



## OPEN ACCESS

## EDITED BY

Ana Mineiro,  
Catholic University of  
Portugal, Portugal

## REVIEWED BY

Alessandra Chiera,  
Roma Tre University, Italy  
Emiliano Loria,  
University of Eastern Piedmont, Italy

## \*CORRESPONDENCE

Peter Gärdenfors  
Peter.Gardenfors@lucs.lu.se

## SPECIALTY SECTION

This article was submitted to  
Language Sciences,  
a section of the journal  
Frontiers in Communication

RECEIVED 15 June 2022

ACCEPTED 14 November 2022

PUBLISHED 05 December 2022

## CITATION

Gärdenfors P (2022) Teaching as  
evolutionary precursor to language.  
*Front. Commun.* 7:970069.  
doi: 10.3389/fcomm.2022.970069

## COPYRIGHT

© 2022 Gärdenfors. This is an  
open-access article distributed under  
the terms of the [Creative Commons  
Attribution License \(CC BY\)](#). The use,  
distribution or reproduction in other  
forums is permitted, provided the  
original author(s) and the copyright  
owner(s) are credited and that the  
original publication in this journal is  
cited, in accordance with accepted  
academic practice. No use, distribution  
or reproduction is permitted which  
does not comply with these terms.

# Teaching as evolutionary precursor to language

Peter Gärdenfors<sup>1,2\*</sup>

<sup>1</sup>Department of Philosophy and Cognitive Science, Lund University, Lund, Sweden, <sup>2</sup>Paleo-Research Institute, University of Johannesburg, Johannesburg, South Africa

The central thesis of this article is that the evolution of teaching is one of the main factors that lead to increasingly complex communicative systems in the hominin species. Following earlier analyses of the evolution of teaching, the following steps are identified: (i) evaluative feedback, (ii) drawing attention, (iii) demonstration and pantomime, (iv) communicating concepts, (v) explaining relations between concepts, and (vi) narrating. For each of these step the communicative and cognitive demands will be analyzed. The focus will be on demonstration and pantomime, since these seem to be the evolutionarily earliest unique human capacities. An important step is the transition from pantomime for teaching to pantomime for informing and how this in turn leads to communicating concepts. As regards explaining relations between concepts, the focus will be of the role of generics in teaching and communication. Analyzing these topics involves combining cognitive science with evolutionary theory, archaeology and theories of communication. Two factors are important as a background: (i) the evolution of prospective planning, that is, planning for future goals, and (ii) the evolution of a theory mind. These capacities are central in explaining how more advanced forms of teaching, communication and cooperation emerged along the hominin line.

## KEYWORDS

evolution of language, pantomime, generics, teaching, gesture, concept relations

## Introduction

The central thesis of this article is that the evolution of teaching is central for the evolution of increasingly complex communicative systems in the hominin species. Building on analyses of the evolution of teaching (D'Errico and Banks, 2015; Kline, 2015; Gärdenfors and Högberg, 2017, 2021), the focus will be on the roles of demonstration and pantomime as stepping stones to forms of communication that go beyond what is found in non-human animals.

## Background

There exist several theories that attempt to explain the evolution of language. It is, however, not sufficient to propose an evolutionary account of the benefits of linguistic communication. The question is not just why the hominins evolved language, but, equally importantly, why chimpanzees and other apes did *not*. Since we share a relatively recent common ancestor, the task is to show that language provided adaptive benefits for

proto-humans but not for proto-chimpanzees. Many theories concerning the evolution of language fail this “chimp test” (Bickerton, 2002, 2009; Gärdenfors, 2004; Johansson, 2005; Dessalles, 2007). For example, Bickerton (2007, p. 514) argues that many species are social, so social dynamics is not enough to explain why language is unique to humans. He therefore rejects Dunbar’s (1996) thesis that language replaces grooming in the larger hominin groups.

When explaining why language is unique to the hominin line, one must, ultimately, rely on the differences in the *ecology* of the early hominins and how it differs from that of proto-chimpanzees. One such ecological factor is that the hominins gradually changed their habitat to more open landscapes. As a consequence, the hominins came to use a wider range of foods than other apes. The availability of these food sources was more transient and scattered in the landscape than the mainly vegetarian food that was collected by the other apes.

Analyzing these topics involves combining cognitive science with evolutionary theory, archaeology and theories of communication. In earlier work, I have argued that the ecological factors led the hominins to evolve more advanced *prospective planning*, that is, to plan for future needs, and to more advanced forms of *cooperation* (Gärdenfors, 2003, 2013; Osvath and Gärdenfors, 2005; Gärdenfors and Osvath, 2010; Gärdenfors et al., 2012; Geurts, 2020). Cooperating about future goals presumes that one is able to refer to non-present objects and actions—what is sometimes called the displacement (Hockett, 1960) or detachment (Gärdenfors, 1995) of communication. Furthermore, advanced forms of cooperation depend on a more developed theory of mind (ToM) (Tomasello, 1999; Gärdenfors, 2004, 2007). The evolutionary pressure generated from the need for cooperation has thus been one factor that drove the evolution of language (Brinck and Gärdenfors, 2003; Gärdenfors, 2004; Gärdenfors et al., 2012; Sterelny, 2014). A more advanced ToM is also required for intentional teaching.

Then the question becomes how the hominin communication systems evolved from a signaling system that we find in other animals to a system of signs—gestures or words—that makes it possible to cooperate about the future. I shall argue that the evolution of teaching, which can be seen as a form of cooperation, can be used to explain many aspects of the emergence of advanced communication. I will not argue, however, that teaching is the sole evolutionary root of language.

Only humans have natural pedagogy, that is, the capacity to transmit cultural and other forms of knowledge through communication (Csibra and Gergely, 2006). To find a clue to what aspects of hominin life have led to this capacity, two other aspects of the adaptation to foraging in open landscapes should be considered: (i) the manufacture and use of tools that are more advanced than those used by other apes; and (ii) the development of food preparation techniques, in particular cooked food (Wrangham, 2009). In order to

be reliably transmitted between generations, these cultural expansions depend on *teaching*.

When comparing different forms of teaching and different forms of communication I apply the following methodological principle:

*Cognitive parsimony*: If the cognitive capacities required for an activity or technique A are included in those required for an activity or technique B, then, A is evolutionarily prior to B (unless the additional capacities required for B are evidenced synchronously or earlier than A)<sup>1</sup>.

Even though this principle does not make any dating possible, it is at least possible to provide evidence that an activity or a technique is evolutionarily older than another.

I shall first present a summary of the steps in the evolution of teaching proposed by Gärdenfors and Högberg (2017, 2021). I agree with Sterelny (2012, p. 2014) who writes that “If language (or protolanguage) evolved as a system of gesture, the evolution of elaborated manual skill and the evolution of gestural communication would support each other”. Then I analyze demonstration and pantomime and compare them, arguing that cognitive parsimony indicates that demonstration is evolutionarily older than pantomime. I then provide arguments indicating that pedagogical uses of pantomime are evolutionarily prior to the more communicative uses. The teaching of concepts and their relations require further developments of expressive capacities. Finally, some ideas on how pantomime and gesture evolved into narratives are outlined.

## The evolution of teaching

Only among humans is teaching intentional, socially structured, and linguistically mediated. Compared to individual learning or trial-and-error activities, the cost of teaching is low in relation to its benefits. Several studies underline intentional teaching as central for cultural transmission when learning complex, cognitively opaque skills such as the making of stone tools (Stout and Chaminade, 2012; D’Errico and Banks, 2015; Kline, 2015; Gärdenfors and Högberg, 2017).

## Imitation, emulation, and facilitation

Before analyzing different forms of intentional teaching, I present three activities that are related to teaching. The first two come from Tomasello’s (1999) distinction between learning by *emulation*, where the learner tries to reach the same outcome as that by the model actions, and learning by *imitation*, where the learner tries to perform the same actions as those of the model (see also Tehrani and Riede, 2008; Caruana et al., 2013).

<sup>1</sup> Progovac (2016, p. 3) proposes a similar principle for the evolution of syntactic structures.

The “artificial fruit” experiments by [Whiten et al. \(2005\)](#) investigated the differences between the two types of learning. At the beginning, it seemed that chimpanzees emulate while children imitate. More recent studies, however, indicate that the learning of the chimpanzees situation is more complex ([Whiten et al., 2009](#)). Imitating familiar motor actions in novel situations seem to be easier for chimpanzees than copying new motor actions ([Myowa-Yamakoshi and Matsuzawa, 1999](#)). A difference may be that chimpanzees as well as children often perceive the *intention* of the model and therefore emulate in such situations ([Froese and Leavens, 2014](#)).

Thirdly, an individual *facilitates* the learning of another by scaffolding the environment so that the learning individual learns faster than it would otherwise have done. An example is meerkats that bring back scorpions, often disabled by removal of their sting ([Caro and Hauser, 1992](#)). As the pups grow older, the meerkat adults give the pups increasingly intact prey. However, the adults do not gauge the competence of the pup, but only rely on changes in pup begging calls ([Thornton and McAuliffe, 2006](#); [Hoppitt et al., 2008](#)). This suggests that the adults cannot evaluate the competence of the pups and that their behavior is non-intentional.

## Levels of teaching

I next present a summary of the six levels of teaching analyzed in [Gärdenfors and Högberg \(2017, 2021\)](#) (for related classifications of teaching, see [D’Errico and Banks, 2015](#); [Kline, 2015](#)). The ordering of the levels reflects the degree of ToM that is required as well as the increase in the complexity of the communication between teacher and learner.

(1) *Evaluative feedback*: The teacher approves or disapproves of the learner’s behavior ([Castro and Toro, 2004](#)).

Animal data on this form of teaching include chimpanzee mothers, who take away poisonous food from infants, and gorilla, chimpanzee and macaque mothers, who facilitate and encourage infants’ locomotion ([Maestripieri, 1995, 1996](#); [Whiten, 1999](#)). The teaching is typically performed by approving or disapproving signals, but is sometimes non-intentional.

(2) *Drawing attention*: Here, the teacher’s intention is that learner focuses on a particular object, feature or action.

