The role of interactional and cognitive mechanisms in the evolution of (proto)language(s)

Michael Pleyer *

Center for Language Evolution Studies, Nicolaus Copernicus University in Toruń, Toruń, Poland
University Centre of Excellence IMSErt - Interacting Minds, Societies, Environments, Nicolaus Copernicus University in Toruń, Toruń, Poland

Abstract

This paper discusses the role of interactional and cognitive mechanisms in the emergence of (proto)linguistic structures and the evolution of (proto)language(s). Both the social, interactive nature of human communication and the interactional timescale have received increasing attention in investigations of how structure emerges in language. This has also led to an increasing focus on the mechanisms involved in the dialogic co-construction of structure and meaning in interaction. These include ad hoc constructionalization, interactive alignment, conceptual pacts, reuse and modification, and local forms of entrenchment, routinisation and schematisation. Interactional and cognitive mechanisms like these do not only play a crucial role in the emergence of structure in modern languages. They can also help explain how the first (proto)constructions came into being in hominin interaction. Frequently re-occuring, temporary, local (proto)constructions acquired increasing degrees of entrenchment, which led to their subsequent diffusion throughout hominin communities. They were then subject to processes of conventionalisation and cumulative cultural evolution. This process is hypothesised to eventually have led to the gradual transition from protolanguage to language.
© 2022 The Author. Published by Elsevier B.V. This is an open access article under the CC BY license (http://creativecommons.org/licenses/by/4.0/).

Keywords: Usage-based approaches; (Proto)language(s); Construction grammar; Interactional mechanisms; Language evolution

1. INTRODUCTION

Social cognition and interaction have increasingly been recognised as key factors in the evolution of (proto)language (s) (e.g. Levinson, 2006; Scott-Phillips, 2015; Tomasello, 2008). This theoretical development is in line with calls for an interactional and social turn in the language sciences, especially in usage-based approaches and cognitive linguistics (Harder, 2010; Zima and Brône, 2015). In this paper, I discuss interactional mechanisms that have been theorised to play an important role in the evolution of protolanguage(s) as well as in the transition to fully complex modern human

language. In doing so, I mostly focus on the role conversational interaction plays in the emergence of linguistic structure and propose that the mechanisms involved in it can also serve as an explanatory framework for the emergence of protolanguage(s). The paper adopts a usage-based construction grammar approach to the terms *language(s)* and *protolanguage(s)*. Usage-based Construction grammar explains language acquisition, use, and indeed evolution in terms of general cognitive mechanisms and does not assume a language-specific innate computational ability (Tomasello, 2003; Pleyer and Hartmann, 2019). In addition, the target of acquisition, and knowledge of language are not treated in terms of linguistic rules, but in term of schemas or generalizations over instances of actual language use in context, that is, constructions (e.g. Langacker, 1987; Tomasello, 2003). From the perspective of construction grammar, then, knowing language means having acquired constructions: form-meaning pairings of different degrees of schematicity, specificity, abstractness and complexity (e.g. Goldberg, 2003). Importantly, these properties are not binary, discrete categories, but are instead dimensional phenomena on a syntax-lexicon continuum. This means that constructions can be more or less simple vs complex, and more or less specific vs abstract. For example, monomorphemic word constructions such as *close*, *my*, *eyes* occupy spaces close to both the specificity and simplicity poles, whereas affix schemas such as [STEM]-[AFFIX] (e.g. *loveless, bloody*) are closer to both the schematicity and the complexity poles. Multiword expressions such as *Valentine's Day* or *soft as snow,* on the other hand, are close both to the specificity and complexity poles, whereas syntactic categories, such as [ADJ] and [N], are closer to the schematicity and simplicity poles (Stefanowitsch and Flach, 2017).

This gradualistic view of language also extends to definitions of language and protolanguage, which are seen to be on a continuum. Importantly, both protolanguage and language are kinds of symbolic systems that not only have a communicative, but also a conceptualising and perspectivising function (cf. Sinha, this volume). The transition from protolanguage to modern language is therefore seen as a gradual change in degrees of complexity and schematicity in the shared symbolic storage of constructions represented on the level of individual (proto)language users and its diffusion through a community. As there "is no sharp line between protolanguage and language" (Planer and Sterelny, 2021, p. 9), there is no clear cut-off point where protolanguage suddenly becomes language, or a specific point where a protolinguistic system becomes a linguistic one. The differences between protolanguage and modern languages, then, are one of degree and not of kind.

Along the same lines, usage-based construction grammar approaches also take the position that many of the processes that led to both the incremental emergence of protolanguages and the gradual transition from protolanguages to modern languages are the same ones that operate during and explain the emergence of structure in the diachronic change of modern languages (Bybee, 2010; Hartmann and Pleyer, 2021; Heine and Kuteva, 2007; Pleyer, 2017) That is, usage-based, constructionist approaches take a unified perspective on the mechanisms that are relevant for both "constructing a protolanguage", the transition from protolanguage to language, and historical language change. Regarding further terminology, the terms protolanguage and protolanguages will be used relatively interchangeably, with *protolanguage* focussing more on the overall status and stage of the different systems employed by different communities of protolanguage users at a given point in time, and *protolanguages* when focussing on the different systems in use by these communities.

The focus of this paper lies on reviewing important mechanisms and processes influencing the emergence of communicative structure in modern humans. In the following I first highlight the importance of the 'interactional' timescale in communicative interactions (Section 2). I then focus on the question what a focus on this timescale can tell us about language evolution and the emergence of (proto)linguistic structure in interaction (Section 3). In Section 4, I discuss how the emergence of protolinguistic structure in interaction can represent a first step in the entrenchment and conventionalisation (cf. Schmid, 2020) of such structures, which is the foundation for their subsequent grammaticalization and constructionalization and the transition from protolanguage(s) to modern language(s). I then describe some of the key interactional and cognitive mechanisms involved in the emergence of (proto)linguistic structure that need to form part of explanations of language evolution (Section 5). Specifically, I describe interactive alignment (Section 5.1), conceptual pacts (Section 5.2), reuse and modification of prior utterances (Section 5.3), the process of resonance activation (Section 5.4), as well as the processes of micro-entrenchment, and local-level routinisation and schematisation (Section 5.5). I end with a short outlook of potential synergies and further research that follow from the perspective on cognitive and interactional mechanisms in the evolution of (proto)language(s) described in this paper (Section 6).

2. EVOLUTIONARY TIMESCALES

There are a number of proposals for different timescales that are relevant in characterizing language (e.g. Kirby, 2012; Larsen-Freeman and Cameron, 2008; MacWhinney, 2005; Steels, 2011; see Enfield, 2014 for a discussion). They all represent converging sources of evidence for questions about the emergence, change, acquisition, processing,

and use of language. Within language evolution research, there are three timescales that have been the focus of most attention:

- **O**ntogenetic Timescale: The timescale of an individual's language development across the lifespan
- **P**hylogenetic: The timescale of the bio-cultural evolution of language
- **D**iachronic: The social/cultural historical timescale of language change

However, there is one further timescale that is of particular relevance to questions about the evolution of (proto)language(s) and its interactional foundations:

- **E**nchronic: The timescale of social interaction and language use in context

Increasingly, theoretical approaches in the language sciences also stress the importance of the 'online', dialogic, interactive nature of human communication in the emergence of linguistic structure. This timescale has been called 'enchronic' by Enfield (2014). For example, the enchronic, conversational level has been shown to play an important role in language change by proponents of the view that grammar is an emergent, interactional phenomenon (e.g. Hopper, 2011, Du Bois, 2014; Pleyer, 2017). This view is best encapsulated in the phrase that "[y]esterday's discourse is today's language; today's discourse is tomorrow's language." (Roberts, 1944, p. 299; cf. Hopper, 2011, p. 26). That is, the way that meanings are expressed in a particular utterance and sequence of words in a concrete situation by individual language users influences the historical development of the shared symbolic storage of a linguistic community (cf. Roberts, 1944, p. 293). This is also echoed by Christiansen and Chater (2022, p. x): "the relatively stable regularities of grammar are not the starting point. They are the result of countless generations of communicative interaction, through which linguistic patterns become ever more entrenched." For this reason, it is of central importance to theories of language evolution to outline the interactional and cognitive mechanisms that influence the emergence of structure in the dialogic co-construction of meaning, as well as the subsequent conventionalization of these emergent patterns.

In the next sections, I especially focus on how interactive mechanisms on the enchronic timescale, that is interactional mechanisms that take place on the level of interactive exchanges, can be related to the evolution of (proto)language(s).

