

This is a draft of a chapter that has been accepted for publication by Oxford University Press in the forthcoming book *Oxford Handbook of Human Symbolic Evolution*, 2nd edition, edited by A. Lock, C. Sinha, N. Gontier, due for publication in 2020.

Pantomimic Conceptions of Language Origins

Sławomir Wacewicz and Przemysław Żywicznyński

Abstract

Could pantomime have been the key step in the evolutionary emergence of symbolic communication? Such a possibility has been consistently present in the intellectual reflection on language origins. What makes pantomime interesting from this perspective is its rich expressive potential, since it can convey open-ended, semantically universal and displaced meanings without relying on semiotic conventions, so that spontaneous pantomimes can be recognized as such and successfully interpreted. Definitions are important in classifying a particular scenario as “pantomimic”. In this chapter, we employ a ‘rich’ definition of pantomime: we describe it as bodily-mimetic communication which is non-conventional, improvised, performed with the whole body, holistic, communicatively and semantically complex. Based on this foundation, we review and evaluate pantomimic accounts of language origins, from the past to the present, and we particularly focus on the contemporary pantomime accounts given by Michael Arbib, Michael Tomasello, and Jordan Zlatev.

Keywords: pantomime, mime, gesture, mimesis, multimodality, language evolution, protolanguage, iconicity, conventionalization, enactment

Pantomimic Conceptions of Language Origins

Pantomime, itself nonlinguistic, provides a credible stepping stone to the first hallmarks of language. In particular, pantomime can communicate an unlimited range of novel messages (is *open-ended*) that are not restricted to a particular topic (is *semantically universal*) nor to the immediate here-and-now (is *displaced*). This has not escaped the notice of influential authors, historical as well as contemporary, whose accounts of language origins can be classified as “pantomimic”.

As with other scenarios in language evolution research, the evaluation of pantomimic scenarios ultimately depends on whether sufficient weight of converging multidisciplinary evidence (cf. Fitch, 2017; Wacewicz & Żywiczyński, 2015) can be mustered to make them convincing. This is the evidential question. But in the case of pantomimic scenarios the problem is more complex, because we cannot answer the evidential question without first answering the logically prior definitional question about the nature of pantomime. In the many literatures relevant to language evolution, the word “pantomime” is frequently used in a loose sense close to “gesture”; distinguishing these two notions, and correspondingly, gestural from *true* pantomimic theories is a difficult task that we address in Section 2, together with other thorny definitional issues. But we wish to note that this problem – to what extent a theory truly refers to the essence of pantomime, and to what extent only to the name – is a fundamental one that will recur throughout this chapter.

Mindful of this caveat, in sections 3 and 4 we provide an overview of historical and contemporary pantomimic accounts of language origins. Of them, Zlatev, Wacewicz, Żywiczyński, & van de Weijer’s (2017) pantomime-first deserves attention as perhaps the only unambiguously pantomimic proposal. However, we start with Michael Arbib’s (2005, 2012, 2016) Mirror System Hypothesis, which is the best known and most detailed, and in many respects can be treated as a prototypical pantomimic scenario. We conclude by

reviewing the criticisms of the role of pantomime in language origins and their polemical discussion.

Definitions and Classifications

One of the main axes of recent debates on the origins of symbolic communication is the question of the modality in which the proximate precursors of language arose (cf. Żywicznyński, Gontier, & Waciewicz, 2017). Pantomime-first and gesture-first theories emphasize the role of bodily/visual signaling, in which they differ from speech-first theories, and to a smaller degree, from multimodal theories (Figure 1). It is common in the literature to apply the adjective “gestural” very broadly (e.g. Corballis, 2013; Goodwin, 2017; Irvine, 2016; Sterelny, 2012a; Tramacere & Moore, 2018), so as to include theories for which we consider the term *visual/bodily* much more accurate. Conversely, pantomime is often defined loosely as “iconic gesturing” (e.g. Brown, Mittermaier, Kher, & Arnold, 2019; Tomasello, 2008), which however is imprecise and easily confusable with iconic *co-speech* gesture. A clear distinction between gesture and pantomime is rarely made, and many authors equivocate between the two (e.g. Corballis, 2014a, 2014b). This leads to the question of how to tell a “gestural” from a “pantomimic” scenario – and is there a meaningful difference?

As we show below, the separation of pantomime-first from gesture-first theories is not only possible but also theoretically productive; it rests both on the difference in meaning between gesture and pantomime, and on the different ranges of argumentation.

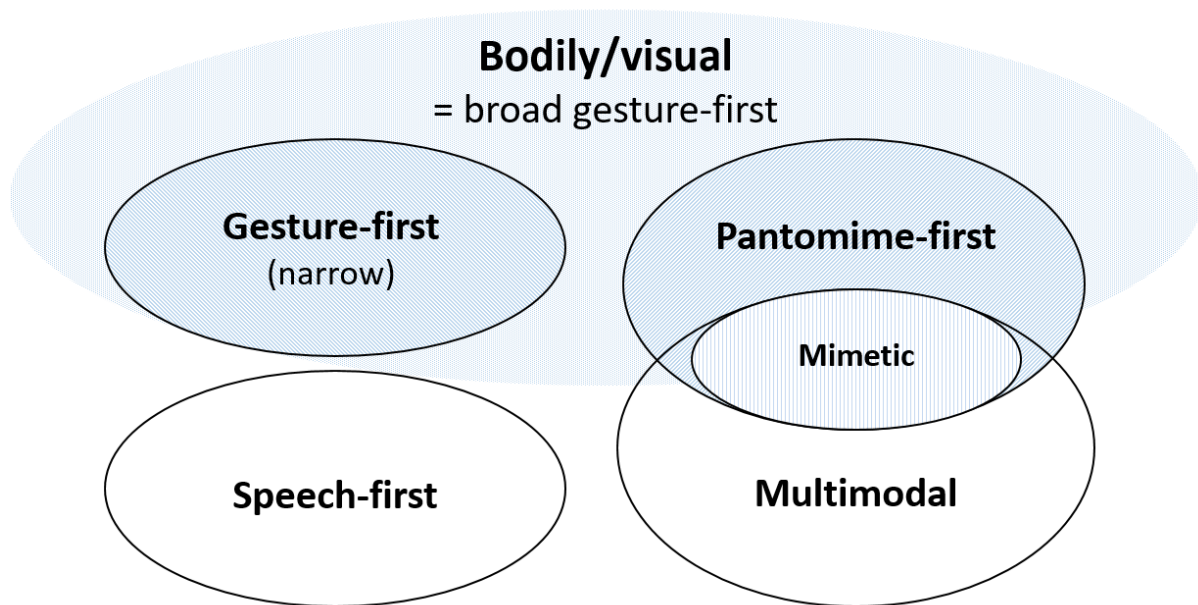


Figure 1. Language origins theories classified by their leading communicative modality.

Speech-first or *phonology-first* theories of language origins maintain that human language arose from ancestral primate vocalizations, i.e. has always been primarily vocal-auditory (Dunbar, 1996; Burling, 2005; MacNeilage, 2008; Fitch, 2010).

Gestural, or *gesture-first*, or *gestural primacy* theories are a group of theoretical accounts or scenarios which postulate that language or protolanguage began as a gestural communication system (Armstrong & Wilcox, 2007; Corballis, 2002a; Hewes, 1977). “Gesture” is sometimes used very broadly, and in most such cases *visual/bodily* theories would be a more accurate description.

Multimodal theories of language origins can be seen as a middle-ground position, on which human communication has always been a complex of bodily movement accompanied by vocalization (Kendon, 2011; Levinson & Holler, 2014; McNeill, 2012; Sandler, 2013; see Goldin-Meadow, this volume).

Mimetic theories, which we classify as essentially pantomime-first, are multimodal but with a much greater emphasis on whole body, visually perceived signaling (e.g. Donald, 1990; Zlatev, 2008).

Gesture

“Gesture” is an umbrella category hard to define in a clear and theory-independent way (Andrén, 2010). Across the diverse interdisciplinary literatures relevant to language evolution research, the uses of the term comprise a heterogeneous variety of qualitatively different forms of bodily action: from spontaneous hand movement accompanying speech, to glances, to postures, and even the orofacial gestures of the mouth area (e.g. Corballis, 2003), and multimodal gestures such as tactile and auditory gestures (e.g. Liebal & Oña, 2018).¹ The founding father of modern gesture-first theories, Gordon Hewes (1996, p. 572), proposes a broad definition of gesture, inclusive of almost all visually perceived bodily communication.

However, the everyday, intuitive meaning of “gesture” is much narrower, and connotes manual communicative movements that are perceived visually. Hewes (1996, p. 572) recognizes this in his definition, noting that prototypical gestures are those performed by the fingers, hands and arms.

Pantomime

The technical meanings of pantomime vary across a considerable number of disciplines such as theatrical studies, anthropology, and neuroscience – some with centuries if not millennia of tradition (Żywicznyński, Waciewicz, & Sibierska, 2018). But pantomime has one main intuitive meaning, well established in popular discourse and understood without much terminological reflection: essentially, pantomime is speechless bodily enactment of events for communication and artistic performance. Prototypical real-life examples of pantomime can be found in theatre, ritual or bodily narratives of hunters-gatherers, and cannot be described as gestures.

¹ Even articulatory gestures of the speech apparatus can sometimes be classified as gestures; for example, Armstrong & Wilcox (2007, p. 67) seem to imply that speech, as “planned sequences of musculo-skeletal actions”, is effectively a form of gesture. But as Kendon (2008, p. 13) observes, such a broad definition would also subsume cutting down trees or driving.