Among humans, drawing attention is often achieved by declarative pointing ([Bates et al., 1975](#); [Brinck, 2004b](#); [Tomasello, 2009](#)) which means that the teacher intends to communicate. Other methods than pointing can also be used. For example, human infants draw attention to an object by showing it even earlier than they point ([Bates et al., 1975](#)). Among non-human animals, attention is often achieved *via* alarm calls. In most cases, however, such signals are presumably non-intentional and not dependent on the knowledge of conspecifics (although see [Crockford et al., 2012](#) for an intentional case).

For both these levels of teaching, one finds non-intentional as well as intentional forms. They do not require any form of ToM, neither on part of the teacher, nor on the learner.

(3) *Demonstration and pantomime*: These forms of teaching involve intentionally showing somebody else how to perform a task or how to solve a problem.

This level seems to be where humans depart from other animals<sup>2</sup>. Demonstrating builds on ToM since it presumes that the teacher understands the lack of knowledge in the learner and that the learner perceives that something can be learnt. Demonstration also requires that the teacher and the learner *jointly attend* to the teacher’s actions ([Bruner, 1995](#); [Tomasello, 1999](#); [Gärdenfors, 2017](#)). In pantomime the intended action is not really performed but just represented.

The abilities to demonstrate and pantomime constitute a breakthrough in the teaching and transmission of culture among hominins. [Stout \(2018, p. 257\)](#) argues that “the emergence of a human *technological niche* increasingly reliant on cooperation sharing, and the intergenerational reproduction of complex subsistence skills” was particularly important [see also the “technological pedagogy hypothesis” in [Stout and Chaminade \(2012\)](#)].

(4) *Communicating concepts*: Unlike animal calls, this involves communicating something that is general about a category and checking that the learner understands the extension of the concept.

Teaching how to categorize plants or animals are central examples of how cultural knowledge is transmitted. Among modern humans, the main method to teach a concept is to use a word (or a gesture) which refers to the concept together with some technique, such as pointing, for drawing the attention to examples of what the concept covers. This involves highlighting characteristic properties. For example, learning that champignons have brown or pink lamellae makes it possible to separate champignons from the deadly poisonous destroying angel mushroom.

Concept teaching also involves ToM since it requires that the learner understands that the teacher is intentionally using a word, sound, or gesture as a communicative sign, that is, that the word, sound, or gesture is used to “stand for” something else ([Zlatev et al., 2005](#)). Purely symbolic communication is, however, not required for this level, although it makes communication about concepts more efficient.

<sup>2</sup> An observation of a chimpanzee showing somebody else how to perform an action is a mother who shows her infant how to hold a stone in order to crack a nut ([Boesch, 1991](#)). This behavior is, in my opinion, better seen as facilitating. Another possible example about demonstration in elephants is presented in [Bates et al. \(2010\)](#). [Russon and Andrews \(2011\)](#) data on spontaneous demonstration in orangutans will be discussed below.

(5) *Explaining relationships between concepts*: The teacher's intention in explaining is typically that the learner understands the semantic or causal relationship between two concepts.

Examples of causal relationships are “if you eat this kind of berry, you will become very sick” and “a track that looks like this is made by an eland antelope”. In Section Narrating, I argue that *generics* fulfill this function. Unlike the previous ones, this level presumes that communication is *detached*, that is, it refers to things not perceivable in the communicative environment.

(6) *Narrating*: This involves chaining descriptions of events into a causally coherent whole.

In all human cultures, narration plays a central role (Barnard, 2011; Ferretti et al., 2017). Typical uses of narratives are gossip and entertainment and such narratives cannot be seen as direct intentional teaching. Narratives often have a moral, however, that is intentionally conveyed and hence it can be seen as a type of teaching.

It is important to note that narratives need not involve spoken language. Pantomiming is sufficient for narration (Sibierska, 2017). However, what distinguishes narration is that it represents a globally coherent sequence of *events* (Ferretti et al., 2017). Thus, it presupposes a well-developed event cognition (Radvansky and Zacks, 2014) which in turn builds on advanced causal thinking (Gärdenfors, 2021).

I next turn to a more detailed analysis of levels (3)—(6) since my thesis is that these forms of teaching have contributed to the evolution of language.

## Demonstration

Demonstration and pantomime form the first level of teaching where such teaching activities seem to be unique to humans and it is therefore natural to focus on this level. This and the following section present analyses of demonstration and pantomime, respectively. My aim is to show that they have been evolutionary precursors to later forms of communication and, in the end, of language. In brief, “showing how” is a stepping-stone for advanced forms of communication.

### The structure of demonstration

Demonstrating can be defined as that a “teacher” intentionally shows a “learner” how to perform a task. Demonstration is central in so called “natural pedagogy” and it can be found all human societies (Csibra and Gergely, 2009, 2011). Showing a child how to tie shoelaces, how to ride a bike, and how to use a smartphone are well-known everyday examples.

When the teacher demonstrates how to a certain set of actions should be performed, for example knapping a Levallois flake or making a fire, I have identified the following characteristic criteria (Gärdenfors, 2017):

- (a) The demonstrator *actually performs* the actions involved in the task.
- (b) The demonstrator makes sure that the learner *attends* to the series of actions.
- (c) The demonstrator's *intention* is that the learner can perceive the right actions in the correct sequence.
- (d) The demonstrator *exaggerates* and slows down some parts of the actions in order to facilitate for the learner to perceive important features.

In comparing demonstration with non-human animal communication, it is easy to find distinctive differences (Waciewicz and Zywczyński, 2021). Demonstration is:

- (e) *Voluntary*, in the sense that the action is deliberately performed. Donald (2012) says that the action is autocued.
- (f) *Intentional*, since the teacher wants that the learner imitates what is demonstrated.
- (g) *Honest*, since there exists no reason to deceive a learner.
- (h) *Directed* to one or a few individuals<sup>3</sup>.

Unlike ordinary action, demonstration also satisfies the following criterion:

- (i) Demonstration is both an action and a *representation* since the motions involved in the act *correspond* to some action (maybe involving objects), namely the action the teacher wants the learner to perform, but at the same time the demonstration differentiated from it by the teacher and the learner.

Therefore, the teacher's actions “point to” what the learner is meant to copy afterwards.

Often the learner gives some form of *feedback*, indicating that he or she has grasped the teacher's intention and what is being demonstrated. When the learner tries to perform the desired actions, the teacher approves or disapproves (level 1 of teaching). A demonstration is sometimes accompanied by verbal comments, but this is often not necessary.

In connection with criterion (b), Csibra and Gergely (2006, p. 149) note that “human communication is often preceded, or accompanied, by ostensive signals that (i) disambiguate that the subsequent action (for example, a tool-use demonstration) is intended to be communicative and (ii) specify the addressee to whom the communication is addressed”. Gergely et al. (2007) have shown experimentally that the ostensive nature of the teacher's actions is central to demonstration.

A consequence of criteria (b) and (c) is that demonstrating presumes advanced ToM for both the teacher and the learner.

<sup>3</sup> Some forms of animal communication also have some (but not all) of these properties.

The best way to satisfy (b) is that the teacher and the learner jointly attend to what is demonstrated. (c) presumes that the teacher realizes that the learner lacks relevant knowledge and also that the learner understands the teacher's intention.

Demonstration is, however, not restricted to direct teaching but is also applicable in other situations. One important case is when a learner can demonstrate to a teacher that she has mastered the actions of a task. For example, a child can show that she now knows how to tie her shoelaces. Another example, typically occurring in a legal case, is that some forms of narratives can be expressed by a witness demonstrating how somebody behaved (the borderline between demonstration and pantomime may not be sharp in such cases).

As regards (d), note that demonstration involves learning by imitation rather than by emulation. The goal of the demonstration is not the most important aspect, but rather the actions which lead to it. By emphasizing the initial and final states of an action, the teacher assists the learner when segmenting the action sequence. Apart from establishing joint attention, the highlighting adds to the ostensive behavior that a teacher uses to make the learner attend to the demonstration (see Gergely et al., 2007).

Furthermore, demonstration can be used in autocued rehearsal (think of a dancer in front of a mirror). Here Donald's (2012) *mimesis hypothesis* is applicable as a starting point. It states that "the ability to produce conscious, self-initiated, representational acts that are intentional but not linguistic" (Donald, 1991, p. 168) mediated between those of the our common ancestor with the apes and modern humans. He claims that mimesis is a main adaptation since it improved tool production. It is clear that demonstration presumes the capability of mimesis.

## Archaeological evidence

From an archaeological point of view, one may ask when in the hominin history one can find some evidence for demonstration. Gärdenfors and Högberg (2017) proposes an answer, by arguing that already the teaching of the techniques used to make Oldowan tools depend on demonstration (see also Morgan et al., 2015, p. 5). Demonstration would then be at least 2.5 million years old.

The argument is based on so-called *core maintenance*. This technique is mastered by knapping flakes from the core so that it will be possible to continuing striking further flakes. Delagnes and Roche (2005) show experimentally that core maintenance depends on planning. For an apprentice to learn core maintenance, a teacher must demonstrate how a flake can be detached in a way so that the knapping of another flake will be possible, which in turn should make it possible for the next flake to be detached, etc. For this purpose, the teacher must demonstrate (or pantomime) (i) how to best hold the core, (ii) the best striking angle, and (iii) the

correct movement of the arm and hand holding the hammer stone when knapping. The learner should then practice, often for an extended period. In support of this position, Stout (2018, p. 260) writes: "For Oldowan knapping, the salient demands are at the level of elementary movements, perhaps comparable to the articulatory/phonological level of speech (or sign) processing. Consistent with this, Oldowan knapping by modern human subjects recruits portions of ventral premotor cortex [...] neighboring those involved in speech production and perception".

Wynn et al. (2011) have argued that the techniques used by hominins in making Oldowan tools can also be achieved by apes. In support of the claim they refer to the knapping behavior of Kanzi and Panbanisha, two bonobos who have been trained to knap by humans. However, they never reached the skill level of Oldowan knappers (Toth et al., 1993). A reason for this is that the bonobos did not rehearse the techniques that had been shown to them, but merely engaged in knapping when encouraged by the teachers, or when a reward was given. Crucially, their knapping showed no signs of core maintenance.