3. LANGUAGE EVOLUTION AND THE EMERGENCE OF (pROTO)lINGUISTIC STRUCTURES IN INTERACTION

One of the key insights of approaches that concentrate on the enchronic, interactional timescale of communication is that conversational interaction can be seen as the "core ecology for language use" (Levinson and Holler, 2014). As argued for in interactional linguistics, this is the key context in which the function of linguistic structures should be investigated, as their main function is that they serve as resources for social interaction (Couper-Kuhlen and Selting, 2018). Interaction is a process in which embodied agents dynamically coordinate, co-create and negotiate meaning and perspectives in a process of participatory sense making (Di Paolo et al., 2018).

From the perspective of language evolution, one further central finding of such approaches is not only that linguistic structures serve as resources to share meaning and to solve communicative problems, but that these structures also emerge in the process of interaction between cooperative agents. Such emergent structures then have the potential to serve as "reusable material" for subsequent interactions and to spread through the community. This is why whenever humans communicate using language, they engage in a "process by which grammar is created" (Croft, 2001; Schmid, 2020). Language, on this view, is a dynamic system of emergent structures that in principle is always changing under the influence of cognitive and interactional mechanisms that operate on the dialogical, enchronic timescale. As Schmid (2020: 9) puts it, the "linguistic system is usage-based and emergent: the design and structure of language emerges from and is continuously updated by usage".

One of the main aims of usage-based and interactional, social-cognitive linguistics approaches is to describe how linguistic knowledge is shaped by social-interactional and cognitive mechanisms at the level of language users engaged in interaction (Divjak, 2019; Harder, 2010). The key commitment of these approaches is that "the emergence of grammatical patterning in on-line production, dialogically, and as a cooperative achievement" (Fried and Östman, 2005), is based on cognitive and interactional processes whose application to communicative utterances in interaction shapes the emergence of structure in language.

On the one hand, describing the processes that take place on the enchronic timescale represents an important part of characterizing the evolutionary foundations of language (Hartmann and Pleyer, 2021; Levinson, 2019; Levinson and Holler, 2014; Pleyer and Winters, 2014). That is, describing these processes and the sociocultural, interactional settings in which they take place, as well as their respective cognitive, as well as evolutionary foundations, are part of specifying

how "language-readiness" evolved in the human lineage (Arbib, 2012; Pleyer and Lindner, 2014; Pleyer and Hartmann, 2020). Together with elucidations of the processes relevant to language acquisition, language change, and so forth, they therefore make up part of the "shopping list" of processes and abilities needed for language to evolve, and whose evolutionary histories the science of language evolution needs to untangle.

Together with elucidations of the processes relevant to language acquisition, language change, and so forth, they therefore make up part of the "shopping list" of processes and abilities needed for language to evolve, and whose evolutionary histories the science of language evolution needs to untangle.

However, the focus on the dialogic, in-vivo, pragmatic and interactive emergence of structure has implications beyond characterising the language-ready brain. Namely, they form part of the explanatory framework for the question of how (proto)language(s) came to have structure in the first place, that is for the question of how (proto)language(s) evolved. Processes of grammaticalization and constructionalization form one important part of this puzzle, as they have been argued to have led to the gradual transition from protolanguage to modern language (e.g. Arbib, 2012; Bybee, 2010; Heine and Kuteva, 2007).[1] Grammaticalization describes a process in language change in which lexical words turn into grammatical structures. For example, in French, the lexical item "pas" (step) has become grammaticalized as part of the French "ne...pas" negation construction (e.g. "Il ne chante pas", *he doesn't sing*), with spoken French increasingly dropping the original "ne" negation part, marking a syntactic change in progress (e.g. "Je ne sais pas"→"Je sais pas", *I don't know*; cf. Ashby, 2001). From a construction grammar perspective, if a new constructional pattern emerges in the history of language change, this is referred to as constructionalization (Traugott and Trousdale, 2013). For example, in the history of English, the BE *going to* future construction ("e.g. I am going to regret this") has emerged out of a construction referring to physical movement (e.g. "I am going to the shop") (Bybee, 2010; Hartmann and Pleyer, 2021; Heine and Kuteva, 2007).

However, the enchronic emergence of structure in interaction represents another central, interlocking puzzle piece in that the mechanisms implicated in the interactive emergence of structure have the potential to explain how the first proto-linguistic structures emerged in interactive agents. In the following, I flesh out this picture more fully with regards to the mechanisms that are of special importance for this process. From the perspective of usage-based, constructionist approaches, these changes take place in terms of the emergence of and changes to constructions, that is form-meaning pairings and their relationship to each other within a dynamic network. In other words, emerging constructions underwent a dynamic process of ad hoc constructionalization (Brône and Zima, 2014; cf. Traugott and Trousdale, 2013). As will be further discussed in Section 4 these processes and their operation on the enchronic timescale also have implications for the subsequent conventionalization and constructionalization of structures in the development of (proto)languages and the transition of modern language(s).

4. LANGUAGE EVOLUTION AND THE ENTRENCHMENT AND CONVENTIONALIZATION OF (pROTO)lINGUISTIC STRUCTURES

The scenario that arises from the perspective outlined here is one of a complex scaffolding process in which protolinguistic structures emerged as temporary problem-solving devices in individual communicative interactions. Recurring problems and tasks led to the frequent re-emergence of particular structures (Dabrowska, 2010; Fischer, 2015), or as Christiansen and Chater (2022, p. 15) put it: "Language is a clumsy, chaotic product of the demands of the moment. Yet the improvisations we concoct to solve each new communicative challenge are shaped by how we solved previous challenges, which then informs how we cobble together the solution to the next." The likelihood of particular structures re-emerging repeatedly was also boosted by frequently re-occuring solutions to communicative problems being increasingly consolidated in memory and becoming increasingly entrenched, thus making them emerge more easily and frequently in subsequent interactions.

This also made them more likely to be used with different communicative partners, leading to diffusion and spread of emergent structures through the community. This process is described as "usualization" by Schmid (2020): "recurrent utterance types become established and are sustained and continually adapted as agreed-upon means of reaching recurrent communicative goals in specific contexts and can therefore function as tacit norms." Usualization then leads

---

[1] Note that while these and other authors agree in taking a gradualist perspective their conceptions of of what protolanguage are quite different from each other (see, e.g., Tallerman, 2012; Progovac, 2019, pp. 31–66 for discussion). In addition, Heine and Kuteva (2007, p. 4) do not use the term "protolanguage" and instead are concerned with changes that occurred from'early language" to 'modern languages' via processes of grammaticalization. That is, what is described as the protolanguage-language continuum here does not fit neatly into Heine & Kuteva's conceptualisation, who distinguish between non-language and early language on the one hand, and early language and modern languages on the other hand.

to the diffusion of emergent, usualised structures through the community, leading to their conventionalisation.[2] According to Schmid (2020) conventionalization refers to the process in which communities of interacting agents, for example members of a (proto)lingusitic community establish and readapt regularities in their communicative behaviours. This process is based on the repetition of usage activities during usage events, and can be defined as the continuous mutual matching and coordination of communicative practices and knowledge (Schmid, 2020).

On the social side, this dynamic can also be seen to have a spiralling, ratcheting effect (Tomasello, 1999), where particular resources and strategies for communicative problem-solving could be cumulatively built upon, changed and adapted in interaction. This is in line with the view that grammatical structures are "best understood as what has been ritualized from interactions" (Thompson and Couper-Kuhlen, 2005). On the cognitive side, more frequent use of (proto)constructions also has an effect of constructional generalisation. As constructions become more entrenched, automated, and more habitual for a (proto)language users, they also gradually become more schematic: "As the use of a construction expands incrementally, its mental representation is being adjusted too: the construction becomes more abstract, more schematic, more productive – more general" (Neels, 2020a). More schematic (proto)constructions also had the potential of widening and expanding their context of use in a similar manner as is found in grammaticalization (Himmelmann, 2004; Neels, 2020a). For example, in the history of English demonstratives grammaticalized into definitive articles (see the analogous development in French, where the Latin distal demonstrative "ille" was grammaticalized into the definitive article "le"). As a consequence, they also expanded the context in which they are used. That is they started "to co-occur regularly with proper names or nouns designating unique entities (such as *sun*, *sky*, *queen*, etc.), i.e. nouns they typically did not co-occur with before" (Himmelmann, 2004, p. 32). Context expansion is influenced by analogical mappings and associations of similar constructions and contexts, which is in line with approaches that assign processes of analogy a catalysing role in the evolution of protolanguage(s) (Brand et al., 2020).

Once these widening contexts became themselves conventionalised within a community, i.e. once the repeated contexts they were used in were absorbed into the conventional semantic meaning of the constructions, they could then potentially serve as points of departure for further elaborations and extensions of (proto)constructional schemas (De Smet, 2016; Neels, 2020a; Traugott and Trousdale, 2013).