In a recent paper (Żywiczyński et al., 2018), we advocated a more technical top-down definition of pantomime that preserves this basic meaning and facilitates terminological agreement; it is tailored for language evolution research, that is it describes spontaneous or naïve pantomime rather than heavily conventionalized systems like in the game of charades or performance of professional mimes. We posit that pantomime, as a communicative system, should be understood as a prototype category with fuzzy membership, and should be defined by the following features, ordered by importance:

- a) *mimetic* (comprising self-initiated bodily actions that are representational. i.e. stand for something else than themselves), with “pantomime” being indeed a near-synonym to mimetic communication;
- b) *non-conventional* (not relying on communicative norms) and *motivated* (pantomimic communication depends on the similarity between its form and intended meaning);
- c) *primarily visual* (making use of the motor-visual channel) but possibly *multimodal* to some extent (i.e. potentially accompanied by non-linguistic vocalization);
- d) *whole-body action* (involving the action of all the limbs, head and torso rather than being exclusively manual);
- e) *holistic* (consisting of unanalyzed continuous expression rather than built from discrete structural components);
- f) *improvized* (impromptu, unstandardized).

We also proposed that for pantomime-first scenarios to work, pantomime should ideally meet the following additional criteria:

- g) communicatively *complex* and *self-sufficient* (being able refer to events and sequences of events without relying on other semiotic resources);

- h) *semantically advanced (displaced, open-ended and domain-universal – capable of expressing potentially unlimited range of meanings that are not limited to the here-and-now or to a predefined set of semantic domains).*²

In short, “gesture” naturally connotes manual movements that are perceived visually, but is underspecified with respect to its remaining characteristics. In contrast, “pantomime” stands for a whole-body communication system that is overspecified as mimetic, non-conventional, holistic, and improvised (and ideally should have considerable expressive power).

Gesture-First Theories: Mostly Manual

Michael Corballis (2013, p. 171) summarizes the arguments favouring the gesture-first theories of language origins in the following way:

- a) the use of the hands as the more natural way to depict events in space and time;
- b) the ability of nonhuman primates to use manual action flexibly and intentionally;
- c) the nature of the primate mirror system and its homology with the language circuits in the human brain;
- d) the relative success in teaching apes to communicate gesturally rather than vocally;
- e) the ready invention of sophisticated signed languages by the deaf;
- f) the critical role of pointing in the way young children learn language; and
- g) the correlation between handedness and cerebral asymmetry for language.

This could be complemented by several more arguments, many of them quite speculative. For example, Hewes (1996) conjectured that *volar depigmentation* (lighter coloration of the inside of the palm in darker-skinned populations) could have been selected for better manual communication. Other proposed links from manual action to language

² The communicative effectiveness of such pantomime of course rests on both the producer’s ability to use pantomime to express meanings and, crucially, the receiver’s ability to interpret pantomime as bearing such meanings (see e.g. Arbib, 2012 and section 4.1.6 below).

involve tool use (and *instruction* of tool use, Stout & Chaminade, 2012) or even ballistic movements; for example, Calvin identified the action of throwing stones to hunt small game as responsible for the emergence of a specialized neural circuit in the brain for manual movements, later exapted for grammatical patterns (Calvin & Bickerton, 2000). There is also a growing body of evidence that action gestures play a significant role in the acquisition of both spoken and signed languages (Volterra, Capirci, Rinaldi, & Sparaci, 2018). Waciewicz and Żywicznyński (2008) underscored the greater *secrecy* of gestural communication. Unlike sound, which is broadcast in all directions and easily overheard, gesture allows the producers to better control who receives the message; the possibility to transfer honest information only to one's kin or allies could have played a role in the competitive social environment at the early stages of language emergence.

Importantly for the present context, all of this argumentation supports specifically manual gesture rather than the broader category of visual/bodily communication. Only the first two points can be extended to reinforce whole-body pantomimic theories. Most importantly, point a), when interpreted as the greater natural *iconicity* of visual than vocal representations, works well for all bodily signals, manual or non-manual.

Pantomime-First Theories: Whole-Body, Primary Iconicity

Iconicity deserves special attention as arguably the most powerful argument summoned by the proponents of pantomimic scenarios. Iconicity is definitionally challenging but in the most general and theory-neutral sense it can be described as the type of representation that depends on some inherent resemblance between the form and intended meaning. Visual representations, including both gestures and pantomimes, are iconic when they “resemble” what they stand for, by means of schematically illustrating the salient structural properties of the represented objects or actions (e.g. a spiraling hand-movement

representing the action of a helicopter's rotor; Müller, 2014; see also Żywicznyński et al., 2018).

A longstanding argument in language evolution is that the iconic potential of visual representations is much greater than that of vocal ones (such as non-linguistic vocalizations), which makes visual representations more “natural” and thus better suited for getting linguistic communication off the ground (Stokoe, 1991; Armstrong & Wilcox, 2007). This line of thinking has recently gained support from empirical studies on novel communication systems (e.g. Fay, Arbib, & Garrod, 2013; Zlatev et al., 2017; although trace iconicity is also present in nonlinguistic vocalization, Perlman, Dale, & Lupyan, 2015). The proponents of gesture-first scenarios elaborated arguments that specifically focus on the iconicity of gestures in the narrow sense, e.g. by showing how the shapes of gesturing hands resemble intended meanings or hypothesizing that gestures may have formed the rudiments of grammar, with the hands and arms instantiating the role of quasi-nouns and hand and arm actions, that of quasi-verbs (Stokoe, 1991; Armstrong, Stokoe, & Wilcox, 1995).

In the case of pantomime, iconicity is closely related to the use of the whole body to represent events. For example, a pantomime of stone knapping involves the mime performing a simplified version of that motor routine without the functional effect, i.e. without actually knapping any stones (Gärdenfors, 2017). Importantly, whole-body representation involves much more *primary iconicity* (Sonesson, 1997) than manual-only representation, and so whole-body representations are less conventional and less schematic but more transparent: compare using the whole-body versus the index and middle fingers to communicate the action of walking. Pantomime understood in this way is a form of mimetic communication.

Pantomime-First Versus Gesture-First Theories: Summary

On the level of theoretical commitments, there are extensive similarities between the pantomime-first and gesture first theories, but also sufficient differences that the two point to

different bodies of potential corroborating or falsifying evidence (see also Zlatev et al., 2017 for pantomime-first versus prototypical multimodal-first). Their main similarity lies in their reliance primarily on the visual modality and on the natural iconic grounding that it provides to signals perceived visually. Here, the difference is in the degree, because this requirement is not as strict for pantomime-first – especially of the mimetic profile – which allow for a significant role of non-linguistic vocalization.

A major difference lies in the emphasis of gesture-first theories on manual signaling, where an important commitment of pantomimic theories is in the use one's entire body (with a nontrivial role of facial expression). This is a very significant difference with further consequences; for example, it makes pantomime highly expressive, but also difficult to produce and resistant to high-fidelity copying, and last but not least, energetically costly – which, as one example, makes pantomime suitable for implementing costly rituals (e.g. Power, 2009, p. 271). A major empirical difference relates to the ape baseline: non-human apes have very rich gestural communication but no pantomime (with possible rare exceptions: Russon, 2018).

In sum, pantomime-first accounts form a coherent and theoretically useful category under two conditions:

- pantomime fits a *narrow definition*, clearly distinguishing it from related notions (most importantly gesture);
- pantomime has *sufficient weight*, i.e. functions as a pivotal rather than marginal element of that theory.

These two conditions are not met in the works of gesture-first theorists such as Armstrong and Wilcox (2007) or Hewes (1996), who do occasionally refer to pantomime, but the bulk of their argumentation relates to manual communicative action. Likewise, Corballis (2002a, b) explicitly refers to manual gesture and not to pantomime, and much of his argumentation is

grounded in neuroscientific data showing links between the neuronal representation of language or orofacial praxis and the cerebral control of the upper limb. In later works, Corballis (2014a, b) uses “pantomime” more frequently, but it is not clear whether this word choice is meant to have theoretical consequences.

To a large extent, the above two conditions are met in several historical accounts and more contemporaneously in the works of Arbib (2012) and Tomasello (2008) as well as in the mimetic scenarios (esp. Donald, 1991; Zlatev, 2008). We review these in turn in the following sections.

Pantomime in Language Origins: Historical Accounts

The definitional problem of separating gestural from pantomimic scenarios was also present in traditional reflection on language origins. Giambattista Vico (1668–1744) is often identified as the father of gestural approaches (see Hewes, 1977). But what he describes in *Scienza Nuova* (*The New Science*) as the earliest form of human communication – the so-called language of gods – in fact depends on a variety of visual semiotic resources, including pantomime, which are used to convey concepts and relations between them.

The philosophers and philologists should all have begun to treat of the origins of languages and letters from the following principles, (i) That the first men of the gentile world conceived ideas of things by imaginative characters of animate and mute substances. (2) That they expressed themselves by means of gestures or physical objects which had natural relations with the ideas; for example, three ears of grain, or acting as if swinging a scythe three times, to signify three years. (3) That they thus expressed themselves by a language with natural significations. (Vico, 1725/1948, pp. 125–126)

Today, it often goes unnoticed that later during the Enlightenment, there emerged a near consensus about the nature of pre-linguistic communication. Accordingly, it was believed that

the first people had used the bi-modal system of communication, which combined non-linguistic vocalization, usually of emotive character, with expressive body movement (Żywicznyński, 2018). The model version of this proposal was presented in the form of a thought experiment about pre-linguistic children who are isolated from the rest of humanity and have to discover language anew. The experiment shows how the first pair and their progeny gradually re-invent language. Although the first version of the experiment came from Bernard de Mandeville (1728), this is rarely mentioned by commentators. They instead focus on the formulation by Condillac (1746/2001), who portrays the children's initial communicative attempts in the following way:

When they lived together they had occasion for greater exercise of these first operations, because their mutual discourse made them connect the cries of each passion to the perceptions of which they were the natural signs. They usually accompanied the cries with some movement, gesture, or action that made the expression more striking. For example, he who suffered by not having an object his needs demanded would not merely cry out; he made as if an effort to obtain it, moved his head, his arms, and *all parts of his body*.³ (pp. 114–115)

Just as Vico, Condillac is regularly portrayed as one of the forefathers of the gestural scenario of language emergence (e.g. Fitch, 2010). However, as the above fragment suggests, what he had in mind when writing about the precursor to language was whole-body multimodal expression, and not isolated expression of the hands and arms. The proposal that language arose out of the communication based on emotional vocalization and pantomime gained huge popularity in the 18th century, and was championed by such influential thinkers

³ Condillac explains that the usual context of these communicative attempts was absence of desired objects (e.g. "One day the sensation of hunger made these children call to mind a tree loaded with fruit which they had seen the day before", Condillac, 1746/2001, p. 114); hence, the expressive movements of "head, arms, all body parts" should be interpreted as pantomime rather than pointing to physically present objects.

as Rousseau, Diderot, Voltaire, Helvétius and the influential Parisian milieu known as *Les Idéologues* (Żywiczyński, 2018).