## Pantomime

In current research concerning the origins of human communication, *pantomime* is proposed as a step in the evolution of symbolic language (Arbib, 2012, 2018; Gärdenfors, 2017, 2020; Abramova, 2018; Zywczyński et al., 2018; Brown et al., 2019). Pantomime is often viewed as a form of communicative gesturing. One definition of pantomime comes from Arbib (2012, Ch. 8) who writes that it "involves expressing a situation, object, action, character, or emotion without words, and using only gestures, and other movements". Brown et al. (2019) defines it as follows: "Pantomime refers to iconic gesturing that is done for communicative purposes in the absence of speech". Zywczyński et al. (2018) provide a broad presentation of different uses of the concept, concluding with the following characterization: "[W]e take pantomime to be a non-verbal, mimetic and non-conventionalized means of communication, which is executed primarily in the visual channel by coordinated movements of the whole body, but which may incorporate other semiotic resources, most importantly non-linguistic vocalizations. Pantomimes are acts of improvised communication that holistically refer to a potentially unlimited repertoire of events, or sequences of events, displaced from the here and now."

In contrast to the communicative function highlighted in these definitions, my aim in this section is to argue that pantomime has developed out of demonstration (Gärdenfors, 2017, 2020) so that pantomime emerges from its use in teaching. This position is in agreement with Sterelny (2012, p. 2146) who writes that "once we have the capacity to demonstrate and to practice, it is available for the stimulus independent production of iconic representations". My definition of pantomime goes as

follows: *an intentional pattern of movements of the body or parts of it (voice not excluded), the intention of which is to represent an action or a sequence of actions*. Notice that my focus is on representing actions, which means that not all forms of gesturing count as pantomime. Pantomime is often restricted to action involving the whole body (e.g., Gullberg, 1998, p. 97), but I take a broader perspective and allow that only parts of the body are used in a pantomime. I also allow non-verbal vocalizations to be parts of a pantomime (Zywiczynski et al., 2018)<sup>4</sup>. The advantage of vocal imitation or iconicity might have been a selective force in driving the control of the vocal apparatus (Sterelny, 2012). On my view, pantomime is the most central form of the mimetic ability that Donald (1991, 2001, 2012) considers to be crucial in the evolution of the human mind.

## The structure of pantomime

Pantomime is similar to demonstration in that it fulfills the criteria above except for (a) and, in the communicative case (g). The crucial difference is that (a) is replaced by the following, since in pantomime the actual actions are not performed, but only simplified versions of them:

(a') The mimer performs the *movements* of the actions in the task without actually performing the actions.

In many teaching situations, the teacher cannot perform the action that the learner is supposed to perform because then the learning opportunity is foregone. For example, teaching somebody how to knap a Levallois flake when only one core is available cannot be made by demonstration because once the flake is made the earlier state of the core cannot be reproduced.

It is cognitively more challenging to grasp the intention of a pantomime than to understand a demonstration. The role of a demonstration is evident as soon as the learner understands that it functions as a form of teaching. In contrast, the learner must also understand that the teacher intends the pantomime to *stand for* a real action and that the teacher wants the learner to understand this. Unlike demonstration, pantomime is not primarily an action, but a representation of an action. Pantomime therefore fulfills the following criterion (Zlatev et al., 2005):

*Communicative sign function:* The agent intends for the act to stand for some action, object or event for an addressee, and for the addressee to realize that the act is a representation.

<sup>4</sup> Cf. Orwell (1948/1968, p. 11): "Primitive man, before he had words, would rely upon gesture, and like any other animal he would cry out at the moment of gesticulating, in order to attract attention. Now one instinctively makes the gesture that is appropriate to one's meaning, and all the parts of the body follow suit, including the tongue". And Armstrong and Wilcox (2007, p. 68) write that "there never was a time when visible gestures were unaccompanied by vocalizations".

The communicative sign function involves a second order intention that the agent intends that the addressee understand the communicative intentions (Gårdenfors, 2003, section 6.3; Bar-On, 2013; Moore, 2018). In line with this, Arbib (2012, p. 217–218) writes: "Where imitation is the generic attempt to reproduce movements performed by another, whether to master a skill or simply as part of a social interaction, pantomime is performed with the intention of getting the observer to think of a specific action or event." Similarly, Mittelberg (2019) claims that pantomimes are inherently metonymic and they therefore create associations to already available knowledge about actions.

Beyond invitations to imitate, demonstration and pantomime may also help the learner to perceive new *effectivities*, that is, new actions that the learner can perform (Zukow-Goldring and Arbib, 2007). For example, when a bringing a new toy to a child, a parent often demonstrates how the toy functions and how it can be used in play. Zukow-Goldring and Arbib (2007) calls such processes "assisted imitation". A more archaeologically directed example is that if you have a stone knife that you know how to use to cut meat, someone can show you how to use it to open mussels.

A central feature of pantomime is that it is open-ended. The autocuing that led to the skill for pantomiming provided a rich source of gestures that could form the basis for a great variety of meanings (see also Stout, 2018, p. 262). In line with this, Arbib (2012, p. 217) writes:

"[B]uilding on the skill for complex imitation, pantomime provided the breakthrough from having just a few gestures to the ability to communicate freely about a huge variety of situations, actions, and objects. Where imitation is the generic attempt to reproduce movements performed by another, whether to master a skill or simply as part of a social interaction, pantomime is performed with the intention of getting the observer to think of a specific action or event."

Pantomime is a form of *pretense*. The pantomimes that are part of pretense play are enactments of a story that is created by the players during their interactions. In this way, play stories are participatory rather than detached narratives that are separated from the play activities.

Pretense is maybe the most basic form of implementing the communicative sign function. The reason for this is that one uses two representations of the same object or action when pretending—one's perception of the object or action and an imagined version of it (Leslie, 1987). For example, when a boy pretends that a red box is a fire engine, he knows that it is a box but at the same time he "sees" it as a fire engine that he can interact with. By *suppressing* his perception, he can use his fantasy instead. His image is a false representation of his perception of the real thing. Leslie (1987) claims that such imagined objects or actions are necessary to be able to pretend.

He writes that small children's pretense "is an early symptom of the human mind's ability to characterize and manipulate its own attitudes to information. [...] In short, pretense is an early manifestation of what has been called a *theory of mind*" (Leslie, 1987, p. 416).

In contrast, demonstration does not build on the double worlds that are used in pantomime. Pantomime can therefore be seen as using pretense to extend demonstration. For this reason, demonstration is less cognitively demanding to perform. According to the principle of cognitive parsimony, it follows that demonstration should appear evolutionarily earlier than pantomime—at least in teaching contexts.

## From pantomime for instruction to pantomime for communication

The previous subsection argued that the evolutionary origin of pantomime is used in teaching as a development of demonstration<sup>5</sup>. However, pantomime is also used for other directly *communicative* aims, for example, in describing a plan, in narrating (Sibierska, 2017) or as part of telling a joke. In this section, I argue for the thesis that pantomime for communication is an exaptation of pantomime for instruction, and thus that the teaching function is evolutionarily primary<sup>6</sup>.

Haiman (2018, p. 46) makes a distinction between (dramatic) discourse which *shows* what is communicated and indirect (narrative) discourse, which *tells* about it [the distinction was originally introduced by Rimmôn-Qênān (1983)]. He makes the distinction in the context of the use of ideophones, but it also applies to demonstration and pantomime. The distinction between showing and telling emerges when determining what is the *intention* of a pantomime. Firstly, I can pantomime an *action* that I want you to *copy*. This is a form of showing and it is typical in a teaching situation. Secondly, I can pantomime an action as part of a *message* (request, command, warning, narrative, play element, etc.). Gesture researchers have focused on the second use of pantomime (communicative act). A demonstration is indeed a gesture, but it only conveys the first type of intention—demonstrating is showing. However, when pantomime is used for narration, it is also telling.

When showing somebody how to perform something, you cannot be dishonest. However, when telling about something, you can. Hence demonstration and pantomime for teaching

are honest forms of communication, while pantomime for communication need not be.

Another important difference between demonstration and pantomime for instruction is that pantomime for communication is often *detached* (Zywiczynski et al., 2018, Sect. 3.8). I can, for example, pantomime how to set up the spring trap that I want you to construct tomorrow. In narrating, the reference of a pantomime can, of course, be far detached in space and time, and even refer to imaginary events. Using pantomime for communication therefore entails a broadening of the mental horizons in space and time, and hence requires more extended cognitive capacities.

Yet another difference is that pantomiming functions as an instruction when it is used for teaching, while pantomiming narratively is part of forming a common ground (Clark, 1992; see also Tomasello, 2009, p. 67). Hence, when pantomime is used for teaching, it is *proto-imperative*, while when a pantomime is used communicatively, it is a *proto-declarative*.

Brown et al. (2019) distinguish between *egocentric* and *allocentric* pantomimes<sup>7</sup>. In egocentric pantomime, the reference is determined in relation to the body of the miming person. When a pantomime is used in teaching, the agent that is supposed to perform the action that is pantomimed need not be specified since this is implicitly assumed to be the learner. In contrast, when pantomime is used in its communicative function, it is often allocentric. For example, you can pantomime a fight between two persons that you witnessed in the street. Then it must be specified *who* is the agent of the action and sometimes also *what* object is acted on or *where* the action is performed (Brown et al., 2019). Detaching the actions from the mimer makes the pantomime more difficult to interpret for the recipient. One way to achieve this is to add signs (gestures or sounds) that refer to the agent, the object and the location.

Representations of events typically contain information about causes and effects. In demonstration both cause (the action) and effect (the result) are manifest, while in pantomime for instruction only the cause is included. If the effect needs to be expressed, another gesture is normally required. For example, if I pantomime how to turn the key in the lock in a box, I must also show that the lid then springs up automatically. And, in pantomime for communication, the actor need not be the agent of the cause. In that case the agent must be specified by other means in order to determine who performs the depicted action.

From an evolutionary point of view, one may ask whether non-human animals, in particular apes, have the ability to pantomime. Some researchers are skeptical, for example Zuberbühler (2013, p. 136), who writes that "pantomiming is conspicuously absent, apart from isolated anecdotes". On

<sup>5</sup> My characterization of pantomime therefore differs from that of Brown et al. (2019) since they delimit it to "iconic gesturing that is done for communicative purposes in the absence of speech".