In the context of cumulative culture, this process therefore has the potential to explain the emergence of the first protolinguistic structures, their subsequent rise in complexity and structuring, and their spread through communities. The gradual transition from protolanguage to modern language is assumed to have happened in the human lineage over an extended period of time, meaning that the processes described here were characteristic of the emergence of protolinguistic communication within the hominin lineage in general. Generally, a usage-based, constructionist view of a gradual transition from a protolinguistic phase to a linguistic phase, or from protolanguages to modern languages is consistent with the view that there were several protolanguages out of which extant modern human languages developed. In the words of Nichols, (2012, p. 572), this view is consistent with the assumption that "language originated gradually over a diverse population of pre-languages and pre-language families." It is also consistent with the view that basic protolanguages emerged independently in different homin species, such as in different homo erectus populations in Africa, Asia and Europe starting around 1.9 mya (Planer and Sterelny, 2021).

In combination with processes of grammaticalization and constructionalization, the processes outlined here also have the potential to explain the gradual transition to modern forms of human language through cumulative systemic changes on the level of 'communities of practice' (Eckert and McConnell-Ginet, 1992). Reoccurring structures turned into routines that were used with more and more communicative partners. This led to temporary, emergent communicative routines turning into an inventory of entrenched community-wide communicative routines – protolanguage.

Over time, these developed increasing degrees of conventionalisation within the community. As argued for by Planer and Sterelny (2021), rich, stable protolanguages depend on large-volume, reliable and accurate social learning within a community. The earliest archaeological signature we have for this type of learning is around 800kya, which suggests that richer protolanguages evolved around this timeframe in *heidelbergensis*. Vocal control increased between 1.9 mya and 800kya and was largely in place by 800kya. This suggests that vocal elements became more central in

---

[2] In the dynamic process of usualization, a linguistic unit or pattern becomes more common within a usage community. Increasing usualization therefore leads to conventionalization. For example, a term like "selfie" became less and less dependent on cotext and context to resolve its meaning, becoming more and more expected to be used as a term when somebody wanted to express they took a picture of themselves. At some point, most members of a usage community make this unit part of their usage, leading to its conventionalization (Schmid, 2020: 129).

multimodal protolanguages during the transition from homo erectus to homo heidelbergensis.[3] Protolanguages increasingly accumulated changes through processes of cumulative innovation, evolving into a fully grammaticalized and conventionalised structured inventory of constructions shared by a community – language. Networks of constructions approaching the linguistic pole on the protolanguage-language continuum likely have existed at least since the emergence of different homo sapiens populations across Africa (Scerri et al. 2018) and in Neanderthals and Denisovans (Dediu and Levinson, 2013; Dediu and Levinson, 2018; Johansson, 2015) starting around < 300 ka (Planer and Sterelny, 2021). They might possibly also date back to our last common ancestor with Neanderthals, 500kya (Dediu and Levinson, 2013).

However, these interactional mechanisms are hypothesised to also have led to the emergence of the first (proto)constructions, which are assumed to have a considerably deeper time depth. Levinson and Holler (2014), for example, argue "that the origin of modern speech, and the prevalence of the vocal tract in human communication, would seem to lie somewhere between 1.4 Ma and ca 600 000 years ago." On this view, many foundations for the emergence of a protolinguistic communication system would have already been in place in the immediate ancestor of Homo heidelbergensis, and in later erectines in general (cf. Planer and Sterelny, 2021). These include, for example, ritualised gestures, pointing, iconic gestures, and voluntary vocal utterances. In addition, they also include many aspects of the "interaction engine" underlying sociocognitively sophisticated bouts of collaborative communication characterised by turn-taking, processes of the negotiating of meaning as well as the stabilisation of communicative constructs in interaction and their diffusion in a community (Levinson and Holler, 2014).

Ancient humans around < 300kya likely had a fully language-ready brain, though, "they may well not yet have had a social life that would support the maintenance and transmission of full language" (Planer and Sterelny, 2021: 179). However, starting around this time there is evidence that homo sapiens groups grew in social scale, engaged in novel forms of complex cooperation, including more cooperation across as well as within groups, including delayed-return and reciprocal cooperation, division of labour, and new hunting technologies using high-velocity projectile weapons requiring smaller but more coordinated hunting parties (Planer and Sterelny, 2021). This in turn also required negotiating the complex interactions and the conflicts that come with these developments. Planer and Sterelny (2021) argue that this was a critical driver for the evolution of full language, involving the development of explicit norms and ways to express them linguistically, social role negotiation, social stress relief, specifying elaborate kinship relations to stabilise cooperative relationships, the reduction of aggression, as well as methods of sharing perspectives, such as storytelling, ritual, and gossip (see also Dunbar, 2009; Sterelny, 2016).

Structures that enabled interactants to negotiate these aspects of social life are therefore likely to have emerged in this transitional period in which (proto)languages moved closer to the linguistic pole of the protolanguage-language continuum, and to not have characterised earlier erectine and heidelbergensian protolanguages. Whereas some form of displacement is likely to have characterised earlier protolanguages, "linguistic resources to unambiguously express claims about past contributions and future expectations" would have been needed in the cooperative social environments of homo sapiens (Sterelny, 2016, p. 182). That is, at least with the advent of homo sapiens, linguistic resources managing more complex forms of displacement would have been needed than those found in earlier protolanguages. Such forms of grammatically marked displacement would also have been necessary for joint action planning (Planer and Sterelny, 2021) as well as for sharing information on the social reputations of others, their actions and the contexts of those actions, in order to explicitly manage and negotiate one's own and others' social 'face' (Goffman, 1967; Pleyer and Pleyer, 2022). In general, the less information can be assumed to be shared between interactants, the more communication needs to be regularised, conventionalised, and explicit, that is there are selective pressures for grammaticalization (Sterelny, 2016). Overall, then, the development of fully modern language and its "grammar of sharing" (Tomasello, 2008) built on more complex social interactions that emerged as *homo sapiens* developed more complex

---

[3] "Multimodal" here refers to the fact that language include "multiple concomitant channels of expression (i.e. gesture, prosody, facial expression and body movement), (Vigliocco et al., 2014), which can be seen as different expressive channels (Oben and Brône, 2016) that contribute to the meaning of an utterance in context. In the words of McNeill (1992, p. 2), elements in other modalities, such as gestures "gestures are an integral part of language as much as are words, phrases, and sentences." This multimodality is also hypothesized to have characterised protolanguages, and to a higher degree than in modern spoken languages, where the vocal modality is dominant (e.g. Arbib, 2012; Levinson and Holler, 2014). Modern spoken languages seem to include multimodal constructions (for example, the combination of an utterance with a headshake, Andrén, 2014; see also Steen and Turner, 2013). Moreover, gestures that are both conventional and familiar within a usage community and are tied to a specific meaning when tied to particular expressions can be argued to be just as much part of language as other constructions (Langacker, 2008) From the perspective of a multimodal protolanguage, such (proto)constructions and conventionalised gestures, would have been even more prevalent in protolanguages (For a critical discussion of the ambiguities of the term "multimodality", see Zlatev, 2019).

culture. This is the general, overall picture of the emergence of protolanguage(s) and their transition to modern languages.

One of the key challenges for further filling in the details of this very broad general picture is it to investigate the different types of processes implicated in it and the levels at which they operate. On the one hand, as already indicated, we need to distinguish between communication on the enchronic timescale, which leads to the online, in vivo emergence of structure in interactions. And on the other hand, we need to look at the effects repeated instantiations of these enchronic processes have across the diachronic level and the level of the history of a population. However, there is a second important distinction that cuts across the different timescales involved in the emergence of (proto)constructions, namely that between speaker-internal cognitive processes, and community-wide processes.

This distinction is also captured in Schmid's (2020) Entrenchment-and-Conventionalization model. This model is a usage-based, cognitive-linguistic model that combines individual, cognitive processes (entrenchment) and social processes (conventionalization) to explain how language usage shapes language. Entrenchment processes refer to the "the continuous routinization and re-organization of associations, depending on exposure to and frequency of identical or similar processing events" that happen on the level of the cognitive systems of individual interacting with each other. Conventionalization processes, on the other hand, refer to community-wide socio-pragmatic communicative processes of how norms and regularities in communicative behaviour diffuse and become established in a community. Conventionalization in this model does not operate on the minds of individual speakers, but over surface level populations of behaviours, or in the case of (proto)language, utterances (Schmid, 2020; cf. Croft, 2000). Both processes interact and reinforce each other, leading to the emergence and change of linguistic structure. As these processes are built on domain-general as well as sociocognitive processes, they also have potential applications for a model of the emergence of protolanguage(s).

