



From Body to Language: Gestural and Pantomimic Scenarios of Language Origin in the Enlightenment

Przemysław Żywiczyński¹ · Sławomir Wacewicz¹

Accepted: 20 December 2021 / Published online: 13 March 2022
© The Author(s) 2022

Abstract

Gestural and pantomimic accounts of language origins propose that language did not develop directly from ape vocalisations, but rather that its emergence was preceded by an intervening stage of bodily-visual communication, during which our ancestors communicated with their hands, arms, and the entire body. Gestural and pantomimic scenarios are again becoming popular in language evolution research, but this line of thought has a long and interesting history that gained special prominence in the Enlightenment, often considered the golden age of glottogony. We highlight several themes key to this line of reflection, such as the idea of innate communication that consists of visual bodily signs and non-linguistic vocalisations; the division of labour between these two semiotic resources (with visual signs used for referential communication and vocal signs for emotional expression); or the developmental trajectory of this bi-modal system of communication determined by its transmission through generations of users and the intervention of social factors. We confront these ideas with problems discussed in current language evolution research, focusing on polysemioticity, sign function, conventionality and modality transition (a postulated transition from gestures to speech in the course of human phylogeny).

Keywords Enlightenment · Pantomime · Gesture · Language origins · Evolution of language

1 Introduction

Gestural and pantomimic accounts of language origin are now center-stage in research on the evolutionary emergence of language. Although this current status derives from the abundance of newly available interdisciplinary evidence, gestural-pantomimic accounts themselves originate from a rich heritage of theoretical arguments, which were based not only on prescientific traditions of philosophical and religious speculation, but also on earlier naturalistic traditions. Of the latter, a particular wealth of ideas linking the origins of language to bodily-visual communication were first expressed in the era of the Enlightenment.

Sections 2–4 are closely based on Chapters 3–5 Przemysław Żywiczyński's book *Language Origins* (2018). The present article, however, addresses a different research problem and remains an independent piece of scholarship.

✉ Przemysław Żywiczyński
przemek@umk.pl

¹ Center for Language Evolution Studies, Nicolaus Copernicus University Toruń, ul. Bojarskiego 1 (Room C3.32), 87-100 Toruń, Poland

The problem of language emergence was one of the fundamental themes in the Enlightenment, to the extent that this era is sometimes regarded as the golden age of the reflection on language origins (Żywiczyński 2018, p. 77). The scale of interest in this problem is demonstrated by the fact that even thinkers whose interests lay far from any concern with language would often articulate their position on the origin of language, as was, for example, the case of the encyclopaedist philosopher Diderot (1713–1784), the political philosopher Voltaire (1694–1778), the economist Adam Smith (1723–1790) or the polymath scientist Mikhail Lomonosov (1711–1765).¹

A defining characteristic of language origins in the Enlightenment was methodological naturalism. In previous epochs, theorising about language origins commonly had religious underpinnings (Żywiczyński 2018). For example, in the Renaissance it was dominated by the Adamic problem, i.e. determining what was the original language of humanity, which was attributed to the biblical Adam (for an extensive discussion of the Adamic problem, see Żywiczyński 2018, pp. 29–54). In opposition to such sentiments, Enlightenment

¹ For a discussion of these thinkers' views on language origins, see Hewes (1975).

thinkers sought for causes of language emergence that are exhausted in nature, without appealing to religious or mythological agencies (Papineau 2017). Naturalistic glottogony so construed, which developed in the seventeenth century and flourished in the eighteenth century, was part of a general intellectual movement in the Enlightenment to eschew supernatural explanations of phenomena in favour of naturalistic ones (Żywiczyński 2018, p. 77).

In this paper, we look at language emergence scenarios tracing the roots of human linguistic capacity back to bodily-visual expression that were articulated during the French Enlightenment. It is important to acknowledge a prominent position of these scenarios in the era (but see Herder 2002 [1772]). As one example, *Les Idéologues*, a highly influential intellectual milieu formed in Paris in the latter part of eighteenth century, set as one of its goals to gather empirical data in support of gestural-pantomimic accounts of language beginnings (Żywiczyński 2018, pp. 118–122). Indeed, towards the end of the Enlightenment this view achieved such popularity that it was often treated not as a hypothesis but a solution to the language-origin problem. In what follows, we extract several core ideas found in the Enlightenment literature on this topic, related to three main themes that are focal in the most recent bodily-visual accounts of language origins:

- *polysemioticity*: combining a variety of vocal and visual means of communicative expression;
- *sign function*: intentional use of a specific form to stand for a specific meaning, and the character of this form-meaning link, which can be either “natural” or conventional and learnt;
- *modality transition problem*: a postulated transition from gestures to speech in the course of human phylogeny.

The term *polysemioticity* (Zlatev 2019) is intended to capture the richness of human communication. Polysemioticity is related to *multimodality*, most importantly the conversational use of different sensory modalities (i.e. sound and vision; cf Levinson 2006); however, describing conversation as *polysemiotic* underscores the fact that it mobilises for communicative purposes a broad variety of distinct semiotic resources, including spoken words, nonlinguistic vocalisations, co-speech gesture, emblems, facial expressions, pantomimic demonstrations and not infrequently other means such as drawing or touching (Zlatev et al. 2020). Contemporary theories of language origins markedly differ from each other on the question of the original division of labour between these resources, in particular with respect to expressing referential-propositional meanings. Although a majority of accounts either implicitly or explicitly assume that vocal signals have always been the dominant substrate of meaning-making, a substantial and growing proportion of language

evolution theories acknowledge an important role of visual-bodily communication.