The type of reflection on language origins that flourished in the Enlightenment is sometimes referred to as naturalistic glottogeny (Żywiczyński, 2018). The term serves to indicate that it was focused on the problem of how humans could have discovered language on the basis of their own – natural – dispositions. As shown by the Mandeville-Condillac thought experiment, these thinkers conceived of pre-linguistic humans as identical to modern humans in terms of their cognitive and social characteristics. Hence, the identification of pantomime as the form of communication used by pre-linguistic people is tantamount to the statement that pantomime is a natural form of human expression. “Natural” means here that it does not depend on learning but appears, to use Condillac’s phrase, “by instinct alone”.

Although naturalistic glottogeny relied mainly on speculation, it did appeal to the empirical evidence that was available at the time, including the evidence afforded by the rehabilitation of feral children and by the early work on sign language. Regarding the former line of evidence, there were two well studied cases – of Peter the Wild Boy (c. 1711–1785) described by Monboddo (1774), and Victor of Aveyron (c. 1788–1828) rehabilitated by Itard (1802). They were often seen as supporting the thesis about naturalness of pantomimic communication: both Adam and Peter were adept at using full body and gestural communication, but even with intensive and planned training they had failed to learn spoken language (Żywiczyński, 2018). Similarly, successful attempts to design sign languages based on deaf communicators’ own communicative movements were interpreted in this way. The strongest claim was made by Pierre Laromiguière, a student of Condillac’s, who considering sign languages and reports on how European travelers communicate with newly discovered populations, concluded that pantomimic communication was innate to humans:

The knowledgeable and the ignorant, everyone understands it, everyone speaks it. Let one of us be transported to the extremities of the globe in the midst of a horde of savages. Do you think that he will not be able to express the most pressing needs of life? Do you think he can mistake the signs of a barbarous refusal or the sign of a generous and compassionate intention? Therefore, there is no question of inventing a language: it already exists made for us by nature. (1826, III, 113; as quoted in Knowlson, 1965, p. 507)

This conclusion led him to postulate that a universal language should be based on this innate capacity to communicate by means of body movements.

With the end of the Enlightenment, naturalistic glottogeny subsided, but the idea that pantomime constitutes a natural form of human expression lived on. Exemplary here is Edward Tylor's conception of natural language (1881). Using the familiar lines of evidence – sign languages and feral cases – complemented with anthropological data, he argues that pantomimes and emotive cries constitute a natural system of communication, because these two semiotic resources rely on natural signs, i.e. there is an evident connection between the form and the referent (Tylor, 1881; Mocerino, 2016, p. 74). In the case of pantomime, this evident connection consists in the iconic similarity between pantomimic movements and objects, actions or – most commonly – whole events⁴ they stand for.

What does historical language origins research tell us about pantomimic scenarios? Abstracting away from the specific lines of evidence described above, they are all based on the observation that pantomime is used by humans when, for a variety of reasons, they cannot use language for communication. It is this element that makes pantomime-first such an attractive proposal. A modern incarnation of this view is found in the tradition of

⁴ The holistic nature of pantomime makes it a good means of communicating events and by the same token leads to the underspecification of individual event components such as agents, patients or actions (see e.g. Zlatev et al., 2017).

experimental semiotics, which studies the emergence of novel communication systems in the laboratory (for an overview see Galantucci, 2017). Many of these studies use "silent gesture", i.e. silent, iconic depictions of individual concepts by means of one's hands and arms; e.g. Fay et al., 2013; Schouwstra & de Swart, 2014. The strong position experimental semiotics has acquired in the science of language evolution intensifies the discussion about pantomime and gesture as a potential starting point of language (see e.g. Roberts, Lewandowski, & Galantucci, 2015). It should be noted that for both the traditional and modern semiotic-experimental lines of evidence supporting the pantomimic origin of language, the departure point are the socio-cognitive endowments of contemporary people. Hence such arguments can only explain processes whereby pantomime culturally evolves into symbolic communication, but they do not shed light on how we came to possess these endowments (see below for Abib's scenario).

Pantomime in Language Origins: Contemporary Accounts

In current language evolution research, the importance of pantomime is revived in two highly influential accounts of language origins: by Michael Arbib and by Michael Tomasello. Arbib's successively updated Mirror System Hypothesis (MSH – Arbib, 2005, 2012, 2016; Arbib et al., 2018)⁵ deserves particular attention: not just for serving as a blueprint for pantomimic theories, but also for being perhaps the most complete and mature language origins scenario. MSH assigns a separate stage for pantomime, where it arises from communicative use of intransitive manual action and later develops into conventionalized signs. Tomasello (2008) proposes that the emergence of the language faculty was bootstrapped by two types of "natural human gesture", used to direct attention (pointing) and to direct imagination (pantomime), and although pointing receives the primary focus, pantomime remains an essential complement. An important third group are bodily-mimetic

⁵ "However, MSH is not a fixed dogma but, rather, an evolving system to be updated as new data and theory become available" (Arbib, 2018a, pp. 131–132).

accounts of language origins by Merlin Donald (1991, 2001) and Jordan Zlatev (2008), whose reliance on primarily visual, whole body expression makes them de facto pantomimic (even if the word “pantomime” is rarely used there, except in Zlatev et al., 2017). The recent pantomimic proposals by Gärdenfors (2017) and Ferretti et al. (2017) are indebted to the mimesis theory.

Arbib: The Mirror System Hypothesis (MSH)

Definition. Pantomime is a vital element in Michael Arbib’s (2012, 2016) Mirror System Hypothesis – a complex model of transition to language from the “baseline” of LCA-m, the human Last Common Ancestor with monkeys, and LCA-c (with chimpanzees). Does pantomime play a sufficiently robust role for MSH to be a truly pantomimic rather than a gestural theory? Arbib’s account lacks a thorough and systematic terminological discussion in this point. Because of the neuroscientific profile of MSH and the fact that it builds on the action of grasping and other mostly⁶ manual praxic actions, pantomime starts “lean” and “manual”, as intransitive execution of a normally mono- or di-transitive action, such as reaching for a tool. This may lead to blurring the distinction between pantomime and gesture. However, Arbib’s definitions and examples suggest that the word “pantomime” is intended in a meaning close to its intuitive sense. Arbib describes pantomime as “the ability to use reduced forms of actions to convey aspects of other actions, objects, emotions, or feelings—the artless sketching of an action to indicate either the action itself or something associated with it” (Arbib, 2012, p. 177). Pantomime “involves expressing a situation, object, action, character, or emotion without words, and using only gestures, especially imitative gestures, and other movements” (2012, p. 217). This is further clarified through examples, including the frequently used example of flapping one’s arms to signify a bird or flying. It appears that

⁶ But not exclusively - so e.g. orofacial praxic and communicative actions are also considered, e.g. Arbib, 2012, p. 127.

“pantomime” is a broad concept for Arbib, which includes iconic gestures but transcends this category in theoretically consequential ways.

The Mirror System Hypothesis (MSH). MSH holds a particular significance for language origins research, not by being any less controversial than its alternatives, but through its remarkable theoretical completeness and the wide range of interdisciplinary data on which it is based. MSH proposes that language arose in a series of stages via *biocultural* evolution: first primarily biological, through which the human “language-ready brain” developed on top of the primate mirror neuron system for action recognition and imitation; then mostly cultural evolution, which transformed the first protolanguages into the full languages of the present (Table 1).

Table 1

Stages of the development of language in Michael Arbib’s Mirror System Hypothesis (based on Arbib, 2005, 2012, 2016). LCA-m, LCA-c = human Last Common Ancestor with monkeys, and with chimpanzees.

1. Mirror Neuron System for grasping and manual praxic actions (LCA-m)
2. Simple imitation (LCA-c)
3. CAR&IM: Complex action recognition and complex imitation
+ communicative intentions, “symbolization”, cultural innovation (possibly latent in LCA-c) =>
- 4. Pantomime**
 - a) of grasping and manual praxic actions, then =>
 - b) of actions outside of own repertoire
=> conventionalization =>
5. Conventional gestures

- a) to enrich and disambiguate pantomime
 - b) to replace pantomime (protosign)
6. Protospeech and multimodal holistic protolanguage
- > fractionation, grammaticalization -->
7. Full human languages (spoken and signed)

The foundation for pantomime: CAR&IM. The first of these stages (see Table 1) was the extension of *simple imitation* already present in LCA-c (as inferred from its presence in extant apes) into *complex action recognition* and *complex imitation*. Whereas simple imitation just makes it possible to repeat an established motor routine or else employ trial and error to achieve one observed sub-goal from another (Byrne, 2003), complex imitation augments this by the ability to imitate aspects of observed movements even if they are not part of the imitator's current stock of actions, thus introducing new variants of actions to one's own repertoire, or "praxicon". Complex action recognition/imitation has its roots and immediate application in praxis, but provides sufficient neuro-motor basis for pantomime. Arbib (2012, 2016) names two additional cognitive conditions for such an extension from praxis to communication - *symbolization* (in a loose sense of learning novel sign-meaning associations) and *communicative intentions*⁷- but claims they are present in the LCA-c at least in a rudimentary form. A final likely requirement is "*cultural innovation* that supports both the creation of novel pantomimes <on the fly,> and the ability to recognize that a novel behavior is indeed an attempt to communicate" (Arbib, 2018).