One particular aspects that is highlighted in this model as well as in other recent work (e.g. Petré and Anthonissen, 2020; Petré and Van de Velde, 2018; Neels, 2020a,b) is that both macro-level diachronic change in communities and micro-level idiolectal lifelong changes in the individual and their interactions are relevant for modelling the factors involved in the emergence of structure across different timescales and levels.

A focus on "individuality in complex systems" (Petré and Anthonissen, 2020) from a constructionist perspective also offers a solution to the gradual shift from protolanguage to modern languages. As research on individual differences in language users' native language attainment has shown (Dąbrowska, 2012; Dąbrowska and Street, 2006), people can differ in their representations of the same token structures. In particular, they might differ in the level of abstraction at which they represent and store a particular instance of linguistic knowledge. Some language users may have "extracted an abstract generalization, others approximate their behaviour by relying on memorised exemplars or lexically specific patterns" (Dąbrowska, 2020).

This presents a possible pathway for a gradual shift from protolanguage to modern languages. That is, at a given point in the evolution of (proto)language(s), some interactants will still have relied mostly on exemplar representations with limited schematicity, thus being closer to the protolanguage pole of the protolanguage-language continuum. Other interactants at the same time will have abstracted higher-order schematic connections between constructs, thus being closer to the modern language pole of the protologanguage-modern language continuum. This means that they will have generalised schematic construction types over token sequences that other interactants do not treat as interrelated to the same degree, if at all. This implies that there was a time in the transition from protolanguage to modern language where more and more interactants recognised particular token structures, - or constructs,[4] as types, whereas others treated them as tokens. The former type of interactants then increasingly and gradually moved towards the modern language pole of the continuum by incrementally adding more abstract generalisations. While doing so, they still retained a significant number of specific constructions such as fixed and semi-fixed multiword expressions, idioms, prefabricated chunks as well as low-level schemas, which still characterise a significant part of modern languages today (Stefanowitsch and Flach, 2017; Pleyer and Hartmann, 2020).

What follows from this scenario is that at the same point in time, there were interactants with protolinguistic/linguistic systems that were similar and overlapped in terms of their token and surface production, but that differed in the underlying cognitive mechanisms of how these surface tokens were produced and interpreted and that differed in the degree of abstraction at which these surface tokens were represented. In other words, they produced similar constructs, but underlying these constructs were different constructions in the (proto)language users' mental inventory of constructions.

---

[4] In construction grammar, 'constructions' are defined as form-meaning pairing of differing degrees of schematicity and abstractness stored in long term memory, whereas 'constructs' are "actual physical realizations of constructions, that is, utterance-tokens (words, phrases, sentences) that instantiate constructions in discourse." (Fried, 2013, p. 423). (proto)language users internalising constructions therefore means that they started to represent generalisations over multiple constructs they were exposed to in discourse.

Such a usage-based perspective also shows that the question of "Who did the first speaker talk with?" or "who did a user of a particular protolanguage stage communicate with?" do not really capture the way hominin interactions work . (Proto)language users with more schematised networks of constructions communicated with (proto)language users with less abstract and schematised constructional repertoires. In addition, successful communication during a given communicative episode was also scaffolded by ostensive-inferential communication, which aided the resolution of potential producer/recipient mismatches (Johansson and Fujita, this volume, Heintz and Scott-Phillips, 2022) as well as repair mechanisms (Dingemanse et al., 2015) and the interactional mechanisms of the alignment of emergent structures described in this paper. This is what enabled hominins to play "(proto)language games" (cf. Wittgenstein, 1953) with each other, "a rich, analogical, metaphorical and potentially highly creative means of conveying clues to one another, which can require any amount of ingenuity and even playfulness to interpret." (Christiansen and Chater, 2022, p. 21). That is, even if the constructional representations of communicators differed from one another, due to the fundamental and inextricably social nature of language, they still had all the necessary requirements to collaboratively construct and improvise meaning in conversation.

Moving on from this stage, generational turnover will have played a role in the gradual protolanguage-language transition. As older members pass and leave the discourse community, certain constraints on what is accepted within that community are lifted. This means that younger (proto)language users with higher degrees of abstraction can produce a wider variety of constructs licensed by their more schematised constructional representation (Petré and Van de Velde, 2018). One of the key research desiderata that follow from the outline described here is to specify the interactional and cognitive mechanisms involved in the in-vivo interactional emergence of constructions. To illustrate this research programme, some of the most important mechanisms involved in this project will be outlined in the following section.

5. INTERACTIONAL AND COGNITIVE MECHANISMS INVOLVED THE EMERGENCE OF CONSTRUCTIONS IN INTERACTION

Research in dialogic construction grammar (Brône and Zima, 2014), emergent grammar (Hopper, 2015), and interactional linguistics (Couper-Kuhlen and Selting, 2018; Su, 2016) suggests that "the beginnings of the emergence of constructions can be found in individual interactions" (Su, 2016).

As argued by Brône and Zima (2014), in interaction language users set up temporary routines for the purposes of communication and mutual understanding. These routines can be conceptualised as "ad hoc constructions" and the process that leads to their emergence as "ad hoc constructionalization." Ad hoc constructions can be described as "locally entrenched form-meaning pairings within the boundaries of an ongoing interaction" (Brône and Zima, 2014). Importantly, these ad hoc constructions are co-created in interaction, as interlocutors "jointly set up local constructional routines with varying degrees of flexibility and fixedness" (Brône and Zima, 2014: 459). Many of the processes evolved in ad hoc constructionalization are the same that generally lead to the abstraction of constructional patterns from usage events that can be observed in all forms of language learning, albeit on a more local level. However, this means that for a characterization of the interactional mechanisms involved in constructionalization in interaction, observations which processes lead to the internalization of constructions are highly relevant as well.

As argued for in Pleyer (2017), ad hoc constructionalization can also serve as a mechanism to get protoconstructions off the ground and lead to the emergence of protolanguage, making these processes even more crucial for specifying a usage-based, constructionist model of language evolution. However, the present paper goes beyond focussing on the role of ad hoc constructionalization, by looking at a wider range of mechanisms involved in the emergence of structure in interaction. Key processes involved in the emergence of structure in conversation are the following, which will be outlined in more detail below: interactive alignment, conceptual pacts, resonance activation, reuse and modification, and routinisation, schematisation, and entrenchment at the local level. Whereas interactive alignment, conceptual pacts and resonance activation are processes that operate at the local level of an interaction, processes of reuse, modification, routinisation, schematisation, and entrenchment take place over different timescales. This means that they are not only relevant for the emergence of constructions in interaction, but also for their subsequent development over longer periods of time. This also makes them especially important from an evolutionary perspective, as they have the potential to explain not only how the first protolinguistic structures emerged in interaction, but also how these structures stabilised into protolinguistic systems, or networks of protoconstructions in communities of protolanguage users.

5.1. Interactive alignment

The process of interactive alignment describes how interactants align their linguistic representations on different levels (Pickering and Garrod, 2005). Alignment can take place across different levels of representation, including align-

ment on the lexical and phrasal level, but also on the level of structures and thematic relations. It helps language users to converge on shared conceptualisations of a situation and reaching communicative goals.

Importantly, interactive alignment extends beyond linguistic levels, as interactants also show coordination in terms of gaze, gesture, body posture, and other nonverbal dimensions. This means that behavioural alignment is in fact a multimodal phenomenon that involves processes of synchronisation in different modalities interacting with each other (e.g. between speech and gesture, or speech and facial expressions, Levinson and Holler, 2014 Oben and Brône, 2016; Rasenberg et al., 2020). As alignment is not confined to linguistic levels of presentation, but a general semiotic and behavioural process, it therefore also has an important place in capturing how hominins converged on aligned expressive behaviour that represented the basis for the emergence of protolinguistic structure.

Two foundational processes discussed in theoretical approaches to alignment are that of priming and grounding (Oben, 2018; Rasenberg et al., 2020). Priming refers to a low-level, individual automatic process in which the instantiation of a particular pattern of activation makes the subsequent activation of that pattern, as well as closely connected patterns, more likely. A further distinction in priming is that between self-priming and other-priming. In self-priming, the activation of a particular structure or pattern of activation increases the subsequent retrievability of that structure of the producer and makes it more likely that they use the pattern again within an interactive bout. In other-priming, use of a pattern or structure by one interlocutor makes it more likely that that pattern or structure will be used by another interlocutor in the dialogic exchange (Neels, 2020a; Pickering and Garrod, 2017). Both processes of priming have been suggested to also play a role in grammaticalization, as they increase the likelihood of particular structures emerging in interaction (Mair, 2017; Neels, 2020a; Pickering and Garrod, 2017). For instance, concerning a linguistic change that was spearheaded by people born in the 1970s in North America, and by now can be considered the dominant quotative construction in Spoken Northern American English (Tagliamonte, 2016), a speaker using the quotative *be like* construction in conversation ("And I was just like, 'Oh my gosh,'" Tagliamonte, 2016, p. 50) would have made it more likely for their interlocutor to use the same construction later in conversation, influencing its subsequent constructionalization. This is another argument for the potential relevance of priming mechanisms for the emergence of protolinguistic structure and the transition to linguistic structure.