Among these, multimodal accounts (e.g. Kendon 2004; McNeill 2012) propose that the original system of human communication relied on a close integration of vocalization and gesticulation, today best reflected in the integration of spoken utterances with co-speech gesture. Pantomimic theories (e.g. Zlatev et al. 2020, Żywiczyński et al. 2021) also understand original system to have been multimodal, but with vocalisation largely subserving emotional expression, and referential-propositional meaning expressed mostly by iconic depictions and enactments produced with one’s entire body. Many influential gestural theories, in contrast, focus on the visual signs articulated mostly with the hands and arms (e.g. Calvin and Bickerton 2000; Corballis 2002; Armstrong and Wilcox 2007). With regard to this point, it is interesting to note that Condillac (2001[1746]) and other Enlightenment thinkers (e.g. Mandeville 1729) are commonly referred to as the forefathers of gestural theories of language origins. Contrary to this belief, in the next section we show that a near-consensus among the intellectual figures of the Enlightenment, including Condillac and Mandeville, was much closer to an essentially pantomimic view on language origins. Unlike proponents of modern gestural theories, they understood the original communication system as an integrated system of bodily visual expression and non-linguistic vocalisation, and also stressed that bodily visual expression had comprised communicative body movement of any body part, often involving the whole body.

For many contemporary researchers, the emergence of the sign function (Zlatev et al. 2005) is one of the watersheds in the evolution of language (e.g. Donald 1991). In short, a sign must involve three entities: an *expression* that represents a *referent* for a conscious *interpreter* (Żywiczyński et al. 2021). Communication of non-human animals is dyadic and typically involves a communicator and an addressee, as in the case of grooming.² In contrast, signs are always triadic, in the sense that they involve a communicator, a denoted object, and an addressee. For example, a pointing gesture is a sign where the communicator intends to bring the attention of the addressee to a relevant object, and for the addressee to recognise this, rather than just to look in a given direction (Tomasello 2008). In Sect. 3, we show that the Enlightenment thinkers understood signs of the original communication system as connecting forms of bodily-visual or vocal expression with distinct meanings. We also discuss in more

² Signs are dyadic even if they involve two dyads (1) communicator and addressee, and (2) object and communicator. If only (1) is the case, signals are clearly dyadic, as in mating calls. If both (1) and (2) are the case, we have so-called “functionally referential” signals (Zlatev et al. 2020), such as the well-known vervet monkey alarm calls (Seyfarth and Cheney 1990).

detail Condillac's account of how our ancestor developed sign function.

A considerable, and perhaps the most important, strength of bodily-visual scenarios of language origins is how it addresses *conventionality*. Since human languages critically depend on semiotic conventions (e.g. de Saussure 1960 [1916]), and since these are absent from other animal systems of communication, their origin is a major explanatory target. Here, bodily-visual theories (gestural, multimodal and pantomimic) offer a compelling explanation: the first signs were iconic, and the strong visual resemblance of the bodily movement to the referent (e.g. Donald 1991; Arbib 2012; Gärdenfors 2017) allowed them to be easily understood even in the absence of preestablished semiotic conventions. Both naturalistic studies (e.g. on emerging sign languages, Mineiro et al. 2021) and experiments (e.g. Motamedi et al. 2019) show that such originally iconic systems have a strong tendency to rapidly conventionalise through repeated interaction. That is, iconic signs decrease in the level of the similarity to their referents, and the relation between the form and meanings ceases to be based on resemblance, but instead relies on the common knowledge of the parties of the communicative convention. Although the Enlightenment authors did not directly refer to the problem of iconicity in the context of language origins, they mustered the evidence available to them to argue for the naturalness of bodily visual communication. In Sect. 2, we also refer to Maupertuis's and Rousseau's explanations that bodily visual communication is an expedient means of bootstrapping communication due to its iconic potential, and present a competing view put forward by Condillac that its original function was capturing the interlocutor's attention.

In contrast, the so called "modality transition problem" is widely regarded as a major difficulty of the bodily-visual theories of language origins (Corballis, 2003; Kendon, 2008; MacNeilage, 2008), to some researchers even a disqualifying one (e.g. Burling, 2005). It is indeed a challenge: "If language arose as a (predominantly) gestural/visual system, why would it now have the (predominantly) spoken/vocal form that it does, backed up by the extensive anatomical and neuroanatomical human adaptations to speech production?" (Orzechowski et al. 2016). In Sect. 4 we show, firstly, that the researchers at that time were already well aware of the modality transition problem. An additional point concerns the proposed solutions, which exploited lines of argumentation related to the superficial advantages of vocal communication (such as communicating after dark) that still reverberate in today's discussion but that are considered highly insufficient: modern arguments in this respect can refer to bodies of interdisciplinary data that were inaccessible in the Enlightenment, such as cerebral connections between the motor control systems of manual movements and vocalisations (see Waciewicz et al. 2021 for review).

