexterity". This potential to flexibly introduce novel signals for novel messages underwrites two other gains in expressive power otherwise typical of language, that is domain-generality and displacement. Pantomime is domain-general in that it can be used to communicate about many semantic domains (rather than being restricted to, e.g., only predator evasion or food), and it can also express meanings displaced in time and space (not concerning the immediate here and now).

As mentioned above, the scope of the evidential basis and the resulting breadth of the language-ready brain approach gives rise to numerous convergences with the other approaches discussed in this paper. For example, Arbib argues that the neural mechanisms supporting language perception and production were first involved in non-communicative actions such as tool production, much in line with Kendon's proposal (see Section 4.1 above). According to Arbib's

hypothesis, the complex imitation of hand movements predates language learning and use and evolved as a form of social sharing of practical skills. One interesting spin-off of Arbib's hypothesis is the claim that language universals are mostly due to cultural factors, with language structure mostly evolving via a cultural mechanism. This resonates with much work in the CAS paradigm (see Section 4.2), and especially by Kirby and others, who come to similar conclusions from a different approach, i.e. computational simulations (e.g., Kirby et al. 2007) later complemented by laboratory experiments (e.g., Kirby et al. 2008).

The concept of language-readiness has also been adopted by other researchers who do not necessarily share Arbib's MSH-CNP account. For instance, Bouchard (2013, 2015), who sees language as a system of signs, including combinatorial signs that underlie syntax (a view that shares many similarities with the Construction Grammar view discussed above), argues that language-readiness is a consequence of the emergence of 'offline brain systems'. These are systems that can be triggered not only by external but also by brain-internal events. These offline brain systems enable a more abstract representational level, which allows for concepts and percepts (or rather: representations thereof) to be linked. He explains the emergence of these systems by an increase in synaptic interactions triggered by a number of interacting developments, for example, the larger brain that entails an increased potential for synaptic interactions, the more globular shape of the brain that affords more cross-modular interactions, and alleles that improve synaptic repair, thus dramatically increasing synaptic interactions (see Bouchard 2015). In addition, he also stresses the importance of bio-cultural coevolution: "the long dependency during infancy feeds more cultural material into these additional brain capacities" (Bouchard 2015). On this view, language, and even language-readiness, can be seen as an exaptation, as a 'side effect', as it were, of other, more general biological (and cultural) developments.

This is only one example that shows that the concept of brain-readiness has become central not only to accounts of language evolution, but also for discussing the nature of language. But Arbib's more specific proposal has remained highly influential as well and has been adopted and further developed in subsequent work. We will now discuss these developments in more detail.

*4.4.2. Antonio Benítez-Burraco and Cedric Boeckx: **T**he **L**anguage-**R**eady **B**rain **R**evisited*

Most of Arbib's neurobiological discussion is focused on brain areas related to language in humans and to visual and auditory perception in both humans and primates. At the same time, as far as language in prehistory is concerned, he mostly focuses on stone technologies in different extinct hominins. However, very recently a more detailed view has been emerging of how the hominin brain was genetically modified in the evolutionary history of our species to support processes involved in language. In particular, a series of related papers by Benítez-Burraco and Boeckx (Boeckx & Benítez-Burraco, 2014a, 2014b; Benítez-Burraco & Boeckx, 2015) has outlined a refined approach to Arbib's view of the language-ready brain, as well as Bouchard's approach based on the notion that our more globular brain resulted in enhanced cross-modal thinking. These authors also adopt a multimodal approach, but mostly relate it to Poeppel's claims of the

multifunctionality of the neural mechanisms involved in language processing (see Poeppel & Embick 2005; Poeppel 2012). Likewise, following Bouchard's 'neurogenetic factors' (e.g., Bouchard 2013: Chap. 4), they are particularly interested in the genetic factors accounting for the changes resulting in our globularity. However, they build on recent paleoneurological and paleogenetic research about changes in the human genome and the human brain (and skull), principally after the split of *Homo sapiens* from the closely related clades of Neanderthals and Denisovans.

The most important reason for this new approach was avoiding what Benítez-Burraco and colleagues saw as an overreliance on speculation in the field of evolutionary linguistics, in particular relying on highly elusive and contentious proxies for language such as 'symbolic behaviour'. Instead, they proposed to focus on the most distinctive and less controversial biological differences distinguishing modern humans from Neanderthals and other extinct hominins. Perhaps the most prominent of such differences is the globular aspect of the human endocranial morphology (Bruner et al. 2003; Neubauer et al. 2010; Gunz et al. 2010, 2012). In their papers, Benítez-Burraco and Boeckx argued that this globularisation of the human brain resulted in a rewiring that improved the connections between subcortical (particularly, the thalamus) and cortical structures, habilitating the neuronal workspace needed for transcending the signature limits of core knowledge systems and ultimately allowing to combine and unify conceptual units that belong to distinct core systems. They further argued that this ability can be equated with the core combinatorial operation in natural language (which may be called Merge by linguists of Chomskyan persuasion), which is at the core of our language-ready brain.

Overall, this is a bridging hypothesis linking considerations of syntax and neuroscience. One reason for this is that this freely combining merging ability is argued to be constrained via its interfacing with other cognitive systems and with the devices involved in speech/gesture production. In addition, compatible with the multimodal approaches to language (see Section 4.1.), this regulation is hypothesised to result from basic neurobiological mechanisms, specifically, from the embedding of high frequency oscillations (e.g., gamma) inside oscillations operating at slower frequencies (e.g., alpha). This embedding ultimately enables the synchronization of distant cortical areas where the diverse core knowledge systems are located, with some subcortical structures, particularly the thalamus, acting, as noted, as a relay centre or switching station connecting the cortical areas. The circuits bidirectionally connecting the thalamus and cortex are at the heart of the language-ready brain and share features of the networks responsible for mind-wandering and inner speech (Gruberger et al. 2011), as well as the top-down attentional regulation network (Miller & Buschman 2013). Likewise, some studies (e.g., Hecht et al. 2013) have related the changes resulting in our language-readiness (particularly, the increase in the ratio between fronto-parietal vs fronto-temporal connectivity from monkeys to apes to humans) to the evolutionary shift from emulation (i.e. a way of copying actions that focuses on the goal rather than the specific movements) to imitation (i.e. a way of copying actions that focuses on the specific movements rather than their ultimate goal).