<sup>6</sup> In line with this thesis, Zywiczynski et al. (2017) distinguish between cognitive and communicative forms of pantomime.

<sup>7</sup> Brown et al. (2019) suggest that this distinction should replace the distinction between character viewpoint and observer viewpoint that has been used by MacNeill and others (Cassell and McNeill, 1990; McNeill, 1992; Cartmill and Goldin-Meadow, 2012).

the other hand, [Russon and Andrews \(2011\)](#) have presented evidence for pantomiming in orangutans. Their subjects lived in a rehabilitation camp and were used to communicate with humans. They conclude that “pantomime could have been within the grasp of the common human-great ape ancestor” (p. 316). Their notion of pantomime is broader than the one used here and includes cases of deception. Furthermore, most of the evidence they analyze agrees with [Gibson \(2013, p. 209\)](#) observation that apes only gesture about actions requested of the addressee. In line with this, 17 out of 18 examples in [Russon and Andrews \(2011\)](#) were classified as imperative (see e.g., [Boesch and Tomasello, 1998](#)). Hence these cases are pantomimes for instruction. Only one case—enacting a shared memory—is classified as declarative: this is the case of the female Kikan acting out how the person next to her had doctored her foot, when it was cut. This episode, however, may also be interpreted as an apprentice showing that she has learnt what to do. Hence, it is not unequivocal that it is a case of pantomime for communication (see also [Ferretti et al., 2017, p. 10](#) for a similar criticism).

[Table 1](#) summarizes the arguments of this section<sup>8</sup>. Pantomime for instruction is contrasted with, on the one hand, demonstration, and, on the other hand, pantomime for communication. All ten factors suggest that the cognitive abilities involved in pantomime for communication are at least as extensive as those for pantomime for instruction. The arguments presented here therefore support, *via* the principle of cognitive parsimony, the thesis that demonstration is evolutionarily prior to pantomime for instruction, which in turn is evolutionarily prior to pantomime for communication.

## Pantomime compared with other forms of gesturing

There exist many proposals for what characterizes gestures, some of which are very general. Most of them are not based on an evolutionary perspective<sup>9</sup>. [Kendon \(2004, p. 15\)](#) definition restricts gestures to “utterance uses”, that is, a communicative function, often performed together with speech. As I argue, pantomime has its primary uses in instructional contexts not connected to utterances. Therefore, a more inclusive definition should be sought for. Following [Warglien et al. \(2012, p. 23\)](#), I therefore consider as gestures “goal-directed communicative body movements, i.e., such that require interpretation from an audience for achieving the gesturer’s goal”<sup>10</sup>. The following quotation from [Kendon \(2017, p. 168\)](#) supports this definition: “In my view, this suggests that gestures can best be understood

as forms of action derived from how one uses one’s hand to show or change the shape of form of things—to pick things up, let them drop from one’s hands, place one’s hands around an object, grasp an object, do something with an object, carry our patterns of action. and so forth.”

Most researchers have treated pantomime as a form of gesture. [McNeill \(2013, p. 483\)](#) describes pantomime as gesture without speech. However, vocal sounds can be parts of a pantomime. Communication is, in its nature, multi-modal ([Gillespie-Lynch et al., 2014](#)). For example, I can, by modulating the pitch of my voice, pantomime an up-and-down movement ([Ekström et al., 2022](#)) or imitate the sound of an animal while pantomiming its movements.

A basic distinction can be made between *indexical* gestures, where the semantic ground is spatio-temporal contiguity (for example, pointing) and *iconic* gestures, where the ground is similarity (for example, pantomime).

My position that pantomimes characteristically express *actions* can be amplified by considering the semantic domains of different types of gestures. One can distinguish three kinds of representational gestures which correspond to three types of domains:

(i) *Location* or *direction*. This involves the domain of *physical space*, which is the characteristic referential domain for indexical gestures.

Indexical gestures are fundamental for drawing attention (level 2 of teaching). Pointing is the primary example, but directing attention can be done by other deictic gestures such as giving, showing (with object in hand), requesting (open palm), and by using gaze direction.

(ii) *Actions*. In previous work ([Gärdenfors, 2007, 2014a, Ch. 8](#); [Gärdenfors and Warglien, 2012](#)), actions have been represented as *patterns of forces*.

A pantomime can then be characterized as a gesture the meaning of which is based on the *force domain*. This entails that the represented action is iconically enacted. This analysis accords with [Kendon’s \(2004, p. 160\)](#), since he identifies pantomime with enactment that is oriented toward actions.

(iii) *Object properties*. Gestures can represent the shape, size, length, height, depth and maybe other properties of an object.

These properties each belong to an object category domain ([Gärdenfors, 2014a, Ch. 6](#)). [Brown et al. \(2019, p. 6\)](#) argue that tracing the outline of an object with a hand (or the hands) is a common way of describing an object with gestures.

It is, of course, difficult to say anything about the evolutionary order of these types of gestures. From a developmental viewpoint, however, pointing develops early in children, so they can then communicate about the spatial domain. As regards actions vs. objects, [Ortega et al. \(2017\)](#) note that signing children have a bias to interpret signs as actions. When children have two signs for the same concept to choose from, they initially prefer an action-based variant. This gives them the opportunity to link a label to schemas grounded in their action experiences.

<sup>8</sup> The table is an extension of the table in [Gärdenfors \(2020\)](#).

<sup>9</sup> See [Brinck \(2004a\)](#) for an analysis of the origins of pointing.

<sup>10</sup> Here I only consider *representational* gestures, so that, for example, beat and emblem gestures are excluded ([Kendon, 2004, Chs. 9–11](#)).



TABLE 1 Ten factors distinguishing demonstration, pantomime for instruction, and pantomime for communication.

	<i>Demonstration</i>	<i>Pantomime for instruction</i>	<i>Pantomime for communication</i>
Sign function	Marginally	Yes	Yes
Communicative form	Showing	Showing	Telling
Honest signal	Yes	Yes	Not necessarily
Detachment	None	Minimal	Varying
Pragmatic form	Imperative (instruction)	Imperative (instruction)	Declarative
Perspective	Egocentric	Egocentric	Often allocentric
Represented action	Demonstrator showing learner's action	Mimer representing learner's action	Mimer representing various roles
Representation of cause and effect	Both cause and effect represented	Only cause represented	Only cause represented actor may not be agent
Pretense	None	Limited	Extensive
Ape capability	Limited evidence	Limited evidence	No unequivocal evidence

## From pantomime to protolanguage

I have characterized pantomime as a way of depicting actions. The fact that the force patterns of a pantomime can simultaneously contain information about the properties of an object complicates the domain analysis presented above. For example, a gesture showing how a bottle is placed on a desk can combine the placing movement and a hand-shape indicating the shape of the object (Gullberg, 2011).

A more inclusive view is to see pantomime as a combination of gestures for actions with gestures for object properties (or objects) and places. Such a proposal can be found in Zywczyński et al. (2018, Section 3.7), who write that “pantomimic acts are ‘the size of’ propositions or utterances rather than smaller component units; rather than being elements of a larger communicative whole, they express complete, *self-contained* communicative acts”.

When pantomimes were exapted from their original function as a request to copy in a teaching situation to having a declarative function, as part of a narrative, or as an element in a play sequence, they formed the seeds for a larger set of conventionalized gestures. Arbib (2012, p. 224) writes that “[p]antomime is not itself part of proto-sign but rather a scaffolding for creating it”. Here proto-sign refers to a communication system that is conventional and combinatorial.

Dividing of gesture references into locations, object properties and actions, allows the three categories to be seen as *proto-demonstratives*, *proto-adjectives* and *proto-verbs*. I thus propose that the three types of semantic domains for gestures constitute the seeds for three of the main word classes<sup>11</sup>. Gestures for *proto-nouns* would, in general, develop by conventionalization of characteristic properties of objects, but

11 Indirect support for the role of proto-verbs comes from Aussems and Kita (2021) who show that seeing gestures that depict verb referents help children generalize verb meanings and learn more verbs from the same subcategory.

they can also develop out of verbs. A gesture for a property or a type of action that is characteristic for a person may by metonymy become a proto-name for the person<sup>12</sup>. For example, a gesture describing the fluffy hair or the limping of a person could come to serve as a reference to the person. Similarly, a gesture for a property or a type of action that is characteristic for a category of objects may, in the same way, become a proto-noun.

Once these proto-word classes are in use, it is natural that gestures are *combined* and thereby more complex messages can be communicated. This would be the first steps toward the compositional communication of a protosign (Arbib, 2012) or protolanguage. Note that the process outlined here deviates from Arbib's (2012) theory about holophrases. Here the pantomime is taken to be the original holophrase, which is then complemented with signs for further thematic roles that are part of the event that is communicated, rather than broken down into smaller elements.

## Communicating concepts and their relations

### Conventionalization

Several authors have noted that pantomime is detached. Pantomime is, however, not conventional. Corballis (2014, p. 190) proposes a next step in the evolutionary process:

“Pantomime, though, is inefficient, and over the course of the Pleistocene, the pressure toward a more efficient and compact system may have driven the process of conventionalization. Iconic or pantomimic gestures were replaced by simpler signals whose meanings were

12 This is how “person signs” are introduced in sign languages.

acquired through association rather than through pictorial representation. Meaning is then carried through cultural transmission, rather than in the signal itself.”

It should be noted that in contrast to pantomime, the conventions of protosign (or protolanguage) must be *learned*.

Bickerton (2007, p. 513) claims that animal calls cannot predicate. So, what is it in human communication that makes it possible for us to predicate, that is, say something *about* entities? The answer can be formulated as that the result of a communicative convention is a *label* (gesture or word)<sup>13</sup>. Gelman and Roberts (2017, p. 7900) analyze the role of labels in cultural inheritance. They argue that “category labels work in an almost paradoxical way to ensure stability in the transmission process, but simultaneously to permit and even foster conceptual change”. Because young children act as if their own knowledge state is the same as that of others, this will improve learning. Sabbagh and Henderson (2007) also argue that children’s understanding that words meanings are conventional makes their word learning more efficient. They write that “these limitation in children’s abilities to reason about others’ epistemic mental state may actually promote the development of an appreciation of conventionality” (Sabbagh and Henderson, 2007, p. 33). Labels for objects also license inferences that depend on underlying causal features of objects. As will be seen in Section Explaining relationships between concepts—the role of generics, such connections are taught *via* generics.