2 Original Communication: Multimodal, Bodily, Innate

A very popular idea in the Enlightenment was that humans possess an innate system of communication, which hence is natural or universal and whose operation does not require any form of learning. The core element of this system is the use of expressive body movements, as described by Bernard de Mandeville (1670–1733):

When a Man's Knowledge is confin'd within a narrow Compass, and he has nothing to obey, but the simple Dictates of Nature, the Want of Speech is easily supply'd by dumb Signs; and it is more natural to untaught Men to express themselves by Gestures, than by Sounds; but we are all born with a Capacity of making ourselves understood, beyond other Animals, without Speech. (1729, pp. 286–287)

Emotive vocalisation was thought to constitute the other key component of the innate system of communication. In this regard, Mandeville writes:

To express Grief, Joy, Love, Wonder and Fear, there are certain Tokens, that are common to the whole Species. Who doubts that the crying of Children was given them by Nature, to call Assistance and raise Pity, which latter it does so unaccountably beyond any other Sound? (Mandeville 1729, pp. 286–287)

The proposal about the bi-modal nature of the innate system was accepted by many important thinkers of the era. For example, Étienne Bonnot de Condillac (1715–1780) describes it as consisting of "cries of the passions and the different motions of the body" (2001[1746], pp. 114–115) and Jean-Jacques Rousseau (1712–1778), of "gesture and ... inarticulate sounds" (1998[1781], p. 305); similar accounts are found in the works of Pierre Louis Maupertuis (1698–1759; 1965[1756]), César Chesneau Du Marsais (1676–1756; 1792) or Pierre Laromiguière (1756–1837; 1826), one of *Les Idéologues*.

But what were the reasons for positing that the bi-modal, gestural-vocal system is innate to humans? First of all, in the Enlightenment it was a common belief that this is how we communicate when deprived of a shared language and that this is the form of communication used by pre-verbal children (Żywiczyński 2018, pp. 100–108). In this regard, the Enlightenment departed from a long-lived idea of the Adamic language, on which children are thought to have (some) inborn linguistic capacity, often identified with the language spoken by the biblical Adam (see Sect. 1). The combination of secular and empiricist sentiments promoted the conviction that children developed language gradually out of protolinguistic gestures

and non-linguistic vocalisations. However, the proponents of a gestural-vocal protolanguage also appealed to empirical data.

One of these lines of evidence was budding research on ontogenetic development and pedagogy, including signed language pedagogy. For example, Joseph Marie Degérando (1772–1842) in a multi-volume work on semiotics *Signs and the Art of Thinking Considered in Terms of Their Mutual Relations* (1799–1800) used naturalistic data to argue that the child's development of symbolic thinking proceeds from visually transmitted symbols to vocal-linguistic signs. A similar point is made by Louis-François Jauffret (1770–1840) in his programmatic works on the study of ontogeny (for details see Benzaquen 2004). One of the key components of Jauffret's programme was to conduct the forbidden experiment on a large scale (see Maupertuis (1965[1756]) for a similar proposal), and he speculated that gestures and pantomime would be the first signs to emerge. In the work *Surdus loquens* (1692, *The Taking Deaf*), Johan Konrad Amman (1669–1724), the author of the first programmes to teach signed languages, underlined the ease with which new signs are devised by the deaf (Żywiczyński 2018, pp. 86–88). Roch-Ambroise Cucurron Sicard (1742–1822), the director of Paris's National Institute of the Deaf (Institution Nationale des Sourds-Muets), argued that manual signs and vocal signs are acquired largely in the same fashion; however, he agreed with Amman that the acquisition of manual signs requires less effort and less instruction than that of spoken signs (1800; Massieu et al. 1815). In the area of pedagogy for typically developing children, Rousseau's treatise on education, *Emile* (1979[1762]), promoted the idea that an extensive use of bodily-visual communication and emotive vocalisation, such as song, has a beneficial impact on the child's intellectual and social development.

The support for the thesis about the innateness of gestures and emotive vocalisation was also drawn from the study of feral children, i.e. children raised in social isolation.³ In the Enlightenment, feral children and attempts at their rehabilitation began to be treated as sources of information about the problem of language acquisition and, more speculatively, about the origin of language. The best documented feral case of the era was that of Victor of Aveyron (c. 1788–1828), who was found in the woods of southern France in his adolescence. He ended up in the custody of Gaspard Itard (1774–1838), a signed language pedagogue from the National Institute of the Deaf in Paris. Itard developed a programme for rehabilitating his charge, the two main objectives of which were teaching him French and teaching him to recognise human emotions, most importantly

empathy. Itard, in great detail, describes the implementation of this programme in *Historical Account of the Discovery and Education of a Savage Man* (1802[1801]). According to this documentation, Victor made immediate progress in understanding spoken French and developing some forms of civilised behaviour, such as table manners. He also showed both eagerness and skill in communicating with his tutor by means of whole-body pantomimes, manual gestures and non-linguistic vocalisations. This observation led Itard to suggest that language must have begun as a combination of communicative body movements and cries.

Although familiar with the signed language that was being used at the Institute, Itard decided that since Victor was neither mute nor deaf, he should be taught spoken French. Still, the imitative exercises turned out to be almost completely unsuccessful, as Victor was able to clearly enunciate only two French items—*lait* (“milk”) and *Mon Dieu* (“My God”). The results obtained by Itard, and other similar attempts, lent support to the view about the innateness of gestures but also led to the conclusion prefiguring the critical age hypothesis (cf. Lenneberg 1967) that the skills required for the acquisition of spoken language disappear in the course of growing up (esp. Itard 1802, p. 144).