Creative open-endedness. For Arbib, the watershed of pantomime is *open-ended semantics*: "a breakthrough at the level of the freedom to create novel associations" (2012, p. 261). Following Stokoe and others, Arbib (2012, p. 219) notes that pantomime has "the ability

⁷ Communicative intentions and more broadly, the motivation to send and receive signals are often taken for granted, including in MSH. However, under standard circumstances such intentions and motivations will make organisms *less* rather than more evolutionarily successful (e.g. Waciewicz & Żywicznyński, 2018), and so their evolutionary appearance is a "central puzzle" in language evolution research (cf. Fitch, 2010).

to create an open-ended set of complex messages exploiting the primates' open-ended manual dexterity", thus extending the relatively closed repertoire of ape gestures to a potentially unlimited repertoire of pantomimic meanings. The potential to flexibly introduce novel signals for novel messages is closely related to other gains in expressive power, i.e. semantic universality and displacement. Pantomime is relatively universal in that it can communicate about many semantic domains (rather than being restricted to e.g. only predator evasion or food), and it can certainly express displaced meanings, i.e. about entities not present in the immediate here and now. "Pantomime, then, allows the transfer of a wide range of action behaviors to communication about action and much more—whereby, for example, an absent object is indicated by outlining its shape or miming its use" (Arbib, 2012, p. 177).

However, pantomimic communication has considerable drawbacks. Without the normative aspect that comes with conventions, pantomimes are spontaneous and impromptu. Arbib (e.g. 2012, 2013) repeatedly makes this point by stressing that first pantomimes are "ad hoc", "artless" or "naïve". Their one-off, unstandardized nature implies that signs must be coined on the spot and interpreted on the spot, rather than simply retrieved from memory, so the creation and comprehension of these basic units take place in the real-time dynamics of the communicative situation. This implies rather substantial costs in terms of time, cognitive effort, communicative effectiveness, and replication fidelity. Such inefficiency motivates the appearance of conventional gestures, first as additions that enrich and disambiguate the pantomimic repertoire but later as more efficient replacements for pantomimes (Arbib, 2012, p. 226). Through conventionalization, the next transition is achieved: from pantomime to protosign.

Conventionalization. Conventionalization (also symbolization, Sulik, 2018) is the process through which rich iconic signs (gestural, pictorial, or otherwise) are transformed in the course of interaction into conventional signs. Iconic signs are based on resemblance, so

their meanings can be guessed from their form through creative inference, but the meanings of conventional signs are fixed to particular forms on the strength of a tacit or explicit agreement, and can be simply retrieved from memory without any need for guessing. Such signs typically, but not necessarily, lose much or sometimes all iconicity in the process and are streamlined into forms that are simpler, standardized and much easier to produce. Like many other authors (e.g. Hutto, 2008; Corballis, 2014a), Arbib notes that pantomimic/gestural forms constitute a perfect substrate for conventionalization, while the inefficiency of pantomime, already mentioned above, provides the motivation. “Pantomime can be both laborious and highly ambiguous, and so the next step is the conventionalization of pantomimes and the addition of otherwise arbitrary gestures to yield a system (protosign) that provided economical and less ambiguous symbols” (Arbib, 2016, p. 13).

Conventionalization is where the pantomimic scenarios have a clear advantage over their competition. Human languages critically depend on semiotic conventions (e.g. de Saussure, 1916), and since these are absent from other animal systems of communication, their origin is a key explanatory target in language evolution research. A major strength of the pantomimic scenarios is that the path to language via pantomime and its subsequent conventionalization provides a solution to this problem that is much more compelling than in any alternative theories.

Equally important and valuable here is the evidential basis, because the conventionalization of iconic signs is one of the few areas in language evolution research where direct empirical evidence exists. There are two principal lines of this evidence – the study of emerging sign languages and experimental semiotic research. Regarding the first of these there is ample of evidence that proto-signs of an emerging sign language are motivated, communicatively complex, and often involve more than just the hands and arms. As a result, communication during this proto-phase is slow, laboured and abounds in redundancy. The

existence of this initial stage is confirmed by both historical reports, for example Cucurron Sicard's treaty on the development of the French sign language, and contemporary ones – on the development of the Nicaraguan and Al-Sayyid Bedouin Sign Languages (Senghas et al., 2005; Sandler, Meir, Padden, & Aronoff, 2005; see also Klima & Bellugi, 1979 and Kendon, 2004, p. 309). This initial stage does not last long: emerging signs quickly become reduced in form, and they gradually lose their iconicity, acquiring instead linguistic properties such as morphological structure. The strong pressure for conventionalization of iconic forms is also documented by experimental semiotic studies. Research on the emergence of novel communication systems in the laboratory (Roberts et al., 2015; Scott-Phillips & Kirby, 2010), including with improvised “silent gesture” (Fay et al., 2013), shows that when such de novo representations are set in an interactive context, their form becomes increasingly economic, and the motivated connection between their form and meaning quickly gives way to fixed, conventional forms (including gestural conventions, e.g. Motamedi, Schouwstra, Culbertson, Smith, & Kirby, 2017).

Parity and the neural grounding. Pantomime is perhaps the most central step in the entire succession of stages, because it serves to complete several key transitions, including the abovementioned emergence of *open-endedness* and then *conventionality*. But equally important to Arbib is the transition from practical to communicative action on the level of neuronal mechanisms. As we know from the work on mirror neurons (e.g. Rizzolatti & Arbib, 1998), the Mirror Neuron System in the primate brain is activated when the primate performs an action such as grasping, but also when it sees a similar action performed by another. Arbib (e.g. 2005, 2012) observes that this ability to “translate” between production and comprehension is also a fundamental requirement in communication: the so-called *parity principle*, whereby the same signal counts for more or less the same meaning to both the producer and the receiver of this signal. Arbib's MSH thus substantially adds to other

pantomimic/gestural models in that it provides a “neural missing link” from the primate manual/bodily dexterity to prelinguistic communication with pantomimes and gestures (see esp. Arbib, 2016, 2018a for more details on the cerebral implementation).

Tomasello: Pointing and Pantomiming

Pantomime and pointing. The element of pantomime is central to Michael Tomasello’s (2008) model of language emergence. However, commentators tend to highlight the other key features of his account, such as prosociality and shared intentionality, or otherwise label Tomasello’s model as ‘gestural’. This is understandable given that Tomasello’s very extensive empirical evidence, gathered from both comparative and developmental (i.e. humans – other apes; human children – adults) studies, mostly relates to *pointing*. Tomasello proposes declarative pointing, and especially informative-declarative pointing (i.e. pointing performed with the intention of providing the recipient with new information) to be the first step distinguishing human ancestors from the generalized LCA-a baseline.

However, pantomime as such also has an important role to play. Tomasello (2008) uses this term in a way that is rather loose but nonetheless corresponds quite closely to the everyday, intuitive meaning:

The second type of human gesture used as a complete communicative act is iconic gestures or pantomimes (depictive, imagistic, characterizing, representational, and symbolic gestures are other terms that have been used) ... In using an iconic gesture the communicator enacts some action with her hands and/or body..., and this is intended to induce the recipient to imagine some corresponding perceptually absent referent..., for example, an action the communicator wants the recipient to perform or an object he wants her to fetch. (p. 66)

This definition and its extensions, as well as his examples (such as miming the grating of cheese on pasta, or the use of a chainsaw; Tomasello, 2008, pp. 67, 68) clearly indicate that a “rich” interpretation of pantomime is intended. Tomasello’s pantomimes are semantically advanced communicative acts, in which the communicator uses the whole body (rather than just the upper limbs) to deliver holistic messages with substantial expressive power.

Action-oriented, standalone, displaced. Characteristically, Tomasello uses the gerund form, “pantomiming”, preferentially to the noun “pantomime”. This is because on his account, pantomiming is action-oriented and dynamic, and it is employed “(i) to indicate that *this* is the action I want you to perform, or that I intend to perform myself, or that I want to tell you about; and (ii) to request or otherwise indicate an object that “does *this*” or an object that “one does *this* with” (Tomasello, 2008, p. 67). Thus, pantomimes do not seem to have internal morphological structures analysable into discrete component parts; rather, pantomiming is a continuous holistic process. Similarly, pantomimes themselves are not replacements for words, but correspond to larger units, at least proposition-size. Words can only complete communicative acts via being combined with other words, but pantomimes are “standalone”, in that a single act of pantomiming can serve as a complete communicative act with its own illocutionary force.

For Tomasello, the special importance of pantomime, and the breakthrough that it accomplishes, appears to be *displacement*, i.e. communication beyond the immediate here-and-now (Hockett, 1960). In fact, enabling displacement makes pantomime the next logical step after pointing. Although pointing is the foundation of human declarative communication, its use is typically limited to situations where the intended referents are perceptually available and salient. Pantomimes, being much more referentially specific, are not subject to similar constraints: “I can only point and intend to indicate a rabbit that is not currently perceptible in

very special circumstances, but I can pantomime an absent rabbit with the same intention quite easily” (Tomasello, 2008, p. 233).