What has been called the *principle of conventionality* (Clark, 1993) states that words are efficient tools for communication when the form–meaning associations are known, shared, and expected within a language community. Already one-and-a-half-year-old children exploit this principle in their learning of language (Graham et al., 2006). Given the limited ToM of small children, it is likely that they believe that labels reflect what objects are called rather than what people call objects.

The gestures of great apes are sometimes learned *via* ontogenetic *ritualization* where individuals learn a gesture *via* regularly occurring dyadic interactions (see Arbib, 2012, Ch. 3 for an analysis of ritualization, and also Abramova, 2018). Non-human ritualizations never become labels that are shared within a society, but they remain signs that are used within a dyad of individuals, typically a mother and an infant. Therefore, the principle of conventionality may be a foundational block in the evolution of human communication that distinguishes it from that of other species.

Nevertheless, the “chimpanzee test” requires an explanation of why conventionalization has not evolved among great apes or other species. One possible explanation, comparing hominins to chimpanzees, is that the hominins increasingly engaged in *collective breeding*. This means that more children spent

time together and with other adults than their mothers, which would make it easier for the children to adopt the principle of conventionality (Volterra et al., 2018, p. 230). This explanation is supported indirectly by two other forms of conventionalization. Firstly, when Nicaraguan home signers met with other deaf children, a conventional sign language quickly emerged (Senghas and Coppola, 2001). Secondly, computer simulations of how a signaling systems become shared among artificial agents show that the more communicators that are involved, the faster the conventionalization occurs (Steels, 1998). In passing, it should be noted that collective breeding makes teaching more effective since more learners can be taught simultaneously and be taught by teachers that are more skilled than the average parent (Henrich, 2004).

## Communicating concepts

Within the cognitive sciences researchers disagreed on how the notion of a concept should be characterized. It is, however, sufficient for my purposes that having a concept involves the ability to recognize a *pattern* (Gärdenfors and Lindström, 2008). Some patterns representing concepts are perceptual and others, for example kinship relations, more abstract.

Evolutionarily, concepts have been exploited for everyday decision making, for example learning to recognize edible mushrooms, distinguishing the tracks of a hyena from that of a leopard, and recognizing an appropriate platform in Acheulean tool manufacture (Sterelny, 2012, p. 2142).

The fourth level of teaching—communicating concepts—presumes a system of conventional labels where categories and perhaps places can be referred to in addition to the actions depicted in pantomime. In human societies, the main methods to teach a concept for a category is to use a label (sound or gesture) standing for the concept together with pointing or some other technique for drawing attention to what is characteristic of the concept.

Csibra and Gergely (2009, 2011) argue that natural pedagogy is unique to humans, because other animals are not capable of learning *generic* knowledge from communication. Animal communication is confined to the here and now. Mattos and Hinzen (2015) suggest that natural pedagogy is one of the main functions of language. They write that humans have a “specific capacity to acquire, through communication, different kinds of information—respectively, *knowledge about kinds* and knowledge about particular events, actions and state of affairs which we will call here simply ‘*knowledge about facts*’<sup>14</sup>.” They also argue that children learn

<sup>13</sup> A label is an icon or a symbol in the terminology of Pierce (1932).

<sup>14</sup> This distinction is the same as the distinction between “knowing what” and “knowing that” in Gärdenfors and Stephens (2018).

knowledge about kinds earlier than they learn knowledge about facts.

## An archaeological example

As an example of a concepts that was necessary to communicate at an early stage, Gärdenfors and Högberg (2017), following Stout (2011), analyze what is required for teaching the concept of a “platform” that is required for manufacturing Acheulean hand axes. Importantly, the subgoals of the manufacturing cannot be perceived directly from the action sequence of a teacher and they are therefore very difficult to learn *via* imitation. The subgoal features can therefore not be identified by the learner by drawing attention or demonstration, but they must be taught with the aid of concepts. The more convoluted a technology is, the more subgoals are involved (Stout, 2011; Lombard and Haidle, 2012; Stout and Chaminade, 2012; Mahaney, 2014). The subgoals involve understanding how the action sequences should be chunked. Such a chunking is difficult to communicate without using concepts. Another example related to stone tool production comes from Wynn and Coolidge (2012, p. 70) who argue the “distal convexity” of a core is a necessary concept to communicate in Levallois technology.

The ability to communicate with concepts is a central component in the evolution of a protosign or a protolanguage. If the above argument is correct, it follows that *Homo erectus* had this ability. Even though we cannot present any clear evidence that they could also communicate relations between concepts, I believe this is very likely (see the following subsection) and I thus submit that *Homo erectus* communicated with protosign or a protolanguage. The same conclusion has been reached by other researchers, for example Bickerton (2009) and Everett (2017) (also Barham and Everett, 2021), albeit their arguments are different from mine. For example, Everett (2017, p. 99) argues that *Homo erectus* needed language to be able to travel.

## Explaining relationships between concepts—The role of generics

The fifth level of teaching—explaining relationships between concepts—requires a fully detached communication system that also can express causal connections.

I claim that generics form the main communicative tool for explaining relations between concepts. In particular, they provide information about the semantic structure of concepts. Gärdenfors and Osta Velez (2022) distinguish between two kinds: (a) *Property* generics dealing with characteristic properties of objects (“Ducks lay eggs”), and (b) *causal* generics (“Sharks kill people”). To this, one can add *normative* generics (“Boys don’t cry”).

Generics have been attested in all languages. Although, on the surface, they look like quantified sentences, generics are not ordinary sentences. Gärdenfors and Osta Velez (2022) argue that they should be seen as expressing *expectations* about relations between concepts. Linguistically, they can simply be formulated as relations between labels rather than as sentences, for example, “finders, keepers; losers, weepers”.

Generics form an efficient method of transmitting information about categories and their relations. In this way they speed up cultural transmission (Gelman and Roberts, 2017). Generics have a central role in teaching, in particular in what is called “natural pedagogy (Csibra and Gergely, 2009). We tell our children, already at a young age, things like: “cats say meow, dogs say woof, and cows say moo<sup>15</sup>”. Later in school, they learn generics like “tigers have stripes”, “copper conducts electricity” and “democracies have freedom of speech”. Such property generics is a way of presenting characteristic properties of various categories (Leslie, 2008). Learning about categories is primarily done *via* their characteristic properties. And when it comes to causal generics such as “smoking causes cancer” and “dogs bite people,” they function as guidelines for caution in actions (Sterken, 2015). Archaeologically and anthropologically, detached communication about the relations between animals, their tracks and their behavior is difficult without using some form of generics (MacDonald and Roebroeks, 2013). Although speculative, this connection could be an explanation of why generics are so central in language.

Several studies argue that young children understand generics and can distinguish them from non-generics on the basis of several types of cues. Gelman and Pelletier (2010) show that already by about the age of two and a half, children start producing generics and at the age of four generics constitute 3% of the sentences produced. This may not seem as a high rate, but it should be considered in the context of all other sorts of utterances that children can make. They rely on generics when they learn about concepts and they use them when drawing inferences. Cimpian and Scott (2012, p. 429) write about how children use generics to build up common knowledge: “Children’s realization that generic facts are widely known may be a crucial step in this direction—that is, a crucial step toward learning what knowledge can reasonably be presupposed when communicating with others.” In line with the position that generics are not ordinary sentences, children also interpret generics as different from quantificational sentences containing “all”, “some” or “most”.

15 Children’s picture books of animals and other object categories highlight the characteristic properties of the categories.

## Narrating

There is one linguistic unit that is not required for the previous levels of teaching, but which is central to narration: the *sentence*. A narrative consists of a sequence of sentences. The sixth level—narrating—therefore requires a protolanguage. As mentioned previously, *Sibierska* (2017) argues that pantomime (in the extended form) is sufficient for narration. Therefore, I agree with *Ferretti et al.* (2017, p. 8) that the origin of narratives coincides with pantomime (for communication).

A central question is what role sentential structures play in the evolution of communication. My proposal is that the primary function of sentences is that they describe *events*<sup>16</sup>. According to an earlier analysis of events (*Gärdenfors and Warglien, 2012; Warglien et al., 2012; Gärdenfors, 2014a,b*), events are based on *causal* relations: an event typically contains information about an agent who is the cause of an action that leads to a result related to a patient<sup>17</sup>. While causal generics express causal relations between concepts, sentences express causes or effects in single events. My hypothesis is that in the transition from protosign to protolanguage, the holophrases represented by a pantomime are supplemented by other components which represent core parts of the event (*Gärdenfors, 2014a; Gärdenfors and Osta Velez, 2022*).

Apes do not narrate. At best, they can express declaratives, mostly in the form of single signs (*Lyn et al., 2011*). During hominin evolution, causal cognition has been extended considerably, from understanding of the causation of your own actions, to understanding hidden causes of different kinds, such as physical causes, medical causes, and the intentions and beliefs of other individuals (*Lombard and Gärdenfors, 2021*). This extension of causal cognition is a necessary requirement for the capacity to narrate.

The ToM involved in understanding narratives includes communicative sign function and joint beliefs. As regards communication, a narrative is typically detached (*Corballis, 2015*). This holds in particular for *gossip*. Gossip normally contains expressions of the type “X did A to Y,” which involves identifying thematic roles such as agent, action and patient. Therefore, narration presumes a communicative system that contains at least a minimal level of syntax (*Gärdenfors, 2012*).

It is obvious that narrating has become a central part of human culture. It is, however, less clear that teaching has been the main driving factor behind the evolution of narration. Firstly, narration is important for planning for cooperation. Unless sentential structures involving agents and objects, it is

difficult to express a sequence of actions that is to be performed by a group of individuals.

Secondly, it has been argued that a key factor in cooperation in humans that distinguishes us from other species is the role of *reputation* (*Nowak and Sigmund, 2005; Gärdenfors et al., 2012*). One important mechanism in determining the reputation of an individual is gossip. Gossip is a form of narrative that often contains information about whether an individual can be trusted or not, which is a crucial factor in selecting who to cooperate with. Furthermore, *myths* are often presented in form of narratives and they function as carriers of common knowledge (joint beliefs). They contain much of the knowledge accumulated by a society (*Donald, 1991*).