Finally, the proponents of a bodily-visual protolanguage appealed to inter-cultural communication. The data primarily came from the travelogues of European discoverers during the Age of Exploration (15th–seventeenth centuries), who recorded their contacts with indigenous populations (Żywiczyński et al. 2021). One of the most heavily studied sources of this genre was the chronicle *The Principall Navigations, Voiages, Traffiques and Discoveries of the English Nation* (2008 [1598–1600]) compiled by Richard Hakluyt (1553–1616), which contains reports of 15th- and 16th-century travellers and is considered one of the most important texts documenting the Age of Exploration (e.g. Laromiguière 1826). Some of the Enlightenment authors, for example Maupertuis (1965 [1756]), also studied the first ethnographic studies, which began to appear in the eighteenth century.⁴ The conclusion commonly drawn from these sources was that when no shared language is available, people are able to successfully communicate by means of bodily and manual gesture with the support of facial expressions and emotive vocalisations (Żywiczyński et al. 2021). Accordingly, the postulate about the innateness of gestural-vocal expression was often coupled with the proposition that it is a form of communication universal to all cultures. Laromiguière describes this point in the following way:

³ Refer to Newton (2002) and Luchte (2012) for more information on feral children.

⁴ Some of these early studies were conducted under the auspices of the Society of Observers of Man (*Société des observateurs de l'homme*), the world's first scientific organisation that had a distinctly anthropological character, which was founded in 1799 by *Les Idéologues* (Żywiczyński 2018, pp. 120–121).

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.³⁵ (1826, III, p. 113; quoted after Knowlson 1965, p. 507)

Laromiguière concludes with the proposal of constructing a universal language based on this innate capacity to communicate by means of body movements.

The hypothesis about the innateness of a multimodal communication was often accompanied by recapitulationist claims that the phylogenetic emergence of language is similar to the process of the child's communicative development. As already noted, Degérando argued that the child's linguistic development begins with bodily-visual communication, and at this juncture he expressed the view that this communicative system also preceded the appearance of language in our phylogeny (1799–1800). Since children are able to use gestures and emotive vocalisations before they acquire a language in the process of socialisation, such communication, argues Degérando, must also have been used by our pre-linguistic ancestors. The same position is taken by a number of the Enlightenment authors, including Mandeville, Maupertuis, Du Marsais, Laromiguière, or Rousseau. This last author, after submitting that gestures and emotive vocalisations are innate to humans, goes on to say:

In the first times, men, scattered over the face of the earth, had no society other than that of the family, no laws other than those of nature, no language other than that of gesture and some inarticulate sounds. (Rousseau 1998[1781], p. 305)

The recapitulationist sentiments may explain the popularity of thought experiments in the Enlightenment, which speculate about the emergence of language by describing pre-verbal children who become isolated from the rest of humanity and have to invent language from scratch. This process of re-inventing language is treated as a language origin model. Accordingly, the authors of such thought experiments assume that the form of communication innate to humans approximates the original system of communication, out of which language developed in the human phylogeny. Since the dominant view on innate communication was that it made use of gesture and pantomime (together with emotional cries), these semiotic resources were identified as the starting point for language, as is the

case with the two best known pieces of thought experiment literature: Mandeville's *Fable of the Bees* (1729) and Condillac's *Essay on the Origin of Human Knowledge* (1746).

An important problem of the scenarios that identify the bi-modal system of communication as the starting point for language is the division of labour between the two modalities: gesture/pantomime vs. vocalisation. Some authors, e.g. Mandeville or Laromiguière, confine themselves to explaining the role of vocalisation, which—as already noted—had the form of emotive cries and accordingly served the transfer of emotive information, but they do not offer any explanation about the function of gestures in the original communication system. According to Maupertuis, gestures mainly served the transfer of rational contents, i.e. information about objects and relations between them (1965[1756]). He further argues that in modern communication these functions have reversed with (vocal) language responsible for expressing referential-propositional information and gesture emotive information, but he does not elaborate on the causes and manner of this transition. In a somewhat similar manner, Rousseau argues that gestures of the original communication system are superior in communicating about objects and needs related to objects, thanks to their iconic potential—gestures are capable of expressing more meanings than vocalisations and, possibly due to their holistic nature, do so faster. These arguments lead him to the conclusion that gestures are an expedient means of bootstrapping communication without pre-existing communicative conventions:

Although the language of gesture and that of the voice are equally natural, nonetheless the first is easier and depends less on conventions: for more objects strike our eyes than our ears and shapes are more varied than sounds; they are also more expressive and say more in less time. (1998[1781], p. 290)

A different account of the relation between gesture and vocalisation in the original system of communication is given by Condillac. According to the author of *Essay on the Origin of Human Knowledge* (2001[1746]), emotive vocalisations, or “cries of passion”, were the primary means of information transfer. Vocalisations were indexical of an emotional state of the communicator, who used them to recruit the addressee's help in satisfying the needs that had engendered a particular emotion. These cries of passion were accompanied by body movements, which emphasised the emotive message and served to better capture the addressee's attention:

When they [the wild pair] 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 accompa-

nied 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. Moved by this display, the other fixed the eyes on the same object, and feeling his soul suffused with sentiments he was not yet able to account for to himself, he suffered by seeing the other suffer so miserably. From this moment he feels that he is eager to ease the other's pain, and he acts on this impression to the extent that it is within his ability. Thus by instinct alone these people asked for help and gave it. I say "by instinct alone," for reflection could not as yet have any share in it. One of them did not say, "I must bestir myself in that particular way to make the other understand what I need and to induce him to help me"; nor the other, "I see by his motions that he wants to have something and I intend to give it to him." But both acted as a result of the need that was most urgent for them. (Condillac 2001[1746], pp. 114–115)

3 Semiotic Evolution

Condillac's account suggests that vocalisations and body movements accompanying them were *natural indexes* (Peirce 1982 Volume 2, pp. 49–58; Mulder and Herve 1972, pp. 13–18), i.e. their interpretation required only the knowledge of the link between emotions, on the one hand, and specific cries and body movements that these emotions induce, on the other, together with the understanding of the communicative context. Although communicators may have exercised some degree of volitional control over their vocal and bodily behaviours and, hence, use them strategically to recruit the addressee's help, Condillac suggests that sign-formation primarily proceeded from comprehension to production, with the addressee establishing which vocalisation and body movements stand for which needs and ways of satisfying them. Such a pattern of sign-formation seems to follow the model of ontogenetic ritualisation (cf. Abramova 2018),

- in the course of which the producer comes to understand that the addressee takes her vocalisations and body movement to stand for specific needs,
- which leads the producer to use vocalisation and body movements as signs for these needs.

The use of these signs under similar circumstances resulted in stabilising their form and meaning: "The frequent repetition of the same circumstances could not fail, however, to make it habitual for them to connect the cries of the passions and the different motions of the body to the

perceptions which they expressed in a manner so striking to the senses" (Condillac 2001[1746], p. 114).

Condillac's view on sign-formation brings us to an important semiotic problem. The majority of the Enlightenment thinkers automatically assumed that our ancestors had been able to use signs, i.e. connect vocalisations and body movements with meanings (see Sect. 1 on sign function). In this regard, they made a proviso in the reductionist epistemology characteristic of genetic empiricism *a la* Locke and sensationism, which was then a popular position emphasising the role of sensory experience in the constitution of knowledge (Żywicznyński 2018, pp. 107–108). According to Mandeville, Rousseau or Laromiguière, the ability to understand and use symbols does not result from the growth of experience but is the innate endowment of human beings, and was in place at the very beginnings of language, when our ancestors used bodily-vocal communication. Condillac's allegiance to sensationism is stronger: the semiotic ability had to be developed by our ancestors incrementally through sense experience and repeated interactions with each other. He further argues that without the assistance of signs, our ancestors (the wild pair of children) would find it difficult to form stable concepts, even though they could associate sensations with memories:

So long as the children I am speaking of lived apart, the exercise of the operations of their soul was limited to that of perception and consciousness, which do not cease so long as we are awake; to that of attention, which occurred whenever some perceptions affected them in a particular manner; to that of reminiscence, when the circumstances which engaged them stayed before their minds before the connections they had formed were destroyed; and to a very limited exercise of the imagination. The perception of a need, for instance, was connected with the object which had served to relieve it. But having been formed by chance and lacking the steady support of reflection, these connections did not last long. One day the sensation of hunger made these children call to mind a tree loaded with fruit which they had seen the day before. The next day this tree was forgotten, and the same sensation called to mind some other object. Thus the exercise of the imagination was not within their power. It was merely the effect of the circumstances in which they found themselves. (Condillac 2001[1746], p. 114)

Condillac insists that the semiotic development was the prime mover of cognitive evolution, with pantomimic-vocal signs providing stable conceptual units for mental operations, such as memory and imagination. In the course of time, the protolanguage gained displacement (cf. Hockett 1960), and the children were able to communicate not only

about their ongoing experiences but also about what had happened to them in the past:

Their memory began to have some exercise; they gained command of their imagination, and little by little they succeeded in doing by reflection what they had formerly done only by instinct. In the beginning both made it a habit to recognize, by those signs, the sentiments which the other felt at the moment; later they used those signs to communicate the sentiments they had experienced. For example, he who came upon a place where he had become frightened, imitated the cries and motions that were the signs of fear to warn the other not to expose himself to the same danger.

The use of signs gradually extended the exercise of the operations of the soul, and they in turn, as they gained more exercise, improved the signs and made them more familiar. Our experience shows that those two things mutually assist each other. Before the discovery of algebraic signs, the operations of the mind had sufficient exercise to lead to their invention; but it is only after the coming into use of these signs that the operations have had the requisite exercise to carry mathematics to the point of perfection at which we find it today. (Condillac 2001[1746], pp. 114–115)

4 The “Modality Transition Problem” and the Transition to Speech

The budding research conducted by Amman and at the National Institute of the Deaf in Paris contributed to the appreciation of the linguistic potential of signed languages in the eighteenth century. However, in the Enlightenment the spoken modality was considered a defining characteristic of language; hence, for example, the attempts to rehabilitate feral children (Itard 1802) or to linguistically educate non-human primates focused on speech training (La Mettrie 1996[1748]). Accordingly, the proponents of the original bimodal system had to account for the predominantly spoken character of language.