Socio-cognitively complex. Despite numerous references to the “naturalness” of pantomime (e.g. 2008, pp. 59, 172), Tomasello recognizes that pantomimic communication only becomes possible when embedded in a rich set of nontrivial socio-cognitive capacities. Firstly, understanding pantomimes presents a difficulty even to contemporary humans, and the property of displacement in particular makes pantomimes even more challenging: “they depend, in a way that pointing does not, on skills involving some kind of imitation, simulation, or symbolizing” (Tomasello, 2008, p. 67). The advanced cognitive-inferential processes required to comprehend pantomimes critically rely on sufficient *common ground* - a framework of knowledge, attentional states, background assumptions, etc. that transcend the egocentric perspective and are shared between the communicators (Tomasello, 2008, pp. 73–88). Even more importantly, pantomimic communication presupposes certain uniquely human social predispositions, such as a natural tendency to cooperatively share honest information with non-relatives (Tomasello, 2008, pp. 85–97). Since cooperative signaling is evolutionarily anomalous and is indeed very rare in animal communication, the human propensity to communicate cooperatively can by no means be taken for granted. Nonetheless, the major explanatory challenge that it poses is equally relevant to all theories of language origins, pantomimic or otherwise (cf. Waciewicz & Żywicznyński, 2018).

Donald and Zlatev: Mimesis

Merlin Donald. A very influential position on the evolution of early, protolinguistic capabilities was formulated by Donald (1991, 2001), and developed by other researchers (most notably Zlatev, 2008, 2014; Zlatev et al., 2017; but see also Gärdenfors, 2017 and Żywicznyński et al., 2018). Donald's proposal was part of his theory of human cognitive and cultural evolution that suggests three stages going from episodic to mimetic to mythic culture.

Episodic culture is characteristic of the life of early hominins and is similar to the behavior of non-human apes. Whereas episodic cognitive culture is concerned with the representation of concrete situations or episodes (1991, pp. 148–153), *mimetic culture* breaks with that tradition and is facilitated by the emergence of abstract representations, something that Donald posits to have occurred with the rise of *Homo erectus*. The defining component of mimetic skill or mimesis is "the ability to produce conscious, self-initiated, representational acts that are intentional but not linguistic" (1991, p. 168). In this way, it is the symbolic nature of mimetic acts that separate them from even complex forms of imitation (see Arbib above). Donald understands mimesis widely as both a cognitive and communicative innovation. It should be stressed that mimesis does not necessarily involve social communication, but can be used for retrieving a particular memory (i.e. mimetic representations serve as internal cues to aid long-term memory, 1991, p. 148; see also Donald, 2001) or rehearsing and practicing a skill such as boxing drills or dancing steps (1991, p. 172).

When it comes to mimetic communication, Donald enumerates a variety of semiotic resources - "[t]ones of voice, facial expressions, eye movements, manual signs and gestures, postural attitudes, patterned whole-body movements of various sorts" (1991, p. 169). He further identifies pantomime and ritual dance as purely mimetic means of communication that are present in contemporary human cultures. However, speaking generally, *mimetic communication is any communicative act in which the body is intentionally used to represent something* (i.e. stand for something other than itself, 1991, pp. 171–173). He also stresses the capacity of mimesis to communicate about an unlimited repertoire of referents, most importantly events. On such an account, pantomime seems to lie at the heart of Donald's mimetic communication. Donald understands mimesis as a precursor to language which emerges in the next stage - *mythic culture*, and his scenario of language evolution can therefore be classified as pantomimic.

Jordan Zlatev. Zlatev (2008; Zlatev et al., 2005) developed Donald's mimesis to build a detailed account of language emergence. For Zlatev, the evolution of language is seen as the evolution of multi-layered socio-cognitive capacities, in which the earlier ones co-exist with the later additions (Zlatev, 2008). Zlatev puts focus on definitional problems, particularly with respect to the foundational concept of bodily mimesis. Accordingly, an act counts as bodily-mimetic if it meets the following criteria (Zlatev, 2008):

- (a) It involves a cross-modal mapping between exteroception (i.e. perception of the environment, normally dominated by vision) and proprioception (perception of one's own body, normally through kinesthetic sense).
- (b) It is under conscious control and corresponds - either iconically or indexically - to some action, object or event, while at the same time being differentiated from it by the subject.
- (c) The subject intends the act to stand for some action, object or event for an addressee (and for the addressee to recognize this intention)
- (d) Without the act being conventional-normative.
- (e) Without the act dividing (semi)compositionally into meaningful sub-acts that systematically relate to each other and other similar acts. (p. 219)

He links this definition to what he calls the mimesis hierarchy (Table 2), which is an evolutionary sequence of mimetic skills (Zlatev, Persson, & Gärdenfors, 2005; Zlatev, 2008). The rudimentary form of mimesis - *proto-mimesis* - is based on the requirement (a) (see above) and is linked to such communicative activities as emotional and attentional contagion or automatic type of imitation (e.g. neonatal imitation). The more advanced form of *dyadic mimesis* involves volition and representation, for example present in more complex imitation (see the requirement (b) above), but only at the next level, referred to as *triadic mimesis*, mimetic acts gain a decidedly communicative function. As stipulated by the requirement (c), a

triadically mimetic act requires the communicative intention, i.e. that a producer's representation is understood as standing "for some action, object or event for an addressee". Finally, *post-mimesis* is explained as the positive versions of the requirements (d) and (e), i.e. triadic acts that are characterized by conventionality and normativity (d), and by semiotic systematicity (e). The endpoint of post-mimesis is spoken or signed languages.

Table 2

Zlatev's mimesis hierarchy

1. Proto-mimesis: rudimentary form of mimesis involving contagion or automatic form imitation;
2. Dyadic mimesis: volition and representation required for complex imitation;
3. **Triadic mimesis**: communicative intentions required for **Pantomime**;
4. Post-mimesis: conventionality, normativity and systematicity.

Crucial for Zlatev's idea of mimetic communication, including pantomime, is the requirement (c) and the corresponding notion of triadic mimesis. He enumerates pantomime sensu Arbib (2005) and gesture sensu Corballis (2002a) as well as iconic gestures and declarative pointing sensu Tomasello (2008; see 4.2 above) as types of triadic mimesis (2008). He also illustrates it with what in our definitional framework are prototypical examples of pantomime, e.g.: "An example of an iconic sign that fulfils all three conditions (a)-(c) is the miming of eating by pretending to move a spoon to one's mouth (e.g. made behind a glass door) in order to communicate to a colleague that it is time for lunch" (Zlatev, 2008, p. 138). Accordingly, pantomime could be defined as an act of bodily communication, in which a bodily action is volitionally used to represent something else than the bodily action itself, i.e. to stand for an object, action or an event, and in which the addressee understands

this intention. Furthermore, pantomime should be free from semiotic conventions and holistic, in the sense that it does not naturally decompose into lower level-units.

In a recent empirical paper Zlatev (Zlatev et al., 2017) accepts Żywicznyński and colleagues' (2018) definitional criteria for pantomime (see 2.2. above) and uses them to formulate a pantomime-first scenario of the emergence of language. It hypothesizes that early form of mimetic communication, although using a variety of modalities, primarily relied on the bodily-visual modality, due to its greater potential for iconic (resemblance-based) representations (Zlatev et al., 2017, p. 457). In this way, Zlatev and colleagues differentiate the *pantomime-first* position from what they call the multimodal-first positions, represented by McNeill (2012) or Kendon (2014), who stress an early integration and mutual development of bodily-visual and vocal-auditory communication.

Related conceptions.

Gärdenfors. Gärdenfors focuses on the evolution of pedagogy, particularly in the context of tool-making (see Gärdenfors & Högberg, this volume). In doing so, however, he looks at a suite of cognitive adaptations that are of great interest to the science of language evolution. In his opinion, the core capacities are mind-reading skills and auto-rehearsal. Mind-reading abilities (Premack & Woodruff, 1978) include what Gärdenfors calls "cooperative forms of mind reading" – joint attention and joint intention (cf. Tomasello, 2008). In explaining auto-rehearsal, he appeals to Donald's mimesis theory (see above) and underlines the role of recall of previous performance for both learning complex activities (such as tool making) and teaching them to others (which is linked to Sterelny's notion of the apprenticeship culture, Sterelny, 2012b; see also Stout, 2018). An analysis of early stone-tool cultures leads him to conclude that beginning with the Oldowan toolkit, teaching became necessary to effectively transmit the technology to other individuals and to maintain it in a population (Gärdenfors & Högberg, 2017). The principal forms of teaching that this

transmission process relied on were two types of enactment – demonstration and pantomime.

For pantomime, he specifies the following criteria (Gärdenfors, 2017):

- (P1) The mimer performs the movements of the actions in the task without actually performing the actions.
- (P2) The mimer makes sure that the learner attends to the series of actions.
- (P3) The mimer's intention is that the learner can perceive the right actions in the correct sequence.
- (P4) The mimer exaggerates and slows down some of the actions in order to facilitate for the learner to perceive important features.⁸

The notable difference between pantomime and demonstration concerns the first point - whereas in demonstration the teacher actually performs the actions involved in the task (e.g. actually knaps a stone), in pantomime the teacher pretends to perform the actions ("Pantomime is a form of pretense", Gärdenfors, 2017), by using "more or less simplified versions of them" (pretends to knap a stone, Gärdenfors, 2017). Contexts, other than teaching, in which pantomime is used include narration and autocued rehearsal, e.g. practicing stone knapping actions without performing the activity of stone knapping (see above, 2.4).

What is the connection between pantomime used for teaching and the hypothetical role pantomime could have played in the emergence of language? Gärdenfors links the history of pantomime to the evolution of mimesis. Using Zlatev's mimesis hierarchy, he describes pantomime as belonging to the third, triadic level of mimesis, i.e. such that has communicative sign function (Gärdenfors, 2017; Zlatev et al., 2005). In consonance with Arbib (2012), Żywicznyński et al. (2018) and Zlatev et al. (2017), Gärdenfors (2017) argues that pantomime was a precursor of language, used to refer to events as pantomimically

⁸ In line with Stout (e.g. 2018), it could be argued that active teaching of this sort constitutes "structured and meaningful community of practice", whereby learners are not just equipped with specific information but also with the motivation to learn specific things in specific ways.

represented holophrases (i.e. signs having the form of unanalyzed expressions; see 2.2) that were later broken down into smaller semantic elements. Notably, the pedagogical use of pantomime as demonstration starts “lean”, from communication of procedural “knowledge-how” information, but has a potential to scale up to richer uses, i.e. communicating semantic-propositional information, or “knowledge-that”.