Grounding, on the other hand, focuses on processes of alignment as a result of high-level, interactive efforts of interlocutors to jointly coordinate referential expressions. That is, grounding refers to temporary agreements reached by interlocutors on specific construals on different expressive channels as they interactively co-create meaning, and negotiate a shared perspective on a situation (Oben, 2018; Rasenberg et al., 2020). When looking at interactive encounters, it is likely that both processes of interactive synchronization are in operation (Oben and Brône, 2016). In fact, Rasenberg et al. (2020) argue that priming and grounding should not be seen as dichotomous, but that instead alignment processes should be seen as taking place on a continuum with priming and grounding representing the respective endpoints of that continuum.

From an evolutionary perspective, what research on interactive alignment indicates is that the interactive mechanisms involved in the emergence of (proto)linguistic structure in communicative encounters involve both more low-level, automatic processes, as well as higher-level, mediated processes that involve hominin's advanced abilities in the domain of perspective-taking, social cognition, and ostensive-inferential cooperative communication (Heintz and Scott-Phillips, 2022; Pleyer, 2017; Scott-Phillips, 2015; Tomasello, 2019; Pleyer and Lindner, 2014).

5.2. Conceptual pacts

As a specific expression of grounding processes in interactive alignment, the notion of conceptual pacts (or 'referential pacts') is of particular importance when considering how interacting agents converge, and first converged on shared referential expressions in the evolution of protolanguage. In interaction, interlocutors often converge on specific lexical referential expressions for an object, situation, or event (for example, if someone is a 'thief' or a 'robber', or something is a 'bush' or a 'tree'). Conceptual pacts are often important for successful communication, especially when interlocutors need to agree on a particular construal, expression or perspective to refer to a referent that is ambiguous, for which no prior agreed upon expression exists (for example in a maze-game task, cf. Garrod and Doherty, 1994) or if multiple expressions can possibly be applied to the same referent. Matthews et al. (2010), for example, showed that by 3 years of age, children are sensitive to the fact that conceptual pacts can be seen as agreements "made between two people to take a given perspective on an object, where this agreement is believed by both interlocutors to be mutual, and its maintenance is understood to be cooperative" (Matthews et al., 2010).

Importantly, such conceptual pacts are at first specific to a given interaction. This is also understood by children, who have an awareness that a conceptual pact agreed upon with one interlocutor (e.g. calling a particular ambiguous roundish toy "a nose"), does not hold for interactions with another interlocutor (who might call the same object "an apple"; cf. Matthews et al. 2010).

According to Brône and Zima (2014), ad hoc constructions can be seen as conceptual pacts between interlocutors that are set up as temporary routines during an unfolding interaction. As such, the process of agreeing on conceptual pacts has obvious implications for language evolution, as it represents one kind of interactional mechanism that can lead to the emergence of protolinguistic structures. From such an evolutionary perspective, it is also highly relevant which factors influence the process of setting up conceptual pacts, as they can shed light on the processes involved in the first protolanguage users setting up conceptual pacts in interaction to solve communicative tasks. Conceptual pacts for situations where interlocutors need to agree on a new referring expression can also be influenced by the affordances of the environment (Nölle et al., 2020). However, as soon as they acquire a history of being successfully used to solve recurrent interactional tasks, especially when used with multiple interactants, just as other processes of interactive alignment they can acquire degrees of conventionalisation (Garrod and Doherty, 1994), and have also been implicated in potentially leading to processes of language change and grammaticalization (Nölle et al., 2020; Pickering and Garrod, 2017). Both interactive alignment and the specific process of forming conceptual pacts therefore are potentially important interactional mechanisms involved in the emergence of (proto)constructions in the evolution of (proto)language(s).

5.3. Reuse and modification

Processes of interactive alignment on a more general level constitute processes of reuse, which also offer participants in an interaction the option to modify reused structures. That is, in interaction interlocutors not only reuse prior utterances, but also modify them. On this view, previous discourse is a resource that language users can draw on in new usage events (Langacker, 1987; Zima, 2013). This is in line with research indicating that formulaic utterances and entrenched chunks of linguistic structure, including speech formulae (e.g. "see you later", "have a nice day!"), idioms (e.g. "to spill the beans", "to kick the bucket"), fixed phrases (e.g. "monkey see, monkey do", "the more, the merrier", "by and large") and semi-fixed phrases (e.g. "pass the N please" "let's give N a round of applause"), make up for a much larger part of our knowledge of use than was previously believed (cf. Wray, 2002). For example, formulaic expressions have been estimated to make up almost 25% of everyday speech (Van Lancker-Sidtis and Rallon, 2004). This makes it likely that the interactions of protolinguistic hominins at one time also consisted to a large degree of such prefabricated chunks or fixed expressions.

However, such processes have also been suggested to represent a way in which grammar can emerge through building on and modifying previous "material" from prior utterances. In the words of Goodwin (2013, p. 9), "building subsequent utterances through decomposition, reuse, and transformation of language structure provided by another is a central locus for grammar as a form of public, social practice" (cf. Su, 2016).

From an evolutionary perspective, this opens up a way to describe the way in which individual interactions are the communicative format in which (proto)constructions first emerged. Interestingly, additional support for this view also comes from usage-based investigations of language acquisition. For example, Lieven et al. (2009) have shown that when looking at the multiword expressions of two-year-old children, between 58–92% of their utterances where either a) verbatim repetitions of a previous utterance, or b) could be generated by a single operation on a previous utterance. This also represents a possible pathway for the gradual emergence of (proto)linguistic structure through reuse and modification of material from previous utterances. Reuse and modification then, are likely mechanisms both for the emergence of constructions in interaction, but also for the diffusion and cumulative accretion of constructional changes in communities of interacting agents.

Reuse and modification take place across different timescales, meaning that on the one hand interactants reuse and modify material from a given interactional exchange. But on the other hand, they also reuse and modify material from previous interactional encounters. In terms of the mechanisms involved in these processes, it is again probable that both low-level individual and more high-level interactive features are at play. Within interactional enchronic encounters, priming (cf. Section 5.1) is an important mechanism that sets the stage for the reuse and modification of prior utterances, by "creating cognitive conditions that facilitate the selective reproduction of recently used linguistic forms and structures" (Du Bois, 2014).

Other such individual processes that influence reuse and modification are the micro-entrenchment of structures in interaction as well as processes of local schematisation and routinisation, which is discussed in more detail in Section 5.5. In terms of high-level interactive mechanisms, reuse and modification are influenced by grounding mechanisms in interactive alignment described in Section 5.1 as well as processes of resonance activation, which is described in the next section.

5.4. Resonance activation

Just as conceptual pacts, resonance activation can be seen as an interactional mechanism that is located more towards the grounding pole of the priming-grounding continuum. This is the case because it relates to the way that

speakers use materials and structure from previous utterances for purposes of stance taking and expressing perspectives on a situation. Resonance activation can broadly be defined as "the activation of affinities across utterances" (Du Bois, 2014). These affinities are intersubjective and conceptual-semantic correspondences that are built upon when interlocutors re-use and exploit in their own utterances forms and meanings expressed by their interlocutor's utterances. As an example of resonance activation, take the following exchange between two preschoolers, which is often used as an example by Deborah Tannen (e.g. Tannen, 2010):

> GIRL 1: Did you know. . .. . . Did you know my babysitter, called Amber, has already contacts?
> GIRL 2: My mom has already contacts and my dad does, too!
> [Girl 1 laughs]
> Girl 1: . . .. . .the same?!

In this example, the syntactic pattern of the second speaker resonates with that of the first speaker, including the non-conventional word order ("My [NP] has already contacts"), which in the terms of DuBois' (e.g. 2014) dialogic syntax can be interpreted as an expression of stance-taking and affiliation. Resonance activation can also be used for multiple different purposes, including different conceptualisations of the same situation or entity, but they share the fact that they represent a mechanism where structure is co-created on-line and in-vivo in interactional encounters to negotiate and express perspectives. As such, resonance activation represents another mechanism implicated in how protolinguistic hominins converged on the first shared constructions. Through resonance activation, interactants also engage in a process of "dialogic bootstrapping" (Du Bois, 2014) where they attend to analogies between others' and their own utterances, which enables them to selectively replicate structures and patterns found in these utterances. From a cognitive perspective, this then also leads to the entrenchment of these structures in language users, thus representing another mechanism by which language users learn linguistic structures (Du Bois, 2014).