Mandeville, Condillac and Rousseau argue that speech emerged from the emotive vocalisation of the original communication system, as Condillac puts it: “... when they had acquired the habit of connecting some ideas to arbitrary signs, the natural cries served as a model for them to make a new language” (Condillac 2001[1746], pp. 115–116). This process required the acquisition of the volitional control over vocalisation (see above/below). To account for this change, Mandeville and Condillac underline the role of children and transgenerational transmission in developing speech, and in doing so, they expressed the view already established in the Enlightenment that the

vocal learning abilities of children are superior to these of adults. Mandeville addresses this problem in the following way:

They would find that the Volubility of Tongue, and Flexibility of Voice, were much greater in their young ones than they could remember it ever to have been in themselves ... Some of these young ones would, either by Accident or Design, make use of this superior Aptitude of the Organs at one time or other; which every Generation would still improve upon; and this must have been the Origin of all Languages, and Speech itself, that were not taught by Inspiration. (Mandeville 1729, pp. 287–288)

Condillac identifies as the principal factor constraining the development of speech adults’ inability to make new sounds, due to the inflexibility of their speech organs:

They articulated new sounds, and by repeating them many times to the accompaniment of some gesture that indicated the objects to which they wished to draw attention, they became accustomed to giving names to things. Still, the first progress of this language was very slow. The organ of speech was so inflexible that it could articulate only very simple sounds with any ease. The obstacles to the pronunciation of other sounds even prevented them from suspecting that the voice could vary beyond the small number of words already imagined. (Condillac 2001[1746], pp. 115–116)

When explaining how new sounds entered the protolanguage, Condillac seems to appeal to a version of the orofacial hypothesis (cf. Waciewicz et al. 2021), whereby the wild pair’s offspring, whose vocal capacities were superior to these of the parents, were able to make their articulators assume new positions that reflected their body movements during the production of pantomimes:

This couple had a child who, when pressed by the needs he could make known only with difficulty, agitated all parts of the body. His very flexible tongue bent itself in some extraordinary manner and pronounced an entirely new word. The need still persisting again caused the same effects; the child moved the tongue as before and once more articulated the same sound. Full of surprise and having at last figured out what the child wanted, the parents gave it to him while at the same time trying to repeat the same word. The trouble they had pronouncing it showed that they would not by themselves have been able to invent it. (Condillac 2001[1746], p. 116)

The transition into vocal language was possible when a repertoire of articulate sounds was large enough to keep the vocal organs of new generations of children busy to such

a degree that the loss of initial articulatory flexibility was prevented:

As the language of articulated sounds became richer, it was better suited to exercise the vocal organ at an early stage and to preserve its initial flexibility. It then became as convenient as the language of action; either one was used with equal ease until the use of articulated sounds became so easy that they prevailed. (Condillac 2001[1746], p. 116)

On Mandeville's and Condillac's view, speech could have grown just in effect of the transmission of the original system of communication through successive generations of communicators, thanks to the fact that the superior articulatory capabilities of children had to be put to practising a growing repertoire of vocal signs. Mandeville believed the transition to vocal language occurred in the context where interlocutors did not see each other. According to Condillac, the development of communicative system proceeded hand in hand with that of the conceptual one (see above). His followers, Du Marsais and Maupertuis, argued that speech was better at expressing a wider array of concepts. Du Marsais contended that at the beginning of language, gesture was more adapt at expressing what he calls "rational content" (roughly equivalent to referential-propositional meaning); however, with the growth of concepts the vocal modality became the preferred means of communication (1792). Maupertuis concentrated on the combinatorial potential of vocal signs, which—as he argued—makes it easier to combine vocal signals to create new meanings, and used this argument to explain the shift of the original communication system to the vocal-auditory modality (1965[1756], pp. 437–438, cf. Hewes 1976, p. 484).

In contrast to these authors, Rousseau stresses that the development of language away from the original pantomimic-vocal mode required a push from outside the domain of communication, and this push was provided by the lifestyle change, whereby man became a social animal. Thinking about the beginnings of this process, Rousseau comes to the conclusion that vocalisation much better served the new demands, as it was more effective than gesture in bringing people together and coordinating their activities; hence it became the dominant modality at this stage of language emergence:

The natural effect of the first needs was to separate men and not to bring them together. This had to have been so for the species to spread and the earth to be populated promptly, otherwise mankind would have been crammed into one corner of the world while the rest of it remained deserted. ... The passions all bring men together, but the necessity of seeking their livelihood makes them flee one another. Neither hunger

nor thirst, but love, hatred, pity, anger wrested the first voices from them. Fruit does not elude our grasp, one can feed on it without speaking, one stalks in silence the prey one wishes to devour; but in order to move a young heart, to repulse an unjust aggressor, nature dictates accents, cries, complaints. The most ancient words are invented in this way, and this is why the first languages were tuneful and passionate before being simple and methodical. (1998[1781], pp. 293–294)

First languages made use of tones, which gave them the "sonorous and harmonious" quality. This quality served to incite appropriate passions and in this way keep people together. They lived in small family groups, *sociétés naissantes*, in the south (i.e. the south of Europe), where lush vegetation generously supplied with all their vial needs, and hence, they could do away with private property (Lovejoy 1923, p. 182). When people migrated north, the new inhospitable lands required that they should form bigger groups and cooperate so as to be able to satisfy vital needs. In these lands, linguistic communication become oriented towards transferring more and more abstract ideas about needs and ways of satisfying them. For Rousseau, this increased precision depended on increased conventionalisation and the irrevocable loss of the original musical quality of language.