Other accounts. Ferretti et al. (2017; see also Ferretti, this volume) mention pantomime in the context of their narrative account of language origins. Following Corballis, they identify Mental Time Travel (MTT) as the cognitive mechanism indispensable to global coherence, which - on their view - allows for the production and comprehension of narrative (Ferretti et al., 2017; Adornetti, 2015). While the emergence of MTT was a cognitive precondition for language, the crucial requirement of the narrative account of language evolution was the presence of a pre-linguistic communication system that could have been used to “tell” stories. Accordingly, they argue that pantomime as defined by Żywiczyński et al. (2018) meets the requirements of such a system, stressing its communicative complexity and self-sufficiency, which translates into the ability of pantomime to describe whole events in the absence of conventional semiotic resources.

The fact that pantomime is often seen as an important sub-type of the communicative manifestation of mimesis makes many of these accounts that stress the role of mimesis in the emergence of language interpretable as pantomimic scenarios. Harnad (2000) sees pantomime as part of a continuum in language origins: praxis → pantomime → propositions. Hutto (2008) underscores the potential social role of pantomime in ceremonies and other group-specific practices:

Regular re-enactments of events of special significance may have eventually become deeply ingrained in the social fabric, thus supporting the establishment of common customs and habits ... [becoming] a powerful substitute means of ensuring social

cohesion, supplanting or at least supplementing the physical grooming of individuals ... [and they] would have helped to solidify within-group identities.” (p. 261)

Steven Mithen’s *Hmmmm* proposal (2005) sees language as emergent from the holistic, manipulative, multimodal, musical and mimetic mode of communication; and although on his account early meaning-making employs a wide polymodal suite of semiotic resources, to the extent that visual pantomimic displays have a key role in this mix, Mithen’s scenario can be classified as a broadly pantomimic scenario. These last accounts underscore the continuity between pantomime and other manifestations of mimesis, such as ritual, dance and music. This continuity is exemplified in some cultural practices found in contemporary hunter-gatherer tribes (e.g. Lewis, 2013).

Finally, we note the recent account by Brown et al. (2019), which stands out by distinguishing several subtypes, or modes, of pantomime. Brown et al. (2019) identify two main such modes: egocentric pantomiming that refers to one’s own peripersonal space (similar to “character viewpoint”) and allocentric pantomiming about referents in objective space, as experienced by an external observer (similar to “observer viewpoint”). Brown et al. link the relative advantages of the egocentric mode to a People-First model of language origins, which prioritizes communication about humans, and the advantages of the allocentric mode to an Environment-First model, i.e. communication about the ecological environment.

Criticisms

The criticisms directly targeting pantomimic conceptions of language origins (Abramova, 2018) or indirectly relevant to them (e.g. Cartmill & Goldin-Meadow, 2012) can be divided into two broad groups. Firstly, pantomime is a mostly visual means of communication, which leaves unexplained the dominant role of vocal-auditory channel in the languages of today (sign language and co-speech gesture notwithstanding, see Goldin-Meadow, this volume). Secondly, critics take issue with the iconic grounding of meaning in

pantomime, suggesting in particular that despite claims about intuitiveness and “naturalness”, iconicity is cognitively complex.

The Transition to Language

Modality shift. All theories postulating a non-vocal stage in language emergence face a fundamental difficulty: If language arose as a (predominantly) gestural/visual system, why would it now have the (predominantly) spoken/vocal form that it does – and how can we explain the extensive anatomical and neuroanatomical human adaptations to speech production? This is the so-called “modality switch” or “modality transition” problem which is commonly identified as the most difficult challenge for gesture-first theories (e.g. Burling, 2005; Corballis, 2003; Kendon, 2008; MacNeilage, 2008; Tallerman, 2011).

Other commentators (e.g. Bickerton, 2007) see this problem as ill-posed, because it relies on a framing of the gesture-speech dichotomy that is unrealistically exclusionary (see Kendon, 2011; Waciewicz, Żywicznyński, & Orzechowski, 2016). For example, even the supporters of gesture-first theories do not shy away from admitting that “there never was a time when visible gestures were unaccompanied by vocalizations” (Armstrong & Wilcox, 2007, p. 68). To what degree gestural and pantomimic theories are affected by this difficulty depends on the division of labour between vocal and visual expression postulated by a particular theory, with the pantomimic accounts based on multimodal mimesis being relatively less vulnerable in this respect:

In general, the less pre-linguistic gestural communication is thought of as a “language”, and the less modern spoken languages are conceived of as purely vocal, the less problematic the why-speech argument appears. While it is indeed damaging for scenarios that frame the transition as one “from hand to mouth” (Corballis, 2002), they are not if stated in the much less idiomatic “from body to mouth-and-body” (Zlatev, Donald, & Sonesson, 2010). That is, from whole-body communication

supported by the human-specific capacity for bodily mimesis to the multi-modal system of linguistic communication which we use today, involving both speech and gesture. (Zlatev, 2014, p. 174; see also 2.1.2)

Pantomime “repels” language. This type of criticism, influentially advocated by McNeill (2012), stresses that pantomime, unlike gesticulation, is separated from modern language use. McNeill’s attack on the view that pantomime could have constituted a precursor of language is motivated by both his definition of pantomime and his scenario of language emergence. McNeill’s influential Growth-Point theory assumes that language relies on the single dialectical unit consisting of gesture and speech. The supposition of the steadfast connection between these two types of communicative actions leads McNeill (2012) to the claim that since its dawn language has made use of both the bodily-visual and vocal-auditory modalities.

When attacking the pantomimic scenarios of language origin, McNeill (2013) in fact attacks the gesture-first theories, which accords with his definition of pantomime as gestures necessarily performed in the absence of speech. The argument runs as follows: given the co-expressiveness of gestures and speech in contemporary language use, positing that language could have evolved from gesture alone (i.e. pantomime in his dictum) is an evolutionary fiction. Without delving into the problem of the correctness of McNeill’s evolutionary logic (e.g. discarding the possibility that at the early, bootstrapping stage, the emerging protolanguage could have been very different from contemporary fully fledged language), this bears on McNeill’s conception of pantomime, which for him cannot just be accompanied by speech but in fact cannot be accompanied by any vocalization – pantomime turns out to rely on a single semiotic resource of gesturing. A broader understanding of pantomime as including a variety of semiotic resources, as accepted by e.g. Zlatev et al. (2017), would therefore blunt McNeill’s critique.

Iconicity

Some criticisms of pantomimic scenarios target the element that has often been championed by their proponents – pantomime’s robust iconicity, which is commonly appealed to when explaining the ease with which pantomime takes communication off the ground (e.g. Arbib, 2012; Zlatev et al., 2017; Gärdenfors, 2017). For example, Irvine (2016) notes that iconicity can be communicatively *disadvantageous*: successful communication requires that participants have sufficient psychological distance from the representations they use, and iconic visual presentations may be too vivid and rich in detail, making it more difficult to decontextualize them and achieve that psychological distance. Empirical results are equivocal. Some authors (e.g. Imai & Kita, 2014) show iconicity to be helpful in the ontogenetic acquisition of lexical representations, whereas others (Cartmill, 2018) remain sceptical of generalizing such results to phylogeny, or conclude that reviews of developmental research in humans and of behavioural and neuroscientific research in non-human primates fail to document a clear facilitating role for iconicity in understanding representations (Irvine, 2016; Cartmill & Goldin-Meadow, 2012, pp. 418–419). More research is certainly needed to adjudicate on this point.

Abramova (2018) argues that iconicity required by pantomime is cognitively demanding, which burdens pantomimic theories with unexplained assumptions. She notes that Arbib’s (2012) account assumes *symbolization* and Gricean *communicative intentions* as necessary elements of the cognitive infrastructure for pantomime, but these features are absent from non-human primate cognition and communication, and are themselves major explanatory targets. Arbib (2018b) responds that firstly, these features seem to be present to some extent in primate gesture, and secondly, there appear to be no convincing learner explanations. In general, Arbib, Tomasello, Donald, Zlatev or Gärdenfors all do take the point

that the emergence of pantomime required a socio-cognitive breakthrough in hominin evolution.

Finally, the expressive power of resemblance-based representations has been questioned. Pantomimic theories of language origins are committed to a “rich” definition of pantomime, since its role as a pivotal step in language emergence would seem to imply it to be an advanced means of communication. This in turn would imply considerable expressive power, such as the ability to communicate a potentially unlimited repertoire of events, or sequences of events, not limited to the here and now or to a confined number of semantic domains. It has been suggested that pantomime is not self-sufficient in this respect, and cannot express certain abstract meanings, except with the aid of other semiotic resources, such as disambiguating verbal clues (Ryan, 2012; see also Cartmill, 2018; but see Sibierska, 2017 for an argument to the contrary). A possible reply to this objection is that at least initially pantomime does not necessarily need to have large expressive power, as long as it can support the breakthrough to open-ended semantics. The limits to the expressive potential of stand-alone pantomime remain an interesting empirical question.

Summary and Conclusion

Similarly to gestural theories of language emergence, pantomimic accounts propose that language arose primarily, if not exclusively, out of visually perceived communicative action. This is their main strength: on the one hand, the natural iconicity of visual displays allows the interactants to invent novel signs whose meanings can be guessed without prior knowledge, and on the other, considerable evidence suggests pantomimes are a good substrate for conventionalization, which can turn them into symbols. Pantomimic and gestural theories of language origins indeed form a continuum and the need to separate them may not be obvious, but we show that such a distinction is theoretically productive. Firstly, while the term “gesture” is broad and ambiguous, “pantomime” (and so “pantomime-first theories”) is more

specific and informative: it is bodily-mimetic communication which is non-conventional, improvised, whole-body rather than manual-only, holistic, and communicatively and semantically complex. Secondly, gesture-first scenarios tend to focus on the hands and arms, whereas pantomime-first theories stress the importance of whole-body communicative action. Thirdly, most mimetic theories heavily rely on pantomime and although classifying them as “gestural” would be rather odd, they comfortably belong within the spectrum of pantomimic theories.