5.5. Micro-entrenchment, local-level routinisation and schematisation

Entrenchment is a key concept in usage-based approaches (Blumenthal-Dramé, 2012; Divjak, 2019; Langacker, 1987; Schmid, 2016, 2020). It captures the way in which repeated encounters in language usage impact cognitive encoding and storage and strengthen the representation of these structures in memory (Divjak, 2019). In Langacker's (1987) original phrasing:

> "Every use of a structure has a positive impact on its degree of entrenchment, whereas extended periods of disuse have a negative impact. With repeated use, a novel structure becomes progressively entrenched, to the point of becoming a unit; moreover, units are variably entrenched depending on the frequency of their occurrence."

Importantly, entrenchment is a matter of degree, with structures exhibiting stronger and weaker degrees of entrenchment in memory. The process of entrenchment is influenced by a whole host of different factors, which include, but are in no way limited to frequency effects, different types of salience, attention, contextual and pragmatic effects and processes of memory consolidation (Divjak, 2019; Schmid and Günther, 2016; Schmid, 2016). According to Divjak (2019), there are three types of effects that entrenchment has on linguistic representations:

1. "What You Do Often, You Do Faster and with Fewer Errors:" A structure that is cognitively entrenched in memory is activated and processed with reduced cognitive effort and conversely, with an increased degree of automatization.
2. "Units that Occur Together, Refer Together:" The frequent association of structures leads to co-occurring structures achieving unit status, which eases their simultaneous access and retrieval.
3. "Units that Occur Together, Blur Together:" Structures that become entrenched are often chunked together, which in some cases can lead to the emergence of a fused structure such as in *I don't know → dunno* and *going to → gonna* (Bybee, 2010).

These processes have wide-ranging implications for usage-based research on language evolution (Pleyer and Hartmann, 2020), language learning (Pickering and Garrod, 2005; Schmid, 2016), language change (Hilpert, 2017; Neels, 2020b), and the conventionalisation of entrenched units in (proto)linguistic communities (Hartmann and Pleyer, 2021).

However, Brône and Zima (2014) also argue that entrenchment at the local level of an interaction, or "micro-entrenchment" also plays a crucial role in the emergence of ad hoc constructions. As they argue, ad hoc constructionalization "essentially involves the same type of entrenchment processes as in language acquisition (Pickering and Garrod, 2005: 99) and grammaticalization" (Brône and Zima, 2014, p. 467). Repeated reuse and modification driven by alignment and

resonance leads to micro-entrenchment, making ad hoc constructions "locally entrenched parts of the repertoire (or 'mini-language', Du Bois, p.c.) of the specific interaction from which they emerge" (Brône and Zima, 2014, p. 483). Framing this repertoire of ad hoc constructions in terms of a 'mini-language' of course has interesting potential implications for conceptualising the emergence of localised repertoires of ad hoc protoconstructions (or protoconstructicons) in interaction as a crucial step in the evolution of protolanguage(s) (Pleyer, 2017).

Furthermore, as Divjak's (2019) distinction of three types of entrenchment shows, entrenchment is not only a process related to simple storage in memory, but can also lead to more routinised and schematised behaviours. Schematisation as a discourse process can create both item-based schemas (such as "Where's-the X?", "I-wanna X", "Look at X"), and as discourse progresses has the potential of turning these schemas into more abstract and elaborate structures (such as [Pronoun] [Verb of Desire] [Action], e.g. "I would like to have a cookie") (cf. Tomasello, 2003, 2014). Schematisations over instances of actual language usage are one of the core mechanisms thought to be responsible for language learning and language change in usage-based approaches (Barlow and Kemmer, 2000; Diessel, 2019; Langacker, 1987; Traugott and Trousdale, 2013). In Brône and Zima's (2014) model, they are also hypothesised to operate on the local level, and in tandem with local routinisation, represent important cognitive and interactional mechanisms involved in ad-hoc constructionalization. In addition, these local processes also represent important candidates for mechanisms involved in modelling the process of the emergence of protoconstructions in hominins.

In all, the interactional and cognitive mechanisms outlined in Section 4 are important processes that have the capacity to explain how protolinguistic structure could have emerged in communicative encounters of interacting hominins on the enchronic timescale. However, as already indicated, this description of possible processes that lead to the emergence of protoconstructions also has implications for a model of how ad hoc protoconstructions became entrenched in the minds of individual protolanguage users. In such a model, if temporary structures emerge and recur frequently, they stabilise across multiple discourse encounters and emerge as schematised community-wide construction without being limited to individual interactional encounters. Building on this, they can then be subject to further cultural evolution and cumulative change such as grammaticalization and constructionalization, leading to the gradual transition of protolanguage(s) to modern language(s), as outlined in Section 4.

6. CONCLUSION

In this paper I have highlighted the potential role of interactional and cognitive mechanisms in the evolution of (proto) language(s). Specifically, I have shown how a process of ad hoc constructionalization on the enchronic timescale driven by these mechanisms could have led to the emergence of (proto)constructions in interactions. These emergent, temporary protoconstructions could then have provided the starting point for increasing entrenchment and conventionalisation of frequently recurring patterns, setting the stage for the further evolution of protolanguage(s) and the gradual transition towards the language pole on the protolanguage-language continuum.

I have discussed some models that promise to be highly relevant in further addressing the dynamics and processes evolved in the emergence of protolanguage(s) and their transition to modern language, such as the Entrenchment-and-Conventionalisation Model (Schmid, 2020), as well as the relevance of factors also involved in grammaticalization and constructionalization. However, this discussion has only scratched the surface of relevant mechanisms. A more extended look at how the various processes in this and other models as well as grammaticalization and constructionalization research may have operated in the evolution of (proto)language(s) promises to have wide-ranging implications.

Overall, the focus on cognitive, social and interactive mechanisms on the enchronic timescale on the one hand, and on community-wide socio-communicative processes of diffusion and conventionalisation on the other hand has a number of important implications for future research. One of them is the desideratum to integrate sociolinguistic research and theorizing into accounts of the emergence, stabilisation, and diffusion of linguistic structure. This point is forcefully made by Roberts and Sneller (2020). They show that a) language evolution research so far has made little contact with sociolinguistics and that b) integrating language evolution research with sociolinguistics and initiating a theoretical debate between these fields has the potential of leading to significant theoretical and empirical progress. Such an integration can prove profitable both for evolutionary linguistic research on the emergence of protolanguage(s) and the transition to modern language(s).

As the present paper has also shown, another research area with important implications for language evolution is research on the changes in the language use of individuals across the life-span. After being neglected for a long time, this aspect has recently been receiving increasing attention in diachronic linguistics (De Smet, 2016; Neels, 2020a; Petré and Anthonissen, 2020; Petré and Van de Velde, 2018). Research on the life-span change in the organisation of the structured network of interconnected items internalised by communicative agents can help in outlining the processes that would have operated in the minds of protolanguage users and how new constructions could have emerged within an individual's life-span. That is, this area of research can help in investigating not only the process of how con-

structions become entrenched in a (proto)language user's (proto)linguistic system, but also how this system changes with age, repeated language usage and interaction with a community. This research can therefore help language evolution address the tension and relation of changes at the micro-level of the individual and the macro-level of a linguistic community (Neels, 2020b: 360; De Smet, 2016).

Lastly, whereas the scenario outlined here drew from different sources of evidence to build a verbal theory, this model would surely profit from further theory building informed by computational modelling of the interplay of cognitive and interactional mechanisms operating on individual (proto)language users as well as populations of utterances within a community across different timescales (Guest and Martin, 2020; van Rooij and Baggio, 2021; Zuidema et al., 2020).

Data availability

No data was used for the research described in the article.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

ACKNOWLEDGEMENTS

References

Andrén, M., 2014. Multimodal constructions in children: Is the headshake part of language?. Gesture 14 141–170.

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

Ashby, W.J., 2001. Un nouveau regard sur la chute du ne en français parlé tourangeau: s'agit-il d'un changement en cours? J. Fr. Lang. Stud. 11, 1–22.

Barlow, M., Kemmer, S., 2000. Usage-based models of language. CSLI Publications, Stanford, CA.

Blumenthal-Dramé, A., 2012. Entrenchment in usage-based theories : what corpus data do and do not reveal about the mind. Mouton De Gruyter, Berlin.