5 Conclusion

Many important assumptions of contemporary bodily-visual accounts of language origins (e.g. Donald 1991; McNeill 2012; Tomasello 2008; Zlatev et al. 2020) can be recognised in the views elaborated in the Enlightenment. Such is the case in particular with the foundational claim about the naturalness of gesture and pantomime as a human-specific form of communication, which we discussed in Sect. 2. Contemporary research may have blunted the strong thesis about the innateness of bodily-visual signals, but lines of evidence coming from the study of language impairments (e.g. Fex et al. 1998, Klippi 1996), home-signing children (e.g. Goldin-Meadow 1998) or silent gesture (e.g. Schouwstra et al. 2019; Ortega and Özyürek 2020) corroborate the view that gesture and pantomime are an essential part of human natural communicative inventory.

Contemporary research has also confirmed that these semiotic resources are an expedient means of bootstrapping semiotic conventions and compositionality (Motamedi et al. 2019). As noted in Sect. 1, research into newly emerging signed languages demonstrates that improvised bodily-visual signs constitute the starting point of what later can become a fully-fledged language (Nicaraguan Sign Language, e.g. Kegl 1999; Al-Sayyid Bedouin Sign Language, Sandler 2012; Sao Tome and Principe Sign Language, Mineiro

et al. 2021). Similar conclusions are drawn from experimental studies that investigate how people improvise communication (Fay et al. 2014; Zlatev et al. 2017; Motamedi et al. 2019). As we saw in Sect. 2, Mandeville and Condillac claimed that the original system of communication becomes gradually elaborated through a process of transgenerational transmission with a special role played by children. Similar intuitions now echo in modern views on the dynamics of language evolution, with empirical support coming, for example, from signed language studies (Senghas et al. 2004; Sandler 2012). There is also a growing emphasis in modern research on the role that social factors played in language evolution (see esp. Dor et al. [eds.] 2014)—a point underlined in Rousseau’s account of language emergence.

As we observed in Sect. 1, important trends in modern language evolution research, mainly affiliated with gesture studies (Kendon 2004; McNeill 2012) and mimesis theory (Donald 1991; Zlatev et al. 2020), assume that protolanguage was polysemiotic and hence its evolution must be viewed as an interplay of continuities between the semiotic systems that originally formed it. Proponents of these scenarios often highlight the polysemiotic nature of the endpoint of the evolutionary continuum: modern linguistic communication, particularly as used in its core ecological niche—face-to-face interaction (cf. Kendon’s notion of *languageing* (2004) or Levinson’s Interaction Engine (Levinson 2006)). Although derived from different theoretical motivations, Mandeville and Condillac’s idea of bi-modal (pantomimic-vocal) protolanguage, which was accepted by many Enlightenment authors, strongly resembles modern polysemiotic accounts of language emergence. It is interesting to observe that unlike the proponents of modern multimodal hypotheses (Kendon 2004; McNeill 2012; see Sect. 1), the Enlightenment authors posit that bodily visual communication was the primary means of transmitting referential-propositional information, while the role of vocalisation was largely limited to emotive expression (cf. Maupertuis’s idea of “rational content”, but also Condillac’s dissenting view on the division of labour between bodily visual communication and vocalisation). Such proposals are instead closer to modern pantomimic accounts of language evolution (Donald 1991; Gärdenfors 2017; Zlatev et al. 2020; see Sect. 1).

We submit that the continuing popularity of the view that language began with bodily-visual communication and the consistency with which it is discussed in the Enlightenment and modern times, derives from its naturalistic underpinnings. An obvious course of reflection to take when faced with the question about language origins is to consider how modern humans bootstrap communication in the absence of a shared spoken language. An appealing answer, informed by our everyday experience, is that in such circumstances we make use of the other major semiotic resources in the arsenal of our communicative behaviours, i.e.—gesture and

pantomime. Once such a view is accepted, it opens a set of ancillary questions:

1. what exactly makes bodily-visual communication an effective means of bootstrapping communication, which leads to the problem of sign function and iconicity;
2. was bodily-visual communication aided at the bootstrapping stage by other semiotic resources, which leads to the consideration of polysemioticity;
3. what was the manner of transition of this original system towards more language-like forms of communication, which leads the problem of conventionalisation; and
4. how did bodily-visual communication change into speech, which leads to the modality transition problem.

Appealing to knowledge available to them, scholars of different epochs have to grapple with these problems, whether they decide to build on the efforts of their predecessors or not.

Acknowledgements We are grateful to anonymous reviewers for insightful comments to an earlier version of this paper.