On Benítez-Burraco and Boeckx's view, the emergence of a language-ready brain does not entail any drastic changes to the generalised primate brain, nor the evolution of entirely new neural devices. Rather, interareal cerebral communication via the synchronizing of spatially distributed oscillations is a generic strategy of the brain, specific neither to humans nor to language. Likewise, most if not all brain areas supporting language are most likely present in other species. Accordingly, the evolution of the language-ready brain essentially involved a change in the dynamic connectivity of the brain resulting from a new anatomical context. Notably, Benítez-Burraco and Boeckx use the available information on ancient genomes to identify some of the genetic changes that may account for the observed differences between hominin species regarding globularisation and cognitive abilities. The candidate genes of interest include *RUNX2* (a master transcription factor during vertebrate development) and several of its effectors; components of two gene networks implicated in vocal learning, clustered around the famous 'language gene' *FOXP2*, and the ROBO and SLITs effectors; and finally, a set of genes clustered around *AUTS2*, strongly linked to autism (which, incidentally, reinforces the intriguing parallelisms between the autistic mind and the hypothesised Neanderthal mind). Similarly to its neurobiological substrate, also the genetic underpinnings of the language-ready brain are mostly shared with other primates, although some human-specific changes can be also identified, seemingly accounting for the changes in neuronal networks described above.

*4.4.3. The **L**anguage-**R**eady **B**rain: **T**aking **S**tock*

On the language-ready brain view, the presence of compositional, open-ended and domain-general semantics is the most important criterial feature of language. This is thought to rely not on an enhanced, language-specific computational ability, but on an unbounded basic combinatorial ability capable of transcending the limits of core conceptual systems. Other components of language, particularly, some forms of phonology and pragmatics, are assumed to predate this human-specific innovation. Accordingly, speech or interaction through (proto)language with communicative or socializing purposes are thought to have been present in other hominin species, particularly Neanderthals. Overall, this is a genuinely non-modular construal of language. Neural devices involved in language processing are hypothesised to perform basic computations that are recruited for language, but also for other cognitive processes. Accordingly, the impairment of any of these neural components of language—either developmentally, resulting from gene mutations, or in the mature state, resulting from brain damage—is expected to give rise to mixed symptoms and diverse pathological, comorbid conditions. As a consequence, too, language is construed as domain-specific only at the term of growth, with the 'language module' resulting from the interaction of diverse ontogenetic and functional brain modules through development.

This view is in line with neuroconstructivist approaches to human cognition (cf. Karmiloff-Smith 2009), where language is understood primarily as a biological capacity, resulting from human-specific gene mutations affecting genes involved in brain development and wiring and language evolution, as the result of minor changes in brain wiring—although it also acknowledges a significant evolution-

ary continuity of language with the communication and cognition of other animals, in particular of non-human apes. This also results in a relative focus on the individual implementation of language, with cognitive changes resulting from neuroanatomical changes, themselves principally resulting from genetic mutations. Ultimately, supraindividual and cultural aspects of language are subordinated to individual and biological aspects, although some feedback effect of culture on cognition is expected, but not with a decisive triggering effect on language evolution. One interesting and distinctive consequence is a conception of language primarily as a tool for thinking (rather than for communicating), since our language-readiness would have initially emerged as a new, improved mechanism of conceptualising. Regarding the question of whether language is species-specific, this is mostly seen as an empirical question in these approaches, with the expectation of relatively few uniquely human components 'on top of' a majority of components shared with other species. Accordingly, this is more a gradable and gradual view of language (evolution) than a categorical one.

5. Discussion

In this paper we have argued against the intuitively appealing and occasionally expressly formulated (especially Botha 2000) dictum that language evolution as a field of research needs an overtly formulated, analytical, top-down definition of language. We have proposed that such a definition is most likely impossible, and further, even if possible, it might not be conducive to any tangible gains. As a particularly forceful illustration of this latter point, we discussed the term FLN, which was conceived in response to the perceived problem of a lack of a technical definition of (the faculty of) language, but which—as we have documented—was defined in two mutually exclusive ways, whose combining results in a tautology. In the subsequent literature the two mutually exclusive definitions widely function interchangeably, without scholars as much as noticing the conflict – a striking demonstration that the bulk of research practice in the field depends on other factors (arguably including an author's larger theoretical orientation) but not on the exact letter of the proposed top-down definitions.[7]

Notably, the claim we advance here does not imply denying the importance of clear definitions in scientific discourse generally, and in language evolution research specifically. This is no contradiction: We distinguish between definitions on two different and clearly separable levels. One is the level of more specific

[7] Worth mentioning here is that while FLN/FLB distinction aims at a precise delineation of (the biological underpinnings of) language, other frameworks such as those that can be grouped under the umbrella of "Complex Adaptive Systems" approaches more or less explicitly acknowledge that language cannot be clearly delineated from other phenomena. In a way, these two views of language can be considered two extreme poles on the continuum from an extremely broad to a maximally narrow conceptualization. The different definitions along this continuum show that the way we think about language as an object of study partly depends on theoretical presuppositions and partly on the epistemological interest of each approach: While Chomsky, Hauser, and Fitch, for example, aim at carving out the biological, species-specific prerequisites for language and therefore narrow down the scope of language as a technical term considerably, proponents of CAS accounts take a macro-perspective on biological and cultural evolution and view language in the broader context of social-interactional phenomena.

technical terms that function as building blocks of theories and especially of hypotheses, which require unambiguous formulations so as to meet the fundamental standards of non-triviality and falsifiability. This level is thus essential for science to make progress by conclusively resolving arguments with recourse to empirical data rather than getting stuck on conceptual differences.[8] The other level, however, is the global level of macroscopic notions, which cannot (without further specification) function as building blocks of specific theories or hypotheses but have a different role, related instead to integrative and classificatory goals.

So, for example, we largely concur with Behme that

> [w]hile it may be neither feasible nor beneficial that all language evolution researchers adopt the same definition of 'language' it would be desirable for them to explicitly state which definition they adopt;
>
> (2016: 8)

and with Fitch (2010: 24) that "unspecified use of […] the word 'language' […] is probably best avoided". A small but essential caveat is that any theory-specific use of *language* will inevitably remain meronymous, in the sense of always relating only to part of the complex phenomenon. Therefore, in our view it is more productive to push definitions one level down: Leave *language* as an unanalysable prime and provide rigorous definitions of particular *components* or *aspects* of language as they function in specific theories under consideration. One very important advantage is that such a strategy prevents attempts to monopolise the word *language* by a particular theory that would claim unique privileged access to a 'correct' understanding of *language*, something that Chomskyan approaches have been criticised for.