Brand, C.O., Mesoudi, A., Smaldino, P., 2020. Analogy as the core of cumulative cultural evolution. PsyArXiv..

Brône, G., Zima, E., 2014. Towards a dialogic construction grammar: Ad hoc routines and resonance activation. Cognitive Linguistics 25, 457–495.

Bybee, J., 2010. Language, Usage and Cognition. Cambridge University Press, Cambridge.

Christiansen, M.H., Chater, N., 2022. The Language Game: How Improvisation Created Language and Changed the World. Penguin Press, London.

Couper-Kuhlen, E., Selting, M., 2018. Interactional Linguistics: Studying Language in Social Interaction. Cambridge University Press, Cambridge.

Croft, W., 2000. Explaining Language Change : An Evolutionary Approach. Longman, Harlow, England, New York.

Croft, W., 2001. Typology. In: Aronoff, M., Rees-Miller, J. (Eds.), The Handbook of Linguistics. Blackwell, Malden, MA, pp. 369–393.

Dabrowska, E., 2010. The mean lean grammar machine meets the human mind: empirical investigations of the mental status of linguistic rules. In: Schmid, H.-.-J., Handl, S. (Eds.), Cognitive foundations of linguistic usage patterns: empirical approaches. De Gruyter, Berlin, pp. 151–170.

Dąbrowska, E., 2012. Different speakers, different grammars: Individual differences in native language attainment. Linguist. Approaches Bilingual. 2, 219–253.

Dąbrowska, E., 2020. Language as a phenomenon of the third kind. Cognitive Linguistics 31, 213–229.

Dąbrowska, E., Street, J., 2006. Individual differences in language attainment: Comprehension of passive sentences by native and non-native English speakers. Lang. Sci. 28, 604–615.

De Smet, H., 2016. How gradual change progresses: The interaction between convention and innovation. Lang. Var. Chang. 28, 83–102.

Dediu, D., Levinson, S.C., 2013. On the antiquity of language: the reinterpretation of Neandertal linguistic capacities and its consequences. Front. Psychol. 4, 397.

Dediu, D., Levinson, S.C., 2018. Neanderthal language revisited: not only us. Curr. Opin. Behav. Sci. 21, 49–55.

Di Paolo, E.A., Cuffari, E.C., De Jaegher, H., 2018. Linguistic bodies: The continuity between life and language. MIT Press, Cambridge, MA London, England.

Diessel, H., 2019. The Grammar Network: How Linguistic Structure is Shaped by Language Use. Cambridge University Press, New York, NY.

Dingemanse, M., Roberts, S.G., Baranova, J., Blythe, J., Drew, P., Floyd, S., Gisladottir, R.S., Kendrick, K.H., Levinson, S.C., Manrique, E., 2015. Universal principles in the repair of communication problems. PLoS One 10, e0136100.

Divjak, D., 2019. Frequency in Language: Memory, Attention and Learning. Cambridge University Press, New York, NY.

Du Bois, J.W., 2014. Towards a dialogic syntax. Cognitive Linguist. 25, 359–410.

Dunbar, R., 2009. Why Only Humans Have Language. In: Botha, R., Knight, C. (Eds.), The Prehistory of Language. Oxford University Press, Oxford, pp. 12–35.

Eckert, P., McConnell-Ginet, S., 1992. Think practically and look locally: Language and gender as community-based practice. Ann. Rev. Anthropol. 21, 461–490.

Enfield, N.J., 2014. Natural Causes of Language: Frames, Biases and Cultural Transmission. Language Science Press, Berlin.

Fischer, K., 2015. Conversation, Construction Grammar, and cognition. Lang. Cogn. 7, 563–588.

Fried, M., 2013. Principles of Constructional Change. In: Hoffmann, T., Trousdale, G. (Eds.), The Oxford Handbook of Construction Grammar. Oxford University Press, Oxford, pp. 419–437.

Fried, M., Östman, J.-O., 2005. Construction Grammar and spoken interaction: the case of pragmatic particles. J. Pragmat. 37, 1752–1778.

Garrod, S., Doherty, G., 1994. Conversation, co-ordination and convention: An empirical investigation of how groups establish linguistic conventions. Cognition 53, 181–215.

Goffman, E., 1967. Interaction Ritual: Essays on Face-to-face Interaction. Anchor Books, Garden City.

Goldberg, A.E., 2003. Constructions: a new theoretical approach to language. Trends Cogn. Sci. 7, 219–224.

Goodwin, C., 2013. The co-operative, transformative organization of human action and knowledge. J. Pragmat. 46, 8–23.

Guest, O., Martin, A.E., 2020. How computational modeling can force theory building in psychological science. PSyArXiv..

Harder, P., 2010. Meaning in Mind and Society: A Functional Contribution to the Social Turn in Cognitive Linguistics. De Gruyter, Berlin New York.

Hartmann, S., Pleyer, M., 2021. Constructing a Protolanguage: Reconstructing Prehistoric Languages in a Usage-Based Construction Grammar Framework. Philosoph. Trans. Roy. Soc. B: Biol. Sci..

Heine, B., Kuteva, T., 2007. The Genesis of Grammar: A Reconstruction. Oxford University Press, Oxford.

Heintz, C., Scott-Phillips, T., 2022. Expression unleashed: The evolutionary & cognitive foundations of human communication. Behavioral and Brain Sciences. https://doi.org/10.1017/S0140525X22000012.

Hilpert, M., 2017. Frequencies in Diachronic Corpora and Knowledge of Language. In: Hundt, M., Mollin, S., Pfenninger, S.E. (Eds.), The Changing English Language: Psycholinguistic Perspectives. Cambridge University Press, Cambridge, pp. 49–68.

Himmelmann, N., 2004. Lexicalization and grammaticization: Opposite or orthogonal? In: Bisang, W., Himmelmann, N., Wiemer, B. (Eds.), What Makes Grammaticalization? A Look from Its Fringes and Its Components. De Gruyter Mouton, Berlin; New York, pp. 21–42.

Hopper, P.J., 2011. Emergent Grammar and Temporality in Interactional Linguistics. In: Auer, P., Pfänder, S. (Eds.), Constructions: Emerging and Emergent. De Gruyter, Berlin, pp. 22–44.

Hopper, P.J., 2015. An Emergentist Approach to Grammar. In: MacWhinney, B., O'Grady, W. (Eds.), The Handbook of Language Emergence. Wiley-Blackwell, Malden, MA, pp. 314–327.

Johansson, S., 2015. Language abilities in Neanderthals. Ann. Rev. Linguist. 1, 311–332.

Kirby, S., 2012. Language is an Adaptive System: The Role of Cultural Evolution in the Origins of Structure. In: Tallerman, M., Gibson, K.R. (Eds.), The Oxford Handbook of Language Evolution. Oxford University Press, Oxford, pp. 589–604.

Langacker, R.W., 1987. Foundations of Cognitive Grammar. Vol. I: Theoretical Prerequisites. Stanford University Press, Stanford, CA.

Langacker, R., 2008. Metaphoric gesture and cognitive linguistics. In: Cienki, A., Müller, C. (Eds.), Metaphor and gesture. John Benjamins, Amsterdam, pp. 249–251.

Larsen-Freeman, D., Cameron, L., 2008. Complex Systems and Applied Linguistics. Oxford University Press, Oxford.

Levinson, S.C., 2006. On the Human "Interaction Engine". In: Enfield, N., Levinson, S.C. (Eds.), Roots of Human Sociality: Culture, Cognition and Interaction. Berg, New York, pp. 39–69.

Levinson, S.C., 2019. Interactional foundations of language: The interaction engine hypothesis. In: Hagoort, P. (Ed.), Human Language: From Genes and Brain to Behavior. MIT Press, Cambridge, MA, pp. 189–200.

Levinson, S.C., Holler, J., 2014. The origin of human multi-modal communication. Philos. Trans. R. Soc., B 369, 20130302.

Lieven, E., Salomo, D., Tomasello, M., 2009. Two-year-old children's production of multiword utterances: A usage-based analysis. Cognitive Linguist. 20, 481–507.

MacWhinney, B., 2005. The emergence of linguistic form in time. Connect. Sci. 17, 191–211.

Mair, C., 2017. From Priming and Processing to Frequency Effects and Grammaticalization? Contracted Semi-Modals in Present-Day English. In: Hundt, M., Mollin, S., Pfenninger, S.E. (Eds.), The Changing English Language: Psycholinguistic Perspectives. Cambridge University Press, Cambridge, pp. 191–211.

Matthews, D., Lieven, E., Tomasello, M., 2010. What's in a Manner of Speaking? Children's Sensitivity to Partner-Specific Referential Precedents. Dev. Psychol. 46, 749–760.