Even though teaching may not have driven the evolution of narration, narration is closely connected to teaching. Narrative teaching is present in all cultures (*Csibra and Gergely, 2009; Kline, 2015*). Narratives are useful in conveying information about causal relations, in particular when the causes are not perceptually present.

My aim in this article has been to analyze the role of teaching in the evolution of communication. As a final remark, I note that a similar trajectory can be found in the development of children’s communication. Following *Mattos and Hinzen (2015)*, one can distinguish three levels in the early development of children’s communication: (i) infants’ initial vocabulary is linked to the different categories of objects they point to; (ii) they express generic information about categories; and (iii) they describe events using a sentential structure. It should be noted that there is a clear parallel to these levels with the communication required for levels 4, 5 and 6 of teaching. This argument provides some indirect support to the ordering of the levels of teaching that have formed the backbone of my analysis.

## Conclusion

My main thesis in this article is that the evolution of teaching is central for the evolution of increasingly complex communicative systems in the hominin species. Building on the argument that demonstration is evolutionarily prior to pantomime for instruction, which in turn is evolutionarily prior to pantomime (*Gärdenfors, 2021*), I have shown that such a series can provide an evolutionary explanation of different functions of communication.

I have focused on the role of demonstration and pantomime since this is the level of teaching where it seems that the hominins separated from the forms of teaching found in other species. In particular, this is the level where the communicative sign function appears. This is also the level where communication becomes detached from the here and now, which is a necessary condition for planning for cooperative actions.

<sup>16</sup> I consider *states* to be special cases of events.

<sup>17</sup> This theory builds on conceptual spaces (*Gärdenfors, 2000, 2013, 2014a*) where actions are modelled as force vectors (or patterns) and results as vectors describing a change in some property of the patient.

I am not claiming, however, that teaching is the only evolutionarily driving force. More generally, my position is that the evolution of communication is connected to advanced cooperation that is unique to the hominins (Gärdenfors, 2003; Gärdenfors et al., 2012). In line with this, teaching can be seen as a form of preparation for future cooperation.

## Data availability statement

All relevant data is contained within the article or in the references.

## Author contributions

The author confirms being the sole contributor of this work and has approved it for publication.

## References

- Abramova, E. (2018). The role of pantomime in gestural language evolution, its cognitive bases and an alternative. *J. Lang. Evol.* 3, 26–40. doi: 10.1093/jole/lzx021
- Arbib, M. A. (2012). *How the Brain Got Language: The Mirror System Hypothesis (Vol. 16)*. Oxford: Oxford University Press. doi: 10.1093/acprof:osobl/9780199896684.001.0001
- Arbib, M. A. (2018). In support of the role of pantomime in language evolution. *J. Lang. Evol.* 3, 41–44. doi: 10.1093/jole/lzx023
- Armstrong, D. F., and Wilcox, S. E. (2007). *The Gestural Origin of Language*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780195163483.001.0001
- Aussem, S., and Kita, S. (2021). Seeing iconic gesture promotes first- and second-order verb generalization in preschoolers. *Child Dev.* 92, 124–141. doi: 10.1111/cdev.13392
- Barham, L., and Everett, D. (2021). Semiotics and the origin of language in the Lower Palaeolithic. *J. Archaeol. Method Theor.* 28, 535–579. doi: 10.1007/s10816-020-09480-9
- Barnard, A. (2011). *Social Anthropology and Human Origins*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511974502
- Bar-On, D. (2013). Origins of meaning: must we ‘go Gricean’? *Mind and Language* 28, 342–375. doi: 10.1111/mila.12021
- Bates, E., Camaioni, L., and Volterra, V. (1975). The acquisition of performatives prior to speech. *Merrill Palmer Q. Behav. Dev.* 21, 205–226.
- Bates, L. A., Handford, R., Lee, P. C., Njiraini, N., Poole, J. H., Sayialel, K., et al. (2010). Why do African elephants (*Loxodonta africana*) simulate oestrus? An analysis of longitudinal data. *PLoS ONE* 5, e10052. doi: 10.1371/journal.pone.0010052
- Bickerton, D. (2002). “From protolanguage to language”, in *The Speciation of Modern Homo Sapiens*, ed T. J. Crow (Oxford: Oxford Academic), 103–120.
- Bickerton, D. (2007). Language evolution: a brief guide for linguists. *Lingua* 117, 510–526. doi: 10.1016/j.lingua.2005.02.006
- Bickerton, D. (2009). *Adam’s Tongue: How Humans Made Language, How Language Made Humans*. New York, NY: Macmillan.
- Boesch, C. (1991). Teaching among wild chimpanzees. *Anim. Behav.* 41, 530–532. doi: 10.1016/S0003-3472(05)80857-7
- Boesch, C., and Tomasello, M. (1998). Chimpanzee and human cultures. *Curr. Anthropol.* 39, 591–614. doi: 10.1086/204785
- Brinck, I. (2004a). “On the evolutionary origin of declarative pointing,” in *Workshop on Mirror Neurons, Mind-Reading and the Emergence of Language* (Parma: University of Parma, Department of Neurophysiology), 1–10.
- Brinck, I. (2004b). The pragmatics of imperative and declarative pointing. *Cogn. Sci.* 3, 429–446.
- Brinck, I., and Gärdenfors, P. (2003). Co-operation and communication in apes and humans. *Mind and Language* 18, 484–501. doi: 10.1111/1468-0017.00239
- Brown, S., Mittermaier, E., Kher, T., and Arnold, P. (2019). How pantomime works: implications for theories of language origin. *Front. Commun.* 4, 9. doi: 10.3389/fcomm.2019.00009
- Bruner, J. (1995). “From joint attention to the meeting of minds: An introduction,” in *Joint Attention: Its Origins and Role in Development*, eds C. Moore and P. J. Dunham (London: Routledge), 1–14.
- Caro, T. M., and Hauser, M. D. (1992). Is there teaching in nonhuman animals? *Q. Rev. Biol.* 67, 151–174. doi: 10.1086/417553
- Cartmill, E. A., and Goldin-Meadow, S. (2012). “Is pantomime a likely stage in language evolution? Evidence from human and primate gesture,” in *The Evolution of Language*, eds T. C. Scott-Phillips, M. Tamariz, E. A. Cartmill, and J. R. Hurford (Singapore: World Scientific), 418–419. doi: 10.1142/9789814401500\_0059
- Caruana, M. V., d’Errico, F., and Backwell, L. (2013). “Early hominin social learning strategies underlying the use and production of bone and stone tools,” in *Tool Use in Animals: Cognition and Ecology*, eds C. Sanz, J. Call, and C. Boesch (Cambridge: Cambridge University Press), 242–285. doi: 10.1017/CBO9780511894800.016
- Cassell, J., and McNeill, D. (1990). “Gesture and ground,” in *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics*, eds K. Hall, J.-P. Koenig, M. Meacham, S. Reinman, and L. A. Sutton (Berkeley, CA: Berkeley Linguistics Society), 57–68. doi: 10.3765/bls.v16i0.1689
- Castro, L., and Toro, M. A. (2004). Cumulative cultural evolution: the role of teaching. *J. Theor. Biol.* 347, 74–83. doi: 10.1016/j.jtbi.2014.01.006
- Cimpian, A., and Scott, R. M. (2012). Children expect generic knowledge to be widely shared. *Cognition* 123, 419–433. doi: 10.1016/j.cognition.2012.02.003
- Clark, E. V. (1993). Conventionality and contrast: pragmatic principles with lexical consequences. *Lehrer and Kittay* 1992a, 171–188.
- Clark, H. H. (1992). *Arenas of Language Use*. Chicago, IL: University of Chicago Press.

## Funding

This research was funded by Lund University.