Exactly such was the nature of the FLN/FLB distinction (again, two different distinctions, as we show in Section 3), categorical about the nature of language and expressly formulated to guide language evolution research as an understanding of the language faculty privileged over other theories. Interestingly, however, much fuzzier notions of language seem to have better served the actual language evolution research. Due to the breadth of research interests in language evolution in the last decade, and the intense interdisciplinarity that cuts across many disparate areas of investigation—from computational modelling, to primate communication, to sign linguistics—language evolution thrives on fuzzy definitions of language and finds categorical, top-down approaches too constraining. This idea is reflected in Section 4, which surveys influential lines of research in language evolution, mostly focusing on the most recent trends. In recent years, the bulk of research in this field revolves around the problems of multimodality, the dynamics of cultural transmission, language as a form of social interaction or biological language-readiness. What emerges from this survey is indeed a breadth of the range of these perspectives that precludes their fitting together under any single definition of language.

---

8 We are grateful to two anonymous reviewers for comments that led us to stress this important point.

This is far from surprising if we look to analogies in other notions in science that, like language, are macroscopic, unobservable and unoperationalisable. One example comes from biologists, who tend to avoid top-down approaches to life (and aprioristic definitions of *life*) and focus instead on the study of the building blocks of living beings, regardless of whether they can be found in other domains (e.g., water) or not (the DNA). Only a successful characterization of these building blocks can lead to achieving a comprehensive view of the nature of life (as in systems biology) and its evolution (as in, e.g., evolutionary developmental biology). Similarly, analyses of the use of terms such as *heat* in physics (Lewis & Linn 1996) or *gesture* in primatology (Bourjade et al. 2020) provide arguments for a beneficial and productive role of conceptual diversity, at least when certain conditions are met such as consistent use of a term within a particular approach.

One particularly interesting motivation for why a lack of a single top-down definition of language in language evolution is not as consequential now as it was 20 years ago may be a methodological change in the profile of this field, from theoretical to empirical research (see especially Dediu & de Boer 2016, Fitch 2017, Żywiczyński 2018, Nölle et al. 2020). In 2017, Wacewicz & Żywiczyński wrote:

> Language evolution researchers no longer stop at being consumers of empirical data, but rather aim at being providers as well, acquiring data by experimentation, observation, or simulation (and a steadily increasing proportion of these results then feed back into more general discussions on the nature of language […]). The maturation of language evolution research has been marked by a steady growth in the proportion of empirical ("new data") research relative to theoretical (synthetic) argumentation […]. In the volume that grew out of the first EVOLANG conference in 1996 (Hurford et al. 1998), all 24 contributions have a decidedly theoretical (synthesising) character, whereas the proceedings of the most recent conference (Roberts et al. 2016) are dominated by empirical research: 123 contributions, as opposed to 25 theoretical. Wacewicz & Żywiczyński (2017: 3)

What follows is that recently, very few publications present comprehensive scenarios of language evolution, and conversely, a vast proportion of studies are more fine-grained, addressing much more specific and narrower Kuhnian "puzzles" such as the efficiency of gestural vs multimodal signals in conveying emotional meanings (Zlatev et al. 2017) or the effect of processed food on the dental configuration and in turn on the production of fricatives (Blasi et al. 2019). A natural consequence is that such specific and bottom-up studies do not directly aim at explaining language evolution *sensu largo*, and so do not need to work with a definition of language *sensu largo*.

An epitome of both the empirical and bottom-up approach is the Causal Hypotheses In Evolutionary Linguistics Database (CHIELD, pronounced 'shield'; Roberts et al. 2020). CHIELD contains crowd-sourced entries for over 400 publications, with over 3,400 causal links between more than 1,700 variables and aims not only at cataloguing hypotheses about language evolution but also making data on them interoperable. As it is unlikely that all 32 authors (much less all 41

contributors to the database) would converge on a single explicit definition of language, CHIELD is possible not *despite* but *because* it deliberately avoids defining language (cf. "A classic example of this is the word 'language' itself, which can be interpreted as anything relating to human communication or only a specific syntactic ability"; Roberts et al. 2020: 3).

Consequently, one way of describing language evolution could be to 'bracket' the notion of *language* and rely solely on content-independent, institutional and scientometric criteria such as conferences, journals, laboratories and citation patterns (cf. Bergmann & Dale 2016). This would delineate a collection of bottom-up approaches and researchers that jointly form a 'community of practice' or a *denkkolektiv* (Fleck 1979). This is an interesting approach with some genuine explanatory power; for example, this strategy would address Haspelmath's (2016) question of why *Journal of Language Evolution* publishes research on language change of apparently non-evolutionary character.

Nevertheless, such an approach would seem deeply unsatisfying to the researchers in the field of language evolution, who have a strong sense of unifying research substance, and in particular the unifying aim of explaining the origin of language. This substance is primary to the content-independent factors, in that it provides identity to the field and gives rise to—as opposed to being secondary and merely resultant from—the patterns and networks of personal and institutional connections. It is the basic, common, intuitive understanding of *language*, and basic human curiosity about how it began, that sets the explanatory goal for the field of language evolution as a whole, and thus shapes its research practices and the resulting *denkkollektiv*—rather than vice versa.

This is the other point that follows from our review: although the recently most influential approaches to language evolution are indeed too diverse to be brought together under a common definition of language, they also do overlap to a large extent in terms of key definitional dimensions. As discussed in section 4, these dimensions are in particular the criterial components of language, its modalities, domain-specificity, biological versus cultural profile, (supra)individual character, gradability, species-specificity and primary function. This complex pattern of numerous similarities and sporadic but significant differences is characteristic of a family-resemblance category (Wittgenstein 1953). Most importantly from the point of view of research practice, this family resemblance pattern underwrites fruitful communication between these approaches, leading to cross-fertilisation and opening new research vistas: for example, there is a growing number of studies in the Iterated Learning paradigm that look at different communicative modalities (e.g., Motamedi et al. 2019); as another example, the self-domestication theory, proposed within the Complex Adaptive Systems camp, is now most actively developed by the proponents of the language ready brain (see especially Benítez-Burraco & Progovac 2020). Thus, based on a tacit and fundamentally ineffable notion of language, all these approaches jointly contribute to the development of the field of language evolution.

Acknowledgements

This research was supported by the Polish National Science Centre under grant agreement UMO-2019/34/E/HS2/00248. This work was also supported from the Excellence Initiative—Research University at Nicolaus Copernicus University in Toruń. We are grateful to Marta Sibierska for her valuable comments and to Aleksandra Poniewierska and Aleksandra Szczepańska for their assistance in editing the manuscript.

References

Aboitiz, Francisco. 2012. Gestures, vocalizations, and memory in language origins. *Frontiers in Evolutionary Neuroscience* 4(2).