Funding PŽ was supported by the Polish National Science Centre (NCN) under grant agreement UMO-2017/27/B/HS2/00642. SW was supported by the Polish National Science Centre (NCN) under grant agreement UMO-2019/34/E/HS2/00248.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

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
- Amman JK (2009) *The Talking Deaf*. Dodo Press, Gloucester
- Arbib MA (2012) *How the brain got language*. Oxford University Press, Oxford
- Armstrong DF, Wilcox S (2007) *The Gestural Origin of Language*. Oxford University Press, New York
- Benzaquen AS (2004) Childhood, identity and human science in the enlightenment. *Hist Work J* 57(1):34–57
- Burling R (2005) *The talking ape: How language evolved*. Oxford University Press, Oxford
- Calvin WH, Bickerton D (2000) *Lingua ex machine: reconciling Darwin and Chomsky with the human brain*. MIT Press, Cambridge, MA

- Condillac ÉB de (2001[1756]) *An Essay on the Origin of Human Knowledge (Essai sur l'origine des connaissances humaines)*. Trans. T Nugent. Nourse, London
- Corballis MC (2002) *From Hand to Mouth: The Origins of Language*. Princeton University Press, Princeton
- Corballis MC (2003) From mouth to hand: gesture, speech, and the evolution of right-handedness. *Behav Brain Sci* 26:199–208
- Cucurron Sicard R-A (1800) *Cours d'instruction d'un sourd-muet de naissance, pour servir à l'éducation des sourds-muets*. Le Clere, Paris
- de Saussure F (1960) [1916] *Cours de Linguistique Générale/ Course in General Linguistics*. Payot/Duckworth, Paris/London
- Dégérando JM (1799–1800) *Des Signes et de l'Art de Penser Consideres dans leurs Rapports Mutuels*, Vols. I–IV. Goujon, Paris
- Donald M (1991) *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition*. Harvard University Press, Cambridge, MA
- Dor DE, Knight CE, Lewis JE (eds) (2014) *The social origins of language*. Oxford University Press, Oxford
- Du Marsais CC (1792) *Logique et principes de grammaire*. Braisson, Breton, Herissant Fils, Paris
- Fay N, Lister CJ, Ellison TM, Goldin-Meadow S (2014) Creating a communication system from scratch: gesture beats vocalization hands down. *Front Psychol* 5:354
- Fex B, Månsson AC (1998) The use of gestures as a compensatory strategy in adults with acquired aphasia compared to children with specific language impairment (SLI). *J Neurolinguistics* 11:191–206. [https://doi.org/10.1016/s0911-6044\(98\)00013-x](https://doi.org/10.1016/s0911-6044(98)00013-x)
- Gärdenfors P (2017) Demonstration and pantomime in the evolution of teaching. *Front Psychol* 8:415
- Hakluyt R (2008 [1598–1600]) *The Principall Navigations, Voiages, Traffiques and Discoveries of the English Nation*. Project Gutenberg, Salt Lake City, UT. <https://www.gutenberg.org/files/25645/25645-pdf.pdf>
- Herder JG (2002) *Treatise on the Origin of Language*. In: Forster MN (ed) *Philosophical Writings*. Cambridge University Press, Cambridge, pp 65–164
- Hewes GW (1975) *Language Origins: A Bibliography*. Mouton, The Hague
- Hewes GW (1976) The current status of the gestural theory of language origin. *Ann N Y Acad Sci* 280(1):482–504
- Hockett CF (1960) The origin of speech. *Sci Am* 203:88–111
- Itard JMG (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*. Richard Phillips, London
- Kegl J, Senghas A, Coppola M (1999) Creation through contact: sign language emergence and sign language change in Nicaragua. In: DeGraff M (ed) *Language creation and language change: creolization, diachrony, and development*. MIT Press, Cambridge, MA, pp 179–237
- Kendon A (2004) *Gesture: Visible Action as Utterance*. Cambridge University Press, Cambridge
- Kendon A (2008) Signs for language origins? *Public J Semiot* 2:2–29
- Klippel A (1996) Conversation as an achievement in aphasics. *Studia fennica linguistica* 6. The Finnish Literature Society, Helsinki
- Knowlson JR (1965) The idea of gesture as a universal language in the 17th and 18th centuries. *J Hist Ideas* 26:495–508
- La Mettrie JO (1996) *Machine Man and Other Writings*. Cambridge University Press, Cambridge
- Laromiguière P (1826) *Leçons de philosophie sur les principes de l'intelligence, ou sur les causes et sur les origines des idées*. Brunot-Labbe, Paris
- Lenneberg EH (1967) *Biological Foundations of Language*. Wiley, New York
- Levinson S (2006) On the human interaction engine. In: Enfield N, Levinson S (eds) *Roots of human sociality: culture, cognition and interaction*. Berg, Oxford, pp 36–69
- Lovejoy AO (1923) The supposed primitivism of Rousseau's 'discourse on inequality.' *Mod Philol* 21(2):165–186
- Luchte J (2012) *Of the Feral Children*. Createspace, London
- MacNeilage PF (2008) *The Origin of Speech*. Oxford University Press, Oxford
- Mandeville B de (1729) *The Fable of the Bees*. Part II. J. Roberts, London, digitised version available at: <https://archive.org/details/fableofbeespart00mand> (DOA: 7 Feb 2021)
- Massieu J, Clerc L, Cucurron Sicard R-A (1815) A collection of the most remarkable definitions and answers of Massieu and Clerc, deaf and dumb, to the various questions put to them, at the public lectures of the Abbé Sicard, in London. Cox and Baylis, London
- Maupertuis PLM de (1965 [1756]) *Dissertation sur les Differentes Moyens dont les Hommes se sont servis pour Exprimer leurs Idees*. Oeuvres. Vol. II. Georg Olms Verlagsbuchhandlung, Hildesheim
- McNeill D (2012) *How language began: Gesture and speech in human evolution*. Cambridge University Press, Cambridge
- Mineiro A, Báez-Montero IC, Moita M, Galhano-Rodrigues I, Castro-Caldas A (2