## Conflict of interest

The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

## Publisher’s note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

- Corballis, M. C. (2014). "The word according to Adam," in *From Gesture in Conversation to Visible Action as Utterance*, ed M. Seyfeddinipur and M. Gullberg (Amsterdam: John Benjamins), 177–198. doi: 10.1075/z.188.09cor
- Corballis, M. C. (2015). *The Wandering Mind*. Chicago, IL: University of Chicago Press. doi: 10.7208/chicago/9780226238753.001.0001
- Crockford, C., Wittig, R. M., Mundry, R., and Zuberbühler, K. (2012). Wild chimpanzees inform ignorant group members of danger. *Curr. Biol.* 22, 142–146. doi: 10.1016/j.cub.2011.11.053
- Csibra, G., and Gergely, G. (2006). Social learning and social cognition: the case for pedagogy. Processes of change in brain and cognitive development. *Atten. Perform. XXI* 21, 249–274.
- Csibra, G., and Gergely, G. (2009). Natural pedagogy. *Trends Cogn. Sci.* 13, 148–153. doi: 10.1016/j.tics.2009.01.005
- Csibra, G., and Gergely, G. (2011). Natural pedagogy as evolutionary adaptation. *Phil. Trans. Royal Soc. B: Biol. Sci.* 366, 1149–1157. doi: 10.1098/rstb.2010.0319
- Delagnes, A., and Roche, H. (2005). Late Pliocene hominid knapping skills: the case of Lokalei 2C, West Turkana, Kenya. *J. Hum. Evol.* 48, 435–472. doi: 10.1016/j.jhevol.2004.12.005
- D'Errico, F., and Banks, W. E. (2015). The archaeology of teaching: a conceptual framework. *Cambridge Archaeol. J.* 25, 859–866. doi: 10.1017/S0959774315000384
- Dessalles, J. L. (2007). *Why We Talk: The Evolutionary Origins of Language* (Vol. 5). Oxford: Oxford University Press.
- Donald, M. (1991). *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Cambridge, MA: Harvard University Press.
- Donald, M. (2001). *A Mind so Rare: The Evolution of Human Consciousness*. New York, NY: Norton and Company.
- Donald, M. (2012). "The mimetic origins of language," in *The Oxford Handbook of Language Evolution*. Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780199541119.013.0017
- Dunbar, R. I. (1996). "Groups, gossip, and the evolution of language," in *New Aspects of Human Ethology* (Boston, MA: Springer), 77–89. doi: 10.1007/978-0-585-34289-4\_5
- Ekström, A., Nirme, J., and Gärdenfors, P. (2022). Motion iconicity in prosody (Submitted). doi: 10.3389/fcomm.2022.994162
- Everett, D. (2017). *How Language Began: The Story of Humanity's Greatest Invention*. London: Profile Books.
- Ferretti, F., Adornetti, I., Chiera, A., Nicchiarelli, S., Magni, R., Valeri, G., et al. (2017). Mental Time Travel and language evolution: a narrative account of the origins of human communication. *Lang. Sci.* 63, 105–118. doi: 10.1016/j.langsci.2017.01.002
- Froese, T., and Leavens, D. A. (2014). The direct perception hypothesis: perceiving the intention of another's action hinders its precise imitation. *Front. Psychol.* 5, 65. doi: 10.3389/fpsyg.2014.00065
- Gärdenfors, P. (1995). Cued and detached representations in animal cognition. *Behav. Processes* 35, 263–273. doi: 10.1016/0376-6357(95)00043-7
- Gärdenfors, P. (2000). *Conceptual Spaces: The Geometry of Thought*. Cambridge, MA: MIT Press. doi: 10.7551/mitpress/2076.001.0001
- Gärdenfors, P. (2003). *How Homo Became Sapiens: On the Evolution of Thinking*. Oxford: Oxford University Press.
- Gärdenfors, P. (2004). "Cooperation and the evolution of symbolic communication," in *The Evolution of Communication Systems*, eds K. Oller and U. Griebel (Cambridge, MA: MIT Press), 237–256.
- Gärdenfors, P. (2007). "Evolutionary and developmental aspects of intersubjectivity," in *Consciousness Transitions: Phylogenetic, Ontogenetic and Physiological Aspects*, eds by H. Liljenström and P. Århem (Amsterdam: Elsevier), 281–305. doi: 10.1016/B978-0-44452977-0/50013-9
- Gärdenfors, P. (2012). "The cognitive and communicative demands of cooperation," in *Games, Actions and Social Software* (Berlin, Heidelberg: Springer), 164–183. doi: 10.1007/978-3-642-29326-9\_9
- Gärdenfors, P. (2013). "The evolution of semantics: sharing conceptual domains," in *The Evolutionary Emergence of Language*, eds C. Knight, M. Studdert-Kennedy, and J. R. Hurford (Cambridge: Cambridge University Press), 139–159. doi: 10.1093/acprof:oso/9780199654840.003.0008
- Gärdenfors, P. (2014a). *The Geometry of Meaning: Semantics Based on Conceptual Spaces*. Cambridge, MA: MIT press. doi: 10.7551/mitpress/9629.001.0001
- Gärdenfors, P. (2014b). The evolution of sentential structure. *Humana. Mente J. Phil. Stud.* 7, 79–97.
- Gärdenfors, P. (2017). Demonstration and pantomime in the evolution of teaching. *Front. Psychol.* 8, 415. doi: 10.3389/fpsyg.2017.00415
- Gärdenfors, P. (2020). From pantomime to protolanguage. *Paradigmi.* 38, 251–268. doi: 10.30460/97897
- Gärdenfors, P. (2021). Causal reasoning and event cognition as evolutionary determinants of language structure. *Entropy* 23, 843. doi: 10.3390/e23070843
- Gärdenfors, P., Brinck, I., and Osvath, M. (2012). "The tripod effect: co-evolution of cooperation, cognition and communication," in *The Symbolic Species Evolved* (Dordrecht: Springer), 193–222. doi: 10.1007/978-94-007-2336-8\_10
- Gärdenfors, P., and Högborg, A. (2017). The archaeology of teaching and the evolution of Homo docens. *Curr. Anthropol.* 58, 188–208. doi: 10.1086/691178
- Gärdenfors, P., and Högborg, A. (2021). "Evolution of intentional teaching," in *Handbook of Human Symbolic Evolution*. Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780198813781.013.9
- Gärdenfors, P., and Lindström, P. (2008). "Understanding is experiencing a pattern," in *A Smorgasbord of Cognitive Science*, eds Peter Gärdenfors and Annika Wallin (Nora: Nya Doxa), 149–164.
- Gärdenfors, P., and Osta Velez, M. (2022). *Generics as Expectations* (submitted).
- Gärdenfors, P., and Osvath, M. (2010). "Prospection as a cognitive precursor to symbolic communication," in *Evolution of Language: Bilingual Approaches*, eds R. K. Larson, V. Déprez and H. Yamakido (Cambridge: Cambridge University Press), 103–114. doi: 10.1017/CBO9780511817755.007
- Gärdenfors, P., and Stephens, A. (2018). Induction and knowledge-what. *Eur. J. Phil. Sci.* 8, 471–491. doi: 10.1007/s13194-017-0196-y
- Gärdenfors, P., and Warglien, M. (2012). Using conceptual spaces to model actions and events. *J. Semant.* 29, 487–519. doi: 10.1093/jos/ffs007
- Gelman, S. A., and Pelletier, F. J. (2010). "Generics as a window onto young children's concepts," in *Kinds, Things, and Stuff: Mass Terms and Generics*, ed J. F. Pelletier (Oxford: Oxford Academic), 100–120. doi: 10.1093/acprof:oso/9780195382891.003.0006
- Gelman, S. A., and Roberts, S. O. (2017). How language shapes the cultural inheritance of categories. *Proc. Nat. Acad. Sci. USA.* 114, 7900–7907. doi: 10.1073/pnas.1621073114
- Gergely, G., Egyed, K., and Király, I. (2007). On pedagogy. *Dev. Sci.* 10, 139–146. doi: 10.1111/j.1467-7687.2007.00576.x
- Geurts, B. (2020). *Evolutionary Pragmatics: From Chimp-Style Communication*. Available online at: [https://www.researchgate.net/publication/345806282\\_Evolutionary\\_pragmatics\\_from\\_chimp-style\\_communication\\_to\\_human\\_discourse](https://www.researchgate.net/publication/345806282_Evolutionary_pragmatics_from_chimp-style_communication_to_human_discourse) (accessed November 22, 2022).
- Gibson, K. R. (2013). Talking about apes, birds, bees, and other living creatures: language evolution in the light of comparative animal behaviour. *Evol. Emerg. Lang.: Evid. Infer.* 17, 204. doi: 10.1093/acprof:oso/9780199654840.003.0011
- Gillespie-Lynch, K., Greenfield, P. M., Lyn, H., and Savage-Rumbaugh, S. (2014). Gestural and symbolic development among apes and humans: support for a multimodal theory of language evolution. *Front. Psychol.* 5, 1228. doi: 10.3389/fpsyg.2014.01228
- Graham, S. A., Stock, H., and Henderson, A. M. (2006). Nineteen-month-olds' understanding of the conventionality of object labels versus desires. *Infancy* 9, 341–350. doi: 10.1207/s15327078in0903\_5
- Gullberg, M. (1998). *Gesture as a Communication Strategy in Second Language Discourse: A Study of Learners of French and Swedish* (Vol. 35). Lund: Lund University.
- Gullberg, M. (2011). "Multilingual multimodality: Communicative difficulties and their solutions in second-language use," in *Embodied Interaction: Language and Body in the Material World*, eds J. Streeck, C. Godwin, and C. LeBaron (Cambridge: Cambridge University Press), 137–151.
- Haiman, J. (2018). *Ideophones and the Evolution of Language*. Cambridge: Cambridge University Press. doi: 10.1017/9781107706897
- Henrich, J. (2004). Demography and cultural evolution: how adaptive cultural processes can produce maladaptive losses: the Tasmanian case. *Am. Antiq.* 69, 197–214. doi: 10.2307/4128416
- Hockett, C. F. (1960). The origin of speech. *Sci. Am.* 203, 88–97. doi: 10.1038/scientificamerican0960-88
- Hoppitt, W. J. E., Brown, R. G., Kendal, R., Rendell, L., Thornton, A., Webster, M. M., et al. (2008). Lessons from animal teaching. *Trends Ecol. Evol.* 23, 486–493. doi: 10.1016/j.tree.2008.05.008
- Johansson, S. (2005). *Origins of Language: Constraints on Hypotheses* (Vol. 5). Amsterdam: John Benjamins Publishing. doi: 10.1075/celcr.5