Amici, Federica, Alex Sánchez-Amaro, Carla Sebastián-Enesco, Trix Cacchione, Matthias Allritz, Juan Salazar-Bonet, & Federico Rossano. 2019. The word order of languages predicts native speakers' working memory. *Scientific Reports* 9(1), 1124.

Arbib, Michael A. 2005. From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences* 28(2), 105–167.

Arbib, Michael A. 2012. *How the Brain Got Language: The Mirror System Hypothesis.* Oxford: Oxford University Press.

Arbib, Michael A. 2016. Towards a computational comparative neuroprimatology: framing the language-ready brain. *Physics of Life Reviews* 16, 1–54.

Arbib, Michael A. 2018. Computational challenges of evolving the language-ready brain: 2. Building towards neurolinguistics. *Interaction Studies* 19(1–2), 22–37.

Arbib, Michael A, Francisco Aboitiz, Judith M. Burkart, Michale C. Corballis, Gino Coudé, Erin Hecht, Katja Liebal, Masako Myowa-Yamakoshi, James Pustejovsky, Shelby S. Putt, Federico Rossano, Anne E. Russon, P. Thomas Schoenemann, Uwe Seifert, Katerina Semendeferi, Chris Sinha, Dietrich Stout, Virginia Volterra, Sławomir Wacewicz, & Benjamin Wilson. 2018. The comparative neuroprimatology 2018 (CNP-2018) road map for research on how the brain got language. *Interaction Studies* 19, 371–388.

Armstrong, David F. & Sherman E. Wilcox. 2007. *The Gestural Origin of Language*. Oxford: Oxford University Press.

Balari, Sergio & Guillermo Lorenzo. 2012. *Computational Phenotypes: Towards an Evolutionary Developmental Biolinguistics*. Oxford: Oxford University Press.

Balari, Sergio & Guillermo Lorenzo. 2016. Evo-devo of language and cognition. In Laura Nuno de la Rosa & Gerd Müller (eds.), *Evolutionary Developmental Biology*. Cham: Springer.

Barceló-Coblijn, Lluís. 2012. Evolutionary scenarios for the emergence of recursion. *Theoria et Historia Scientiarum* 9, 171–199.

Beckner, Clay, Richard Blythe, Joan Bybee, Morten H. Christiansen, William Croft, Nick C. Ellis, John Holland, Jinyun Ke, Diane Larsen-Freeman, & Tom Schoenemann. 2009. Language is a complex adaptive system: Position paper. *Language Learning* 59, 1–26.

Behme, Christina. 2016. Evolution of what? In Sean G. Roberts, Christine Cuskley, Luke McCrohon, Lluis Barceló-Coblijn, Olga Fehér, & Tessa Verhoef (eds.), *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG11)*. http://evolang.org/neworleans/papers/140.html

Benítez-Burraco, Antonio 2020. Prehistoric languages and human self-domestication. *Language Dynamics and Change* 10(1), 27–58.

Benítez-Burraco, Antonio & Cedric Boeckx. 2014. Universal Grammar and biological variation: An evodevo agenda for comparative biolinguistics. *Biological Theory* 9(2), 122–134.

Benítez-Burraco, Antonio & Cedric Boeckx. 2015. Possible functional links among brain- and skull-related genes selected in modern humans. *Frontiers in Psychology* 6: 794.

Benítez-Burraco, Antonio & Ljiljana Progovac. 2020. A four-stage model for language evolution under the effects of human self-domestication. *Language & Communication* 73, 1–17.

Bergen, Benjamin K. 2012. *Louder than Words: The New Science of How the Mind Makes Meaning*. New York: Basic Books.

Bergmann, Till & Rick Dale. 2016. A scientometric analysis of evolang: Intersections and authorships. In Sean G. Roberts, Christine Cuskley, Luke McCrohon, Lluis Barceló-Coblijn, Olga Fehér, & Tessa Verhoef (eds.), *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG11)*. http://evolang.org/neworleans/papers/182.html

Berwick, Robert C. & Noam Chomsky. 2016. *Why Only Us: Language and Evolution*. Cambridge, MA: MIT Press.

Bickerton, Derek. 1990. *Language and Species*. Chicago, IL: University of Chicago Press.

Blasi, Damián E., Steven Moran, Scott R. Moisik, Paul Widmer, Dan Dediu, & Balthasar Bickel. 2019. Human sound systems are shaped by post-Neolithic changes in bite configuration. *Science* 363(6432).

Boeckx, Cedric. 2012. The I-language mosaic. In Cedric Boeckx, María del Carmen Horno Chéliz, & José Luis Mendívil Giró (eds.), *Language, from a Biological Point of View: Current Issues in Biolinguistics*, 23–51. Newcastle upon Tyne: Cambridge Scholars.

Boeckx, Cedric & Antonio Benítez-Burraco. 2014a. Globularity and language-readiness: Generating new predictions by expanding the set of genes of interest. *Frontiers in Psychology* 5(1324), doi: 10.3389/fpsyg.2014.01324.

Boeckx, Cedric & Antonio Benítez-Burraco. 2014b. The shape of the human language-ready brain. *Frontiers in Psychology* 5(282), doi: 10.3389/fpsyg.2014.00282.

Bolhuis, Johan J., Ian Tattersall, Noam Chomsky & Robert C. Berwick. 2014. How could language have evolved? *PLoS Biology* 12(8), doi: 10.1371/journal.pbio.1001934.

Botha, Rudolf P. 2000. Discussing the evolution of the assorted beasts called language. *Language & Communication* 20(2), 149–160.

Bouchard, Denis. 2013. *The nature and origin of language*. Oxford: Oxford University Press.

Bouchard, Denis. 2015. Brain readiness and the nature of language. *Frontiers in Psychology* 6. doi:10.3389/fpsyg.2015.01376

Bourjade, Marie, Hélène Cochet, Sandra Molesti & Michèle Guidetti. 2020. Is conceptual diversity an advantage for scientific inquiry? A case study on the concept of 'gesture' in comparative psychology. *Integrative Psychological and Behavioral Science*, doi: 10.1007/s12124-020-09516-5.

Bowling, Daniel L. 2017. The continuing legacy of nature versus nurture in biolinguistics. *Psychonomic Bulletin & Review* 24(1). 140–141.

Bruner, Emiliano, Giorgio Manzi & Juan L. Arsuaga. 2003. Encephalization and allometric trajectories in the genus Homo: evidence from the Neandertal and modern lineages. *Proceedings of the National Academy of Sciences* 100. 15335-15340.