Among the most critical defining features of language are open-endedness, displacement, and conventionality. Pantomime enables the first two features (4.1.4, 4.2.2) and provides a credible entry point for bootstrapping the third: a “bridge” from icons to communicative conventions (4.1.5). Two more bridges are a “neural missing link” from practical action to pantomimic communication via the primate Mirror Neuron System (4.1.6), and the bootstrapping of the cognitively richer propositional communication with the cognitively leaner demonstration of motor routines (4.3.3.1).

Like all language origins theories, pantomimic models make background assumptions about the socio-cognitive infrastructure of our prelinguistic ancestors, and these assumptions can be questioned (e.g. Abramova, 2018). Arbib (e.g. 2012) provides a comprehensive account of the neurocognitive mechanisms involved in the transition from simple imitation to communication via pantomime, but his model takes for granted some social requirements for informative communication, such as common ground or cooperative dispositions. Tomasello’s (2008) account, not as detailed on the specifics of implementation, provides a much more complete discussion of these social requirements. As of now, the discussions remain mostly theoretical, since with rare exceptions (Zlatev et al., 2017; Brown et al., 2019) little direct empirical evidence exists to inform the evaluations of pantomimic models of language origins.

Acknowledgements: This research was supported by grant UMO-2017/27/B/HS2/00642 from the Polish National Science Centre. We are grateful to the editors of this volume, Nathalie Gontier and Christopher Sinha, as well as to Michael Arbib, for valuable and insightful comments. We thank Karolina Nowicka for her help in editing. All remaining errors and omissions are our own.

References

- Abramova, E. (2018). The role of pantomime in gestural language evolution, its cognitive bases and an alternative. *Journal of Language Evolution*, 3(1), 26–40.
- Adornetti, I. (2015). The phylogenetic foundations of discourse coherence: A pragmatic account of the evolution of language. *Biosemiotics*, 8(3), 421–441.
- Arbib, M. A. (2005). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavior and Brain Sciences*, 28, 105–167.
- Arbib, M. A. (2012). *How the brain got language*. Oxford: Oxford University Press.
- Arbib, M. A. (2013). Complex imitation and the language-ready brain. *Language and Cognition*, 5(2–3), 273–312.
- Arbib, M. A. (2016). Towards a computational comparative neuroprimatology: Framing the language-ready brain. *Physics of Life Reviews*, 16, 1–54.
- Arbib, M. A. (2018a). From cybernetics to brain theory, and more: A memoir. *Cognitive Systems Research*, 50, 83–145. <https://doi.org/10.1016/j.cogsys.2018.04.001>
- Arbib, M. A. (2018b). In support of the role of pantomime in language evolution. *Journal of Language Evolution*, 3(1), 41–44. <https://doi.org/10.1093/jole/lzx023>
- Arbib, M. A., Aboitiz, F., Burkart, J. M., Corballis, M., Coudé, G., Hech, E., ... Wilson, B. (2018). The comparative neuroprimatology 2018 (CNP-2018): Road map for research on *How the brain got language*. *Interaction Studies*, 19(1–2), 370–387.
- Armstrong, D. F., & Wilcox, S. E. (2007). *The gestural origin of language*. Oxford: Oxford University Press.
- Armstrong, D. F., Stokoe, W. C., & Wilcox, S. E. (1995). *Gesture and the nature of language*. Cambridge: Cambridge University Press.
- Bickerton, D. (2007). Language evolution: A brief guide for linguists. *Lingua*, 117, 510–526.

- Brown, S., Mittermaier, E., Kher, T., & Arnold, P. (2019). How pantomime works: Implications for theories of language origin. *Frontiers in Communication*, 4, 9.
- Burling, R. (2005). *The talking ape: How language evolved*. Oxford: Oxford University Press.
- Byrne, R. W. (2003). Imitation as behaviour parsing. *Philos Trans R Soc Lond B Biol Sci*, 358(1431), 529–536. doi:10.1098/rstb.2002.1219
- Calvin, W. H., & Bickerton, D. (2000). *Lingua ex machina: Reconciling Darwin and Chomsky with the human brain*. Cambridge, MA: MIT Press.
- Cartmill, E. A. (2018). Iconicity and convention in the manual modality: Pantomime in language origins. In C. Cuskley et al. (Eds.), *The evolution of language: Proceedings of the 12th International Conference on the Evolution of Language (Evolang12)* (pp. 73–75). Torun: NCU Press. doi: 10.12775/3991-1.014
- Cartmill, E. A., & Goldin-Meadow, S. (2012). Is pantomime a likely stage in language evolution? Evidence from human and primate gesture. In T. C. Scott-Phillips, M. Tamariz, E. A. Cartmill, & J. R. Hurford (Eds.), *The evolution of language: Proceedings of the 9th International Conference (EVOLANG9)* (p. 418). Singapore: World Scientific.
- Corballis, M. C. (2002a). *From hand to mouth: The origins of language*. Princeton, NJ: Princeton University Press.
- Corballis, M. C. (2002b). Did language evolve from manual gestures? In A. Wray (Ed.), *The transition to language* (pp. 161–179). Oxford: Oxford University Press.
- Corballis, M. C. (2003). From mouth to hand: Gesture, speech, and the evolution of right-handedness. *Behavioral and Brain Sciences*, 26(2), 199–208.
- Corballis, M. C. (2013). Gestural theory of the origins of language. In C. Lefebvre, B. Comrie, & H. Cohen (Eds.), *New perspectives on the origins of language* (pp.171–184). Oxford: Oxford University Press.

- Corballis, M. C. (2014a). The gradual evolution of language. *Humana Mente: Journal of Philosophical Studies*, 27, 39–60.
- Corballis, M. C. (2014b). The word according to Adam. In M. Seyfeddinipur & M. Gullberg (Eds.), *From gesture in conversation to visible action as utterance* (pp. 177–198). Amsterdam: John Benjamins.
- de Condillac, É. B. (1756). *An essay on the origin of human knowledge* (T. Nugent, Trans.). London: J. Nourse. (Original work published 1746)
- de Mandeville, B. (1728). *The fable of the bees. Part II*. London: J. Roberts.
- de Saussure, F. (1959). *Course in general linguistics* (W. Baskin, Trans.). New York, NY: Philosophical Library. (Original work published 1916)
- 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: Norton.
- Fay, N., Arbib, M., & Garrod, S. (2013). How to bootstrap a human communication system. *Cognitive Science*, 37(7), 1356–1367.
- Ferretti, F. (this volume). The narrative origins of language. In N. Gontier & C. Sinha (Eds.), *Handbook of human symbolic evolution*. Oxford: Oxford University Press.
- Ferretti, F., Adornetti, I., Chiera, A., Nicchiarelli, S., Magni, R., Valeri, G., & Marini, A. (2017). Mental Time Travel and language evolution: A narrative account of the origins of human communication. *Language Sciences*, 63, 105-118.
- Fitch, T. (2010). *The evolution of language*. Cambridge: Cambridge University Press.
- Fitch, T. (2017). Empirical approaches to the study of language evolution. *Psychonomic Bulletin & Review*, 24(1), 3–33.

- Galantucci, B. (2017). Experimental semiotics. In *Oxford Research Encyclopedia of Linguistics*. doi: 10.1093/acrefore/9780199384655.013.210
- Gärdenfors, P. (2017). Demonstration and pantomime in the evolution of teaching. *Frontiers in Psychology*, 8, 415. doi:10.3389/fpsyg.2017.00415
- Gärdenfors, P., & Högberg, A. (2017). The archaeology of teaching and the evolution of *Homo docens*. *Current Anthropology*, 58(2), 188–208. doi: 10.1086/691178
- Gärdenfors, P., & Högberg, A. (this volume). In N. Gontier & C. Sinha (Eds.), *Handbook of human symbolic evolution*. Oxford: Oxford University Press.
- Goldin-Meadow, S. (this volume). Gesture is an intrinsic part of modern-day human communication and may always have been so. In N. Gontier & C. Sinha (Eds.), *Handbook of human symbolic evolution*. Oxford: Oxford University Press.
- Goodwin, C. (2017). *Co-operative action (Learning in doing: Social, cognitive and computational perspectives)*. Cambridge: Cambridge University Press.
- Harnad, S. (2000). From sensorimotor praxis and pantomime to symbolic representations. *Proceedings of 3rd International Conference on the Evolution of Language. Paris 3–6 April 2000*, 118–125.
- Hewes, G. W. (1977). A model for language evolution. *Sign Language Studies*, 15, 97–168.
- Hewes, G. W. (1996). A history of the study of language origins and the gestural primacy hypothesis. In A. Lock & C. R. Peters (Eds.), *Handbook of human symbolic evolution* (pp. 263–269). Oxford: Oxford University Press.
- Hockett, C. F. (1960). The origin of speech. *Scientific American*, 203, 88–111.
- Hutto, D. (2008). First communions: Mimetic sharing without theory of mind. In J. Zlatev, T. Racine, C. Sinha, & E. Itkonen (Eds.), *The shared mind: Perspectives on intersubjectivity* (pp. 245–276). Amsterdam: John Benjamins Publishing Company.