- Kendon, A. (2004). *Gesture: Visible Action as Utterance*. Cambridge: Cambridge University Press. doi: 10.1017/CBO9780511807572
- Kendon, A. (2017). Reflections on the “gesture-first” hypothesis of language origins. *Psychon. Bull. Rev.* 24, 163–170. doi: 10.3758/s13423-016-1117-3
- Kline, M. A. (2015). How to learn about teaching: an evolutionary framework for the study of teaching behavior in humans and other animals. *Behav. Brain Sci.* 38, 1–17. doi: 10.1017/S0140525X14000090
- Leslie, A. M. (1987). Pretense and representation: the origins of “theory of mind.” *Psychol. Rev.* 94, 412. doi: 10.1037/0033-295X.94.4.412
- Leslie, S. J. (2008). Generics: cognition and acquisition. *Phil. Rev.* 117, 1–47. doi: 10.1215/00318108-2007-023
- Lombard, M., and Gärdenfors, P. (2021). Causal cognition and theory of mind in evolutionary cognitive archaeology. *Biol. Theor.* 1–19. doi: 10.1007/s13752-020-00372-5
- Lombard, M., and Haidle, M. N. (2012). Thinking a bow-and-arrow set: cognitive implications of Middle Stone Age bow and stone-tipped arrow technology. *Cambridge Archaeol. J.* 22, 237–264. doi: 10.1017/S095977431200025X
- Lyn, H., Greenfield, P. M., Savage-Rumbaugh, S., Gillespie-Lynch, K., and Hopkins, W. D. (2011). Nonhuman primates do declare! A comparison of declarative symbol and gesture use in two children, two bonobos, and a chimpanzee. *Lang. Commun.* 31, 63–74. doi: 10.1016/j.langcom.2010.11.001
- MacDonald, K., and Roebroeks, W. (2013). “Neanderthal linguistic abilities: an alternative view,” in *The Evolutionary Emergence of Language: Evidence and Inference*, eds R. Botha and M. Everaert (Oxford: Oxford Academic), 97–117. doi: 10.1093/acprof:oso/9780199654840.003.0006
- Maestriperi, D. (1995). First steps in the macaque world: do rhesus mothers encourage their infants’ independent locomotion? *Anim. Behav.* 49, 1541–1549. doi: 10.1016/0003-3472(95)90075-6
- Maestriperi, D. (1996). Maternal encouragement in nonhuman primates and the question of animal teaching. *Hum. Nat.* 6, 361–378. doi: 10.1007/BF02734206
- Mahaney, R. A. (2014). Exploring the complexity and structure of Acheulean stonknapping in relation to natural language. *PaleoAnthropology* 2014, 586–606. doi: 10.4207/PA.2014.ART90
- Mattos, O., and Hinzen, W. (2015). The linguistic roots of natural pedagogy. *Front. Psychol.* 6, 1424. doi: 10.3389/fpsyg.2015.01424
- McNeill, D. (1992). *Hand and Mind: What Gestures Reveal About Thought*. Chicago, IL: University of Chicago Press.
- McNeill, D. (2013). “The co-evolution of gesture and speech, and downstream consequences,” in *Body - Language - Communication: An International Handbook on Multimodality in Human Interaction. Vol. 1*, eds C. Miller, A. Cienki, E. Fricke, S. Ladewig, D. McNeill, and S. Tessendorf (Berlin: De Gruyter Mouton), 480–512. doi: 10.1515/9783110261318.480
- Mittelberg, I. (2019). Visuo-kinetic signs are inherently metonymic: how embodied metonymy motivates forms, functions, and schematic patterns in gesture. *Front. Psychol.* 10, 254. doi: 10.3389/fpsyg.2019.00254
- Moore, R. (2018). Gricean communication, language development, and animal minds. *Philosophy Compass* 13, e12550. doi: 10.1111/phc3.12550
- Morgan, T. J. H., Uomini, N. T., Rendell, L. E., Chouinard-Thuly, L., Street, S. E., Lewis, H. M., et al. (2015). Experimental evidence for the co-evolution of hominin tool-making teaching and language. *Nat. Commun.* 6, 1–8. doi: 10.1038/ncomms7029
- Myowa-Yamakoshi, M., and Matsuzawa, T. (1999). Factors influencing imitation of manipulatory actions in chimpanzees (*Pan troglodytes*). *J. Comp. Psychol.* 113, 128. doi: 10.1037/0735-7036.113.2.128
- Nowak, M. A., and Sigmund, K. (2005). Evolution of indirect reciprocity. *Nature* 437, 1291–1298. doi: 10.1038/nature04131
- Ortega, G., Sümer, B., and Özyürek, A. (2017). Type of iconicity matters in the vocabulary development of signing children. *Dev. Psychol.* 53, 89. doi: 10.1037/dev0000161
- Orwell, G. (1948/1968). *New Words. The Collected Essays, Journalism and Letters of George Orwell*. San Diego, CA: Harcourt.
- Osvath, M., and Gärdenfors, P. (2005). Oldowan culture and the evolution of anticipatory cognition. *Lund Univ. Cogn. Stud.* 122, 1–16.
- Pierce, C. S. (1932). *Collected Papers*. Cambridge, MA: Charles Hartshorne and Paul Weiss.
- Progovac, L. (2016). A gradualist scenario for language evolution: precise linguistic reconstruction of early human (and Neandertal) grammars. *Front. Psychol.* 7, 1714. doi: 10.3389/fpsyg.2016.01714
- Radvansky, G. A., and Zacks, J. M. (2014). *Event Cognition*. Oxford: Oxford University Press. doi: 10.1093/acprof:oso/9780199898138.001.0001
- Rimmón-Qèñán, Š. (1983). *Narrative Fiction*. London: Routledge. doi: 10.4324/9780203130650
- Russon, A., and Andrews, K. (2011). Orangutan pantomime: elaborating the message. *Biol. Lett.* 7, 627–630. doi: 10.1098/rsbl.2010.0564
- Sabbagh, M. A., and Henderson, A. M. (2007). How an appreciation of conventionality shapes early word learning. *New Dir. Child Adolesc. Dev.* 2007, 25–37. doi: 10.1002/cd.180
- Senghas, A., and Coppola, M. (2001). Children creating language: how Nicaraguan Sign Language acquired a spatial grammar. *Psychol. Sci.* 12, 323–328. doi: 10.1111/1467-9280.00359
- Sibierska, M. (2017). Storytelling without telling: the non-linguistic nature of narratives from evolutionary and narratological perspectives. *Lang. Commun.* 54, 47–55. doi: 10.1016/j.langcom.2016.10.005
- Steels, L. (1998). The origins of ontologies and communication conventions in multi-agent systems. *Auton. Agent Multi. Agent Syst.* 1, 169–194. doi: 10.1023/A:1010002801935
- Sterelny, K. (2012). Language, gesture, skill: the co-evolutionary foundations of language. *Phil. Trans. Royal Soc. B: Biol. Sci.* 367, 2141–2151. doi: 10.1098/rstb.2012.0116
- Sterelny, K. (2014). “Constructing the cooperative niche,” in *Entangled Life* (Dordrecht: Springer), 261–279. doi: 10.1007/978-94-007-7067-6\_13
- Sterken, R. K. (2015). Generics in context. *Philosopher’s Imprint* 15, 1–30.
- Stout, D. (2011). “Stone tool making and the evolution of human culture and cognition.” *Phil. Trans. Royal Soc. London, B: Biol. Sci.* 366:1050–1059. doi: 10.1098/rstb.2010.0369
- Stout, D. (2018). Archaeology and the evolutionary neuroscience of language: the technological pedagogy hypothesis. *Interact. Stud.* 19, 256–271. doi: 10.1075/is.17033.sto
- Stout, D., and Chaminade, T. (2012). Stone tools, language and the brain in human evolution. *Phil. Trans. Royal Soc. B: Biol. Sci.* 367, 75–87. doi: 10.1098/rstb.2011.0099
- Tehrani, J. J., and Riede, F. (2008). Towards an archaeology of pedagogy: learning, teaching and the generation of material culture traditions. *World Archaeol.* 40, 316–331.
- Thornton, A., and McAuliffe, K. (2006). Teaching in wild meerkats. *Science* 313, 227–229. doi: 10.1126/science.1128727
- Tomasello, M. (1999). *The Cultural Origins of Human Cognition*. Cambridge, MA: Harvard University Press. doi: 10.4159/9780674044371
- Tomasello, M. (2009). *The Cultural Origins of Human Cognition*. Harvard university press. doi: 10.1080/00438240802261267
- Toth, N., Schick, K. D., Savage-Rumbaugh, S., Sevik, R. A., and Rumbaugh, D. M. (1993). Pan the tool-maker: investigations into the stone tool-making and tool-using capabilities of a bonobo (*Pan paniscus*). *J. Archaeol. Sci.* 20, 81–91. doi: 10.1006/jasc.1993.1006
- Volterra, V., Capirci, O., Rinaldi, P., and Sparaci, L. (2018). From action to spoken and signed language through gesture: some basic developmental issues for a discussion on the evolution of the human language-ready brain. *Interact. Stud.* 19, 216–238. doi: 10.1075/is.17027.vol
- Waciewicz, P., and Zywczyński, P. (2021). *Pantomimic Conceptions of Language Origins. Oxford Handbook of Human Symbolic Evolution*. Oxford: Oxford University Press. doi: 10.1093/oxfordhb/9780198813781.013.30
- Warglien, M., Gärdenfors, P., and Westera, M. (2012). Event structure, conceptual spaces and the semantics of verbs. *Theor. Linguist.* 38, 159–193. doi: 10.1515/tl-2012-0010
- Whiten, A. (1999). “The evolution of deep social mind in humans,” in *The Descent of Mind: Psychological Perspectives on Hominid Evolution*, eds M. C. Corballis and S. E. G. Lea (New York, NY: Oxford University Press), 173–193. doi: 10.1093/acprof:oso/9780192632593.003.0010
- Whiten, A., Horner, V., and Marshall-Pescini, S. (2005). 11 Selective imitation in child and chimpanzee: a window. *Perspectives on Imitation, Volume 1: From Neuroscience to Social Science-Volume 1: Mechanisms of Imitation and Imitation in Animals* 1, 263.
- Whiten, A., McGuigan, N., Marshall-Pescini, S., and Hopper, L. M. (2009). Emulation, imitation, over-imitation and the scope of culture for child and chimpanzee. *Phil. Trans. Royal Soc. B: Biol. Sci.* 364, 2417–2428. doi: 10.1098/rstb.2009.0069
- Wrangham, R. (2009). *Catching Fire: How Cooking Made us Human*. Basic books.

Wynn, T., and Coolidge, F. L. (2012). *How to Think Like a Neandertal*. Oxford University Press.

Wynn, T., Hernandez-Aguilar, R. A., Marchant, L. F., and McGrew, W. C. (2011). "An ape's view of the Oldowan" revisited. *Evolutionary Anthropology: Issues, News, and Reviews* 20, 181–197. doi: 10.1002/evan.20323

Zlatev, J., Persson, T., and Gärdenfors, P. (2005). Bodily mimesis as "the missing link" in human cognitive evolution. *Lund Univ. Cogn. Stud.* 121, 1–45.

Zuberbühler, K. (2013). Acquired mirroring and intentional communication in primates. *Lang. Cogn.* 5, 133–143. doi: 10.1515/langcog-2013-0008x

Zukow-Goldring, P., and Arbib, M. A. (2007). Affordances, effectivities, and assisted imitation: Caregivers and the directing of attention. *Neurocomputing* 70, 2181–2193. doi: 10.1016/j.neucom.2006.02.029

Zywicznyński, P., Gontier, N., and Wacewicz, S. (2017). The evolution of (proto-) language: focus on mechanisms. *Lang. Sci.* 63, 1–11. doi: 10.1016/j.langsci.2017.06.004

Zywicznyński, P., Wacewicz, S., and Sibierska, M. (2018). Defining pantomime for language evolution research. *Topoi* 37, 307–318. doi: 10.1007/s11245-016-9425-9