Byrne, Richard & Andrew Whiten (eds.). 1988. *Machiavellian intelligence: Social expertise and the evolution of intellect in monkeys, apes and humans*. Oxford: Oxford University Press.

Cheney, Dorothy L. & Robert M. Seyfarth. 2007. *Baboon metaphysics: The evolution of a social mind*. Chicago: Chicago University Press.

Christiansen, Morten H. & Simon Kirby. 2003. Language evolution: Consensus and controversies. *Trends in Cognitive Sciences* 7(7). 300–307.

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

Corballis, Michael C. 2013. Gestural theory of the origins of language. In Claire Lefebvre, Bernard Comrie & Henri Cohen (eds.), *New perspectives on the origins of language*, 171–184. Oxford: Oxford University Press.

Cornish, Hannah. 2010. Investigating how cultural transmission leads to the appearance of design without a designer in human communication systems. *Interaction Studies* 11(1). 112–137.

Cowie, Fiona. 2008. Innateness and language. *The Stanford Encyclopedia of Philosophy*. https://plato.stanford.edu/archives/fall2017/entries/innateness-language/. (21 July, 2020.)

Croft, William. 2000. *Explaining Language Change: An Evolutionary Approach.* Harlow: Pearson Longman.

Croft, William. 2011. Language structure in its human context: New directions for the language sciences in the twenty-first century. In Patrick Colm Hogan (ed.), *The Cambridge encyclopedia of the language sciences*, 1–11. Cambridge: Cambridge University Press.

Darwin, Charles. 1871. *The descent of man: and selection in relation to sex.* London: J. Murray.

Deacon, Terrence W. 1997. *The Symbolic Species: The Co-Evolution of Language and the Brain.* New York & London: Norton.

de Boer, Bart. 2000. Self-organization in vowel systems. *Journal of Phonetics* 28. 441–465.

Dediu, Dan & Bart de Boer. 2016. Language evolution needs its own journal. *Journal of Language Evolution* 1. 1–6.

Divjak, Dagmar. 2019. *Frequency in language: memory, attention and learning*. Cambridge: Cambridge University Press.

Donald, Merlin. 1991. *Origins of the modern mind: Three stages in the evolution of culture and cognition.* Cambridge, MA: Harvard University Press.
Donald, Merlin. 2001. *A mind so rare: The evolution of human consciousness*. New York: Norton.
Dor, Daniel. 2015. *The Instruction of Imagination: language as a social communication technology.* Oxford: Oxford University Press.
Dor, Daniel, Chris Knight, Jerome Lewis (eds.). 2014. *The social origins of language*. Oxford: Oxford University Press.
Dunbar, Robin I. M. 1993. Coevolution of neocortical size, group size and language in humans. *Behavioral and Brain Sciences.* 16(04). 681–694.
Dunbar, Robin I. M. 1996. *Grooming, gossip, and the evolution of language*. Cambridge, MA: Harvard University Press.
Ejiri, Keiko & Nobuo Masataka. 2001. Co-occurrences of preverbal vocal behavior and motor action in early infancy. *Developmental Science* 4(1). 40–48.
Ekman, Paul & Wallace V. Friesen. 1969. The repertoire of nonverbal behavior: Categories, origins, usage, and coding. *Semiotica* 1(1). 49–98.
Elman, Jeffrey L. 1999. The emergence of language: A conspiracy theory. In Brian MacWhinney (ed.), *The Emergence of Language*. 1–28. Mahwah, New Jersey: Lawrence Erlbaum Associates.
Evans, Nicholas & Stephen C. Levinson. 2009. The myth of language universals: Language diversity and its importance for cognitive science. *Behavioral and Brain Sciences* 32(5). 429–448.
Fitch, Tecumseh W. 2010. *The evolution of language.* Cambridge: Cambridge University Press.
Fitch, Tecumseh W. 2017. Empirical approaches to the study of language evolution. *Psychonomic Bulletin & Review* 24(1). 3–33.
Fitch, Tecumseh W., Marc D. Hauser & Noam Chomsky. 2005. The evolution of the language faculty: clarifications and implications. *Cognition* 97(2). 179–210.
Fleck, Ludwik. 1979. *Genesis and Development of a Scientific Fact.* Chicago: The University of Chicago Press.
Garrod, Simon, Nicolas Fay, Shane Rogers, Bradley Walker & Nik Swoboda. 2010. Can iterated learning explain the emergence of graphical symbols? *Interaction Studies* 11(1). 33–50.
Green, Jennifer. 2014. *Drawn from the ground: Sound, sign and inscription in Central Australian sand stories.* Cambridge: Cambridge University Press.
Grice, Herbert Paul. 1975. Logic and conversation. In Peter Cole & Jerry L. Morgan (eds.), *Syntax and Semantics 3: Speech Acts*, 41–58. New York: Academic Press.
Gruberger, Michal, Eti Ben-Simon, Yechiel Levkovitz, Abraham Zangen & Talma Hendler. 2011. Towards a neuroscience of mind-wandering. *Frontiers in Human Neuroscience* 5(56), doi: 10.3389/fnhum.2011.00056.
Gunz, Philipp, Simon Neubauer, Bruno Maureille & Jean-Jacques Hublin. 2010. Brain development after birth differs between neanderthals and modern humans. *Current Biology* 20. R921-R922.
Gunz, Philipp, Simon Neubauer, Liubov V. Golovanova, Vladimir Doronichev, Bruno Maureille & Jean-Jacques Hublin. 2012. A uniquely modern human pattern of endocranial development. insights from a new cranial