- Imai, M., & Kita, S. (2014). The sound symbolism bootstrapping hypothesis for language acquisition and language evolution. *Philosophical Transactions of the Royal Society B*, 369(1651), 20130298.
- Irvine, E. (2016). Method and evidence: Gesture and iconicity in the evolution of language. *Mind & Language*, 31(2), 221–247.
- Itard, J. M. G. (1802). *An historical account of the discovery and education of a savage man: Or, the first developments, physical and moral, of the young savage caught in the woods near Aveyron in the year 1798*. London: Richard Phillips.
- Kendon, A. (2004). *Gesture: Visible action as utterance*. Cambridge: Cambridge University Press.
- Kendon, A. (2008). Signs for language origins? *The Public Journal of Semiotics*, 2, 2–29.
- Kendon, A. (2011). Some modern considerations for thinking about language evolution: A discussion of *The evolution of language* by Tecumseh Fitch. *The Public Journal of Semiotics*, 3(1), 79–108.
- Kendon, A. (2014). The 'poly-modal' nature of utterances and its relevance for inquiring into language origins. In D. Dor, S. Knight, & J. Lewis (Eds.), *The social origins of language* (pp. 67–76). Oxford: Oxford University Press.
- Klima, E., & Bellugi, U. (1979). *The signs of language*. Harvard, MA: Harvard University Press.
- Knowlson, J. R. (1965). The idea of gesture as a universal language in the 17th and 18th centuries. *Journal of the History of Ideas*, 26, 495–508.
- Laromiguière, P. (1826). *Leçons de philosophie sur les principes de l'intelligence, ou sur les causes et sur les origines des idées*. Paris: Brunot-Labbe.
- Lewis, J. (2013). A cross-cultural perspective on the significance of music and dance to culture and society insight from BaYaka pygmies. In M. A. Arbib (Ed.), *Language*,

- music, and the brain: A mysterious relationship* (pp. 45–65). Cambridge, MA: MIT Press.
- Liebal, K., & Oña, L. (2018). Different approaches to meaning in primate gestural and vocal communication. *Frontiers in Psychology, 9*, 478.
<https://doi.org/10.3389/fpsyg.2018.00478>
- MacNeilage, P. F. (2008). *The origin of speech*. Oxford: Oxford University Press.
- McNeill, D. (2012). *How language began: Gesture and speech in human evolution*. Cambridge: Cambridge University Press.
- McNeill, D. (2013). The co-evolution of gesture and speech, and downstream consequences. In C. Müller et al. (Eds.), *Body–language–communication* (Vol. 1) (pp. 480–512). Berlin: Walter de Gruyter.
- Mithen, S. (2005). *The singing Neanderthals: The origins of music, language, mind and body*. London: Weidenfeld and Nicholson.
- Mocerino, R. P. (2016). Gesture, interjection and onomatopoeia in Edward Burnett Tylor’s theory of the origin and development of language. *Theoria et Historia Scientiarum, 13*, 71–84.
- Monboddo (Burnett, J.). (1774). *Of the origin and progress of language*. Edinburgh-London: Balfour, Cadell.
- Motamedi, Y., Schouwstra, M., Culbertson, J., Smith, K., & Kirby, S. (2017). The cultural evolution of complex linguistic constructions in artificial sign languages. In *Proceedings of the 39th Annual Conference of the Cognitive Science Society*. Seattle, WA: Cognitive Science Society.
- Müller, C. (2014). Gestural modes of representation as techniques of depiction. In C. Müller et al. (Eds.), *Body-language-communication: An international handbook on multimodality in human interaction* (pp. 1687–1702). Berlin: De Gruyter Mouton.

- Perlman, M., Dale, R., & Lupyan, G. (2015). Iconicity can ground the creation of vocal symbols. *Royal Society Open Science*, 2(8). doi:10.1098/rsos.150152
- Power, C. (2009). Sexual selection models for the emergence of symbolic communication: Why they should be reversed. In R. Botha & C. Knight (Eds.), *The cradle of language* (pp. 257–280). Oxford: Oxford University Press.
- Rizzolatti, G., & Arbib, M. (1998). Language within our grasp. *Trends in Neurosciences*, 21(5), 188–194.
- Roberts, G., Lewandowski, J., & Galantucci, B. (2015). How communication changes when we cannot mime the world: Experimental evidence for the effect of iconicity on combinatoriality. *Cognition*, 141, 52–66.
- Russon, A. (2018). Pantomime and imitation in great apes: Implications for reconstructing the evolution of language. *Interaction Studies*, 19(1–2), 200–215.
- Ryan, M.-L. (2012). Narration in various media. In P. Hühn, J. C. Meister, J. Pier, & W. Schmid (Eds.), *The living handbook of narratology*. Retrieved from <http://www.lhn.uni-hamburg.de/article/narration-various-media>
- Sandler, W. (2013). Vive la différence: Sign language and spoken language in language evolution. *Language and Cognition*, 5(2-3), 189-203.
- Sandler, W., Meir, I., Padden, C., & Aronoff, M. (2005). The emergence of grammar: Systematic structure in a new language. *Proceedings of the National Academy of Sciences of the United States of America*, 102(7), 2661–2665.
- Schouwstra, M., & de Swart, H. (2014). The semantic origins of word order. *Cognition*, 131, 431–436.
- Scott-Phillips, T. C., & Kirby, S. (2010). Language evolution in the laboratory. *Trends in Cognitive Sciences*, 14(9), 411–417.

- Senghas, R. J., Senghas, A., & Pyers, J. E. (2005). The emergence of Nicaraguan Sign Language: Questions of development, acquisition, and evolution. In S. T. Parker, J. Langer, & C. Milbrath (Eds.), *Biology and knowledge revisited: From neurogenesis to psychogenesis* (pp. 287–306). London: Lawrence Erlbaum Associates.
- Sibierska, M. (2017). Storytelling without telling: The non-linguistic nature of narratives from evolutionary and narratological perspectives. *Language & Communication*, 54, 47–55.
- Sonneson, G. (1997). The ecological foundations of iconicity. In I. Rauch, G. Carr, & F. Gerald (Eds.), *Semiotics Around the World: Synthesis in Diversity* (pp. 739–742). Berlin & New York: Mouton de Gruyter.
- Sterelny, K. (2012a). Language, gesture, skill: The co-evolutionary foundations of language. *Phil. Trans. R. Soc. B*, 367(1599), 2141–2151.
- Sterelny, K. (2012b). *The evolved apprentice. How evolution made human unique*. Cambridge, MA: MIT Press.
- Stokoe, W. C. (1991). Semantic phonology. *Sign Language Studies*, 71, 107–114.
- Stout, D. (2018). Archaeology and the evolutionary neuroscience of language: The technological pedagogy hypothesis. *Interaction Studies*, 19(1–2), 256–271.
- Stout, D., & Chaminade, T. (2012). Stone tools, language and the brain in human evolution. *Phil. Trans. R. Soc. B*, 367(1585), 75–87.
- Sulik, J. (2018). Cognitive mechanisms for inferring the meaning of novel signals during symbolization. *PLoS ONE*, 13(1), e0189540.
<https://doi.org/10.1371/journal.pone.0189540>
- Tallerman, M. (2011). Protolanguage. In K. R. Gibson & M. Tallerman (Eds.), *The Oxford handbook of language evolution* (pp. 479–491). Oxford: Oxford University Press.
- Tanner, J., & Perlman, M. (2017). Moving beyond ‘meaning’: Gorillas combine gestures into sequences for creative display. *Language Communication*, 54, 56–72.

- Tomasello, M. (2008). *Origins of human communication*. Cambridge, MA: MIT Press.
- Tramacere, A., & Moore, R. (2018). Reconsidering the role of manual imitation in language evolution. *Topoi*, 37(2), 319–328.
- Tylor, E. B. (1881). *Anthropology: An introduction to the study of man and civilization*. London: Macmillan & Co.
- Vico, G. (1948). *The new science of Giambattista Vico* (T. G. Bergin & M. H. Fish, Trans.). Ithaca, NY: Cornell University Press. (Original work published 1725)
- Volterra, V., Capirci, O., Rinaldi, P., & Sparaci, L. (2018). From action to spoken and signed language through gesture: Some basic issues for a discussion on the evolution of the human language-ready brain. *Interaction Studies*, 19(1–2), 216–238.
- Waciewicz, S., & Żywicznyński, P. (2015). From the narrow to the broad. Multiple perspectives on language evolution. *Theoria et Historia Scientiarum*, 11, 5–18.
- Waciewicz, S., & Żywicznyński, P. (2018). Language origins: Fitness consequences, platform of trust, cooperation, and turn-taking. *Interaction Studies*, 19(1–2), 167–182.
- Waciewicz, S., Żywicznyński, P., & Orzechowski, S. (2016). Visible movements of the orofacial area. *Gesture*, 15(2), 250–282.
- Zlatev, J. (2008). The co-evolution of intersubjectivity and bodily mimesis. In J. Zlatev, T. Racine, C. Sinha, & E. Itkonen (Eds.), *The shared mind: Perspectives on intersubjectivity* (pp. 215–244). Amsterdam: John Benjamins.
- Zlatev, J. (2014). Bodily mimesis and the transition to speech. In M. Pina & N. Gontier (Eds.), *The evolution of social communication in primates* (pp. 165–178). Lisbon: Springer.
- Zlatev, J., Persson, T., & Gärdenfors, P. (2005). Triadic bodily mimesis is the difference. *Behavioral and Brain Sciences*, 28(05), 720–721.

Zlatev, J., Wacewicz, S., Żywicznyński, P., & van de Weijer, J. (2017). Multimodal-first or pantomime-first? Communicating events through pantomime with and without vocalization. *Interaction Studies*, 18, 465–488.

Zuberbühler, K. (2013). Acquired mirroring and intentional communication in primates. *Language and Cognition*, 5(2–3), 133–143.

Żywicznyński, P. (2018). *Language origins: From mythology to science*. Frankfurt am Main: Peter Lang.

Żywicznyński, P., Gontier, N., & Wacewicz, S. (2017). The evolution of (proto-)language: Focus on mechanisms. *Language Sciences*, 63, 1–11.

Żywicznyński, P., Wacewicz, S., & Sibierska, M. (2018). Defining pantomime for language evolution research. *Topoi*, 37(2), 307–318. doi:10.1007/s11245-016-9425-9