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

Neels, J., 2020a. Lifespan change in grammaticalization as frequency-sensitive automation: William Faulkner and the let alone construction. Cognitive Linguistics 31, 339–365.

Neels, J., 2020b. Refining frequency-effect explanations of grammaticalization. Universität Leipzig, Leipzig.

Nichols, J., 2012. Monogenesis or polygenesis: A single ancestral language for all humanity? In: Tallerman, M., Gibson, K.R. (Eds.), The Oxford Handbook of Language Evolution. Oxford University Press, Oxford, pp. 558–573.

Nölle, J., Fusaroli, R., Mills, G.J., Tylén, K., 2020. Language as shaped by the environment: linguistic construal in a collaborative spatial task. Palgrave Commun. 6, 1–10.

Oben, B., Brône, G., 2016. Explaining interactive alignment: A multimodal and multifactorial account. J. Pragmat. 104, 32–51.

Oben, B., 2018. Gaze as a predictor for lexical and gestural alignment. In: Brône, G., Oben, B. (Eds.), Eye-tracking in interaction. John Benjamins, Amsterdam Philadelphia, pp. 233–264.

Petré, P., Anthonissen, L., 2020. Individuality in complex systems: A constructionist approach. Cognitive Linguist. 31, 185–212.

Petré, P., Van de Velde, F., 2018. The real-time dynamics of the individual and the community in grammaticalization. Language 94, 867–901.

Pickering, M.J., Garrod, S., 2005. Establishing and using routines during dialogue: Implications for psychology and linguistics. In: Cutler, A. (Ed.), Twenty-first Century Psycholinguistics: Four Cornerstones. Lawrence Erlbaum Associates, London, pp. 85–101.

Pickering, M.J., Garrod, S., 2017. Priming and Language Change. In: Hundt, M., Mollin, S., Pfenninger, S.E. (Eds.), The Changing English Language: Psycholinguistic Perspectives. Cambridge University Press, Cambridge, pp. 173–190.

Planer, R.J., Sterelny, K., 2021. From Signal to Symbol: The Evolution of Language. MIT Press, Cambridge, MA.

Pleyer, M., 2017. Protolanguage and Mechanisms of Meaning Construal in Interaction. Lang. Sci. 63, 69–90.

Pleyer, M., Hartmann, S., 2019. Constructing a consensus on language evolution? Convergences and differences between biolinguistic and usage-based approaches. Front. Psychol.: Language Sci. 10, 2537.

Pleyer, M., Hartmann, S., 2020. Construction grammar for monkeys? Animal communication and its implications for language evolution in the light of usage-based linguistic theory. Evol. Linguist. Theory 2, 153–194.

Pleyer, M., Lindner, N., 2014. Constructions, Construal and Cooperation in the Evolution of Language. In: Cartmill, E.A., Roberts, S., Lyn, H., Cornish, H. (Eds.), The Evolution of Language. Proceedings of the 10th Conference. World Scientific, Singapore, pp. 244–251.

Pleyer, M., Pleyer, M. 2022. An Evolutionary Perspective on Im/politeness and the Pragmatic Concept of Face. In: Ravignani, A. et al. (Eds.): Proceedings of the Joint Conference on Language Evolution (JCoLE). https://doi.org/10.17617/2.3398549.

Pleyer, M., Winters, J., 2014. Integrating Cognitive Linguistics and Language Evolution Research. Theoria et Historia Scientiarium 11, 19–44.

Progovac, Ljiljana., 2019. *A critical introduction to language evolution: Current controversies and future prospects*. Springer, Berlin.

Rasenberg, M., Özyürek, A., Dingemanse, M., 2020. Alignment in multimodal interaction: An integrative framework. Cognit. Sci. 44, e12911.

Roberts, M.H., 1944. The Science of Idiom: A Method of Inquiry into the Cognitive Design of Language. Proc. Modern Language Assoc. 59, 291–306.

Roberts, G., Sneller, B., 2020. Empirical foundations for an integrated study of language evolution. Language Dynamics Change 10, 188–229.

Scerri, E.M., Thomas, M.G., Manica, A., Gunz, P., Stock, J.T., Stringer, C., Grove, M., Groucutt, H.S., Timmermann, A., Rightmire, G.P., 2018. Did our species evolve in subdivided populations across Africa, and why does it matter? Trends Ecol. Evol. 33, 582–594.

Schmid, H.-J., Günther, F., 2016. Toward a unified socio-cognitive framework for salience in language. Front. Psychol. 7, 1110.

Schmid, H.-J.r., 2016. Entrenchment and the psychology of language learning: how we reorganize and adapt linguistic knowledge, Language and the human lifespan. De Gruyter, Berlin, p. pages cm.

Schmid, H.-J.r., 2020. The Dynamics of the Linguistic System: Usage, Conventionalization, and Entrenchment. Oxford University Press, New York.

Scott-Phillips, T.C., 2015. Speaking our Minds: Why Human Communication is Different, and How Language Evolved to Make it Special. Palgrave Macmillan, Houndmills.

Steels, L., 2011. Modeling the cultural evolution of language. Phys. Life Rev. 8, 339–356.

Steen, F., Turner, M., 2013. Multimodal Construction Grammar. In: Borkent, M., Dancygier, B., Hinnell, J. (Eds.), Language and the Creative Mind. CSLI Publications, Stanford, CA, pp. 255–274.

Stefanowitsch, A., Flach, S., 2017. A corpus-based perspective on entrenchment. In: Schmid, H.-J.r. (Ed.), Entrenchment and the psychology of language learning: How we reorganize and adapt linguistic knowledge. De Gruyter, Berlin, pp. 101–127.

Sterelny, K., 2016. Cumulative cultural evolution and the origins of language. Biol. Theory 11, 173–186.

Su, D., 2016. Grammar emerges through reuse and modification of prior utterances. Discourse Stud. 18, 330–353.

Tagliamonte, S.A., 2016. Teen Talk: The Language of Adolescents. Cambridge University Press, Cambridge.

Tallerman, M., 2012. Protolanguage. In: Tallerman, M., Gibson, K.R. (Eds.), The Oxford Handbook of Language Evolution. Oxford University Press, Oxford, pp. 479–491.

Tannen, D., 2010. He said, she said. Sci. Am. Mind 21, 54–59.

Thompson, S.A., Couper-Kuhlen, E., 2005. The clause as a locus of grammar and interaction. Discourse Stud. 7, 481–505.

Tomasello, M., 1999. The Cultural Origins of Human Cognition. Harvard University Press, Cambridge, MA.

Tomasello, M., 2003. Constructing A Language: A Usage-Based Theory of Language Acquisition. MIT Press, Cambridge, MA.

Tomasello, M., 2008. Origins of Human Communication. MIT Press, Cambridge, MA.

Tomasello, M., 2014. A Natural History of Human Thinking. Harvard University Press, Cambridge, MA.

Tomasello, M., 2019. Becoming Human: A Theory of Ontogeny. Harvard University Press, Harvard, MA.

Traugott, E.C., Trousdale, G., 2013. Constructionalization and constructional changes, First. Oxford University Press, Oxford.

Van Lancker-Sidtis, D., Rallon, G., 2004. Tracking the incidence of formulaic expressions in everyday speech: Methods for classification and verification. Lang. Commun. 24, 207–240.

van Rooij, I., Baggio, G., 2021. Theory before the test: How to build high-verisimilitude explanatory theories in psychological science 1745691620970604 Perspect. Psychol. Sci..

Vigliocco, G., Perniss, P., Vinson, D., 2014. Language as a multimodal phenomenon: implications for language learning, processing and evolution. Philosoph. Trans. Roy. Soc. B 369, 20130292.

Wittgenstein, L., 1953. Philosophical Investigations. Transl. by G.E.M. Anscombe, Macmillan, London.

Wray, A., 2002. Formulaic Language and the Lexicon. Cambridge University Press, Cambridge.

Zima, E., 2013. Cognitive Grammar and Dialogic Syntax: Exploring potential synergies. Rev. Cognitive Linguist. 11, 36–72.

Zima, E., Brône, G., 2015. Cognitive Linguistics and interactional discourse: time to enter into dialogue. Lang. Cogn. 7, 485–498.

Zlatev, J., 2019. Mimesis theory, learning, and polysemiotic communication. In: Peters, M.A. (Ed.), Encyclopedia of Educational Philosophy and Theory. Springer, Singapore, pp. 1–6.

Zuidema, W., French, R.M., Alhama, R.G., Ellis, K., O'Donnell, T.J., Sainburg, T., Gentner, T.Q., 2020. Five ways in which computational modeling can help advance cognitive science: Lessons from artificial grammar learning. Top. Cogn. Sci. 12, 925–941.