reconstruction of the Neandertal newborn from Mezmaiskaya. *Journal of Human Evolution* 62. 300–313.
Haspelmath, Martin. 2016. The evolution (or diachrony) of "language evolution". http://dlc.hypotheses.org/894. (21 July, 2020.)
Haspelmath, Martin. 2020. Human linguisticality and the building blocks of languages. *Frontiers in Psychology* 10. 3056. doi: 10.3389/fpsyg.2019.03056.
Hauser, Marc, Noam Chomsky & Tecumseh Fitch. 2002. The faculty of language: What is it, who has it, and how did it evolve? *Science* 298. 1569–1579.
Hauser, Marc, David Barner & Tim O'Donnell. 2007. Evolutionary linguistics: A new look at an old landscape. *Language Learning and Development* 3(2). 101–132.
Hecht, Erin E., David A. Gutman, Todd M. Preuss, Mar M. Sánchez, Lisa A. Parr & James K. Rilling. 2013. Process versus product in social learning: comparative diffusion tensor imaging of neural systems for action execution-observation matching in macaques, chimpanzees, and humans. *Cortex* 23. 1014–1024.
Hilpert, Martin. 2013. *Constructional Change in English: Developments in Allomorphy, Word Formation, and Syntax*. Cambridge: Cambridge University Press.
Hewes, Gordon Winant. 1977. Language origin theories. In Duane M. Rumbaugh (ed.), *Language learning by a chimpanzee: The Lana project*, 3–53. New York: Academic Press.
Hrdy, Sarah Blaffer. 2009. *The woman that never evolved*. Cambridge, MA: Harvard University Press.
Hurford, James R. 1989. Biological evolution of the Saussurean sign as a component of the language acquisition device. *Lingua* 77(2). 187–222.
Hurford, James R. 2003. The language mosaic and its evolution. In Morten H. Christiansen & Simon Kirby (eds.), *Language Evolution* 3, 38–57. Oxford: Oxford University Press.
Hurford, James R. 2007. *The Origins of Meaning: Language in the Light of Evolution, Vol. 1*. Oxford: Oxford University Press.
Hurford, James R. 2012. *The Origins of Grammar: Language in the Light of Evolution, Vol. 2*. Oxford: Oxford University Press.
Hurford, James R., Michael Studdert-Kennedy & Chris Knight (eds.). 1998. *Approaches to the evolution of language: social and cognitive bases.* Cambridge: Cambridge University Press.
Jackendoff, Ray. 2002. *Foundations of Language: Brain, Meaning, Grammar, Evolution*. Oxford: Oxford University Press.
Jackendoff, Ray. 2010. Your theory of language evolution depends on your theory of language. in Richard K. Larson, Viviane Déprez & Hiroko Yamakido (eds), *The evolution of human language: Biolinguistic perspectives*, 63–72. Cambridge: Cambridge University Press.
Jackendoff, Ray & Steven Pinker. 2005. The nature of the language faculty and its implications for evolution of language (Reply to Fitch, Hauser, and Chomsky). *Cognition* 97(2). 211–225.
Johansson, Sverker. 2005. *Origins of language: Constraints on hypotheses*. John Benjamins Publishing.

Karmiloff-Smith, Annette. 2009. Nativism versus neuroconstructivism: Rethinking the study of developmental disorders. *Developmental Neurocognition* 45. 56–63.

Kendon, Adam. 1980. Gesticulation and speech: Two aspects of the process of utterance. In Mary Ritchie Key (ed.), *The Relationship of verbal and nonverbal communication*, 207–227. The Hague & New York: Mouton.

Kendon, Adam. 1990. *Conducting interaction: Patterns of behavior in focused encounters*. Cambridge: Cambridge University Press.

Kendon, Adam. 2004. *Gesture: Visible action as utterance.* Cambridge: Cambridge University Press.

Kendon, Adam. 2008. Signs for language origins? *The Public Journal of Semiotics* 2(2). 2–29.

Kendon, Adam. 2011. Some modern considerations for thinking about language evolution: A discussion of the evolution of language by Tecumseh Fitch. *The Public Journal of Semiotics* 3(1). 79–108.

Kendon, Adam. 2014a. Semiotic diversity in utterance production and the concept of 'language.' *Philosophical Transactions of the Royal Society B: Biological Sciences* 369(1651). doi:10.1098/rstb.2013.0293

Kendon, Adam. 2014b. The "poly-modalic" nature of utterances and its relevance. In Daniel Dor, Chris Knight & Jerome Lewis (eds.), *The social origins of language*, 67–76. Oxford: Oxford University Press.

Kendon, Adam. 2017. Reflections on the "gesture-first" hypothesis of language origins. *Psychonomic Bulletin & Review* 24(1). 163–170.

Kinsella, Anna R. 2009. *Language evolution and syntactic theory*. Cambridge: Cambridge University Press.

Kirby, Simon. 2012. Language is an adaptive system: The role of cultural evolution in the origins of structure. In Maggie Tallerman & Kathleen R. Gibson (eds.), *The Oxford Handbook of Language Evolution*, 589–604. Oxford: Oxford University Press.

Kirby, Simon, Hannah Cornish & Kenny Smith. 2008. Cumulative cultural evolution in the laboratory: An experimental approach to the origins of structure in human language. *Proceedings of the National Academy of Sciences* 105(31). 10681–10686.

Kirby, Simon, Mike Dowman & Thomas L. Griffiths. 2007. Innateness and culture in the evolution of language. *Proceedings of the National Academy of Sciences* 104(12), 5241–5245.

Kirby, Simon, Tom Griffiths & Kenny Smith. 2014. Iterated learning and the evolution of language. *Current Opinion in Neurobiology* 28. 108–114.

Kurcz, Ida. 2004. Communicative competence and theory of mind. *Psychology of Language and Communication* 8(2). 5–18.

Larsen-Freeman, Diane. 2017. Complexity theory: The lessons continue. In Lourdes Ortega & ZhaoHong Han (eds.), *Complexity Theory and Language Development: In celebration of Diane Larsen-Freeman*, 11–50. Amsterdam, Philadelphia: John Benjamins.

Levinson, Stephen C. 2006. On the human "interaction engine". In Nick J. Enfield & Stephen C. Levinson (eds.), *Roots of Human Sociality*. 39–69. Oxford: Berg.

Levinson, Stephen C., & Judith Holler. 2014. The origin of human multi-modal communication. *Philosophical Transactions of the Royal Society B: Biological Sciences* 369(1651). 20130302.

Lewandowska-Tomaszczyk, Barbara. 2008. Czym jest język? Dzisiejsze kontrowersje w paradygmatach generatywnych i kognitywnych. In Piotr Stalmaszczyk (ed.), *Metodologie językoznawstwa: współczesne tendencje i kontrowersje*, 9–26. Kraków: Lexis.

Lewis, Eileen L. & Marcia C. Linn. 1996. Where is the heat? A response to David Pushkin. *Journal of Research in Science Teaching* 33. 335–337.

Lohmann, Heidemarie & Michael Tomasello. 2003. The role of language in the development of false belief understanding: A training study. *Child Development* 74(4). 1130–1144.

Little, Hannah, Kerem Eryılmaz & Bart de Boer 2017. Signal dimensionality and the emergence of combinatorial structure. *Cognition* 168. 1–15.

Lupyan, Gary & Rick Dale. 2010. Language structure is partly determined by social structure. *PLoS One* 5. e8559.

MacNeilage, Peter F. 2008. *The origin of speech.* Oxford & New York: Oxford University Press.

Martins, Pedro Tiago, Evelina Leivada, Antonio Benítez-Burraco & Cedric Boeckx. 2016. Biological pluralism in service of biolinguistics. In Koji Fujita & Cedric Boeckx (eds.), *Advances in biolinguistics: The human language faculty and its biological basis*, 153-169. New York: Routledge.

McNeill, David. 1992. *Hand and mind: What gestures reveal about thought.* Chicago: University of Chicago Press.

McNeill, David. 2012. *How language began: Gesture and speech in human evolution*. Cambridge: Cambridge University Press.

Mead, George H. 1974. *Mind, Self and Society*, edn. with introduction by Charles W. Morris. Chicago: University of Chicago Press.

Mendívil-Giró, José-Luis. 2019. Did language evolve through language change? On language change, language evolution and grammaticalization theory. *Glossa* 4(1), https://www.glossa-journal.org/articles/10.5334/gjgl.895/.

Mesoudi, Alex. 2011. *Cultural evolution: how Darwinian theory can explain human culture and synthesize the social sciences.* University of Chicago Press.

Miller, Earl K. & Timothy J. Buschman. 2013. Cortical circuits for the control of attention. *Current Opinion in Neurobiology* 23. 216–222.

Motamedi, Yasamin, Marieke Schouwstra, Kenny Smith, Jennifer Culbertson & Simon Kirby. 2019. Evolving artificial sign languages in the lab: From improvised gesture to systematic sign. *Cognition* 192. 103964.

Neubauer, Simon, Philipp Gunz, Jean-Jacques Hublin. 2010. Endocranial shape changes during growth in chimpanzees and humans: a morphometric analysis of unique and shared aspects. *Journal of Human Evolution* 59. 555–566.

Nölle, Jonas, Stefan Hartmann & Peter Tinits. 2020. Language evolution research in the year 2020. *Language Dynamics and Change* 10(1). 3–26.

Okanoya, Kazuo. 2007. Language evolution and an emergent property. *Current Opinion in Neurobiology* 17(2). 271–276.

Parker, Anna R. 2006. Evolving the narrow language faculty: was recursion the pivotal step? *The Evolution of Language*. 239–246.

Pinker, Steven & Ray Jackendoff. 2005. The faculty of language: what's special about it? *Cognition* 95(2). 201–236.

Pleyer, Michael & James Winters. 2014. Integrating cognitive linguistics and language evolution research. *Theoria et Historia Scientiarum* 11. 19–43.

Pleyer, Michael & Stefan Hartmann. 2019. Constructing a consensus on language evolution? Convergences and differences between biolinguistic and usage-based approaches. *Frontiers in Psychology: Language Sciences*. doi:10.3389/fpsyg.2019.02537

Poeppel, David. 2012. The maps problem and the mapping problem: Two challenges for a cognitive neuroscience of speech and language. *Cognitive Neuropsychology* 29. 34–55.

Poeppel, David & David Embick. 2005. Defining the relation between linguistics and neuroscience. In Anne Cutler (ed.), *Twenty-first Century Psycho-linguistics: Four Cornerstones*, 103–120. Hillsdale: Lawrence Erlbaum.

Pomerantz, Anita. 1984. Agreeing and disagreeing with assessments: Some features of preferred/dispreferred turn shapes. In Maxwell Atkinson & John Heritage (eds.), *Structures of Social Action: Studies in Conversation Analysis*, 57–101. Cambridge: Cambridge University Press.

Richerson, Peter J. & Robert Boyd. 2005. *Not by genes alone: how culture transformed human evolution*. Chicago: University of Chicago Press.

Ritt, Nikolaus. 2004. *Selfish Sounds and Linguistic Evolution: A Darwinian Approach to Language Change*. Cambridge: Cambridge University Press.

Roberts, Seán G., Christine Cuskley, Luke McCrohon, Lluis Barceló-Coblijn, Olga Fehér & Tessa Verhoef (eds.). 2016. *The Evolution of Language: Proceedings of the 11th International Conference (EVOLANG11).* doi:10.17617/2.2248195

Roberts, Seán G., Anton Killin, Angarika Deb, Catherine Sheard, Simon J. Greenhill, Kaius Sinnemäki, José Segovia-Martín, Jonas Nölle, Aleksandrs Berdicevskis, Archie Humphreys-Balkwill, Hannah Little, Christopher Opie, Guillaume Jacques, Lindell Bromham, Peeter Tinits, Robert M. Ross, Sean Lee, Emily Gasser, Jasmine Calladine, Matthew Spike, Stephen Francis Mann, Olena Shcherbakova, Ruth Singer, Shuya Zhang, Antonio Benítez-Burraco, Christian Kliesch, Ewan Thomas-Colquhoun, Hedvig Skirgård, Monica Tamariz, Sam Passmore, Thomas Pellard & Fiona Jordan. 2020. CHIELD: the causal hypotheses in evolutionary linguistics database. *Journal of Language Evolution*. doi:10.1093/jole/lzaa001

Samuels, Bridget. 2009. The third factor in phonology. *Biolinguistics* 3(2–3). 355–382.

Saussure, Ferdinand de. 1916. *Cours de linguistique générale*. Paris: Payot.

Scott-Phillips, Thomas C. 2015. *Speaking our minds: Why human communication is different, and how language evolved to make it special*. Houndmills: Palgrave Macmillan.

Scott-Phillips, Thomas C. & Simon Kirby. 2010. Language evolution in the laboratory. *Trends in cognitive sciences* 14(9). 411–417.

Smith, Kenny, Simon Kirby & Henry Brighton. 2003. Iterated learning: A framework for the emergence of language. *Artificial Life* 9. 371–386.

Smith, Kenny. 2018. The cognitive prerequisites for language: insights from iterated learning. *Current Opinion in Behavioral Sciences* 21. 154–160.

Smith, Maynard John & Eörs Szathmáry. 1995. *The major transitions in evolution*. Oxford & New York: W.H. Freeman Spektrum.

Smith, Kenny & Elizabeth Wonnacott. 2010. Eliminating unpredictable variation through iterated learning. *Cognition* 116(3). 444–449.

Sonesson, Göran. 1997. The ecological foundations of iconicity. In Irmengard Rauch & Gerald Carr (eds.), *Semiotics Around the World: Synthesis in Diversity*, 739-742. Berlin & New York: Mouton de Gruyter.

Sperber, Dan & Deirdre Wilson. 1986. *Relevance: Communication and cognition.* Cambridge, MA: Harvard University Press.

Steels, Luc. 2000. Language as a complex adaptive system. In Marc Schoenauer, Kalyanmoy Deb, Günter Rudolph, Xin Yao, Evelyne Lutton, Juan Julian Merelo & Hans-Paul Schwefel (eds.), *Parallel Problem Solving from Nature*, 17–26. Berlin, Heidelberg: Springer.

Számadó, Szabolcs & Eörs Szathmáry. 2006. Selective scenarios for the emergence of natural language. *Trends in Ecology & Evolution* 21(10). 555–561.

Tallerman, Maggie & Kathleen R. Gibson (eds.). 2012. *The Oxford Handbook of Language Evolution.* Oxford: Oxford University Press.

Thomas, James & Simon Kirby. 2018. Self domestication and the evolution of language. *Biology & Philosophy* 33(1). 9.

Tincoff, Ruth & Marc D. Hauser. 2006. Cognitive basis for language evolution in nonhuman primates. In Keith Brown (ed.), *The encyclopedia of language & linguistics*, 533–538. 2nd edn. Amsterdam: Elsevier.

Tinits, Peter, Jonas Nölle & Stefan Hartmann. 2017. Usage context influences the evolution of overspecification in iterated learning. *Journal of Language Evolution* 2(2). 148–159.

Tomasello, Michael. 1999. *The cultural origins of human cognition.* Cambridge, MA: Harvard University Press.

Tomasello, Michael. 2003. On the different origins of symbols and grammar. In Morten H. Christiansen & Simon Kirby (eds.), *Language Evolution*, 94–110. Oxford University Press.

Tomasello, Michael. 2006. Why don't apes point? In Nick J. Enfield & Stephen C. Levinson (eds.), *Roots of human sociality: Culture, cognition and interaction*, 506–524. Oxford: Berg.

Tomasello, Michael. 2008. *Origins of human communication.* Cambridge, MA: MIT Press.

Tomasello, Michael. 2009. The usage-based theory of language acquisition. In Edith Laura Bavin (ed.), *The Cambridge Handbook of Child Language*, 69–87. Cambridge: Cambridge University Press.

Tomasello, Michael. 2019. *Becoming human: A theory of ontogeny.* Cambridge, MA: Belknap Press.

Tomasello, Michael & Malinda Carpenter. 2007. Shared intentionality. *Developmental Science* 10(1). 121–125.

Tomasello, Michael, Ann Cale Kruger & Hilary Horn Ratner. 1993. Cultural learning. *Behavioral and Brain Sciences* 16(3). 495–511.

Tomasello, Michael, Malinda Carpenter, Josep Call, Tanya Behne & Henrike Moll. 2005. Understanding and sharing intentions: The origins of cultural cognition. *Behavioral and Brain Sciences* 28(5). 675–691.

Trijp, Remi van. 2011. Can iterated learning explain the emergence of case marking in language? *Proceedings of the 23rd Benelux Conference on Artificial Intelligence (BNAIC 2011), Volume 1*. 288–295.

Tuomela, Raimo. 2007. *The philosophy of sociality: The shared point of view*. Oxford: Oxford University Press.

Verhoef, Tessa, Simon Kirby & Bart de Boer. 2014. Emergence of combinatorial structure and economy through iterated learning with continuous acoustic signals. *Journal of Phonetics* 43. 57–68.

Vigliocco, Gabriella, Pamela Perniss & David Vinson. 2014. Language as a multimodal phenomenon: implications for language learning, processing and evolution. Philosophical Transactions of the Royal Society B: Biological Sciences 369(1651). 1–7. . https://royalsocietypublishing.org/doi/10.1098/rstb.2013.0292. (24 July, 2020.)

Wacewicz, Sławomir. 2007. Debata Hauser, Chomsky, Fitch – Pinker i Jackendoff. Nowoczesny spór o pochodzenie ludzkiej zdolności językowej. In Szymon Wróbel (ed.), *Modularność umysłu*, 79–94. Kalisz: WPA UAM.

Wacewicz, Sławomir. 2012. The narrow faculty of language: What is it, who has it, and how is it defined? *Theoria et Historia Scientiarum* 9. 217–229.

Wacewicz, Sławomir & Przemysław Żywiczyński. 2017. The multimodal origins of linguistic communication. *Language & Communication* 54. 1–8.

Wescott, Roger W. 1991. Defining Language. In Walburga von Raffler-Engel, Jan Wind & Abraham Jonker (eds.), *Studies in Language Origins: Volume 2*. John Benjamins Publishing. 77–84.

Willems, Roel M., Aslı Özyürek & Peter Hagoort. 2007. When language meets action: The neural integration of gesture and speech. *Cerebral Cortex* 17(10). 2322–2333.

Wittgenstein, Ludwig 1953. *Philosophical Investigations*. Translated by Elisabeth Anscombe. Oxford: Basil Blackwell.

Wróbel, Szymon. 2012. Rethinking language faculty: Has language evolved for other than language related reasons? *Theoria et Historia Scientiarum* 9, 201–216.

Zlatev, Jordan. 2008. From proto–mimesis to language: Evidence from primatology and social neuroscience. *Journal of Physiology–Paris* 102(1–3), 137–151.

Zlatev, Jordan. 2014. Human uniqueness, bodily mimesis and the evolution of language. *Humana. Mente Journal of Philosophical Studies* 7(27), 197–219.

Zlatev, Jordan. 2016. Preconditions in human embodiment for the evolution of symbolic communication In Gregor Etzelmüller & Christian Tewes (eds.), *Embodiment in Evolution and Culture*, 151–174. Tübingen: Mohr Siebeck.

Zlatev, Jordan. 2019. Mimesis theory, learning and polysemiotic communication. In Michael Peters (ed.), *Encylcopedia of Educational Philosophy and Theory*. Dordrecht: Springer.

Zlatev, Jordan, Sławomir Wacewicz, Przemysław Żywiczyński, & Joost van de Weijer. 2017. Multimodal-first or pantomime-first? Communicating events through pantomime with and without vocalization. *Interaction Studies* 18(3), 465–488.

Zlatev, Jordan, Przemysław Żywiczyński & Sławomir Wacewicz. 2020. Pantomime as the original human-specific communication system. *Journal of Language Evolution* 5(2), 156–174. doi:10.1093/jole/lzaa006

Żywiczyński, Przemysław. 2018. *Language Origins: From Mythology to Science.* Berlin: Peter Lang.

Żywiczyński, Przemysław & Jordan Zlatev. In press. The role of gesture in debates on the origins of language. In Alan Cienki (ed.), *Handbook of Gesture Studies*. Cambridge: Cambridge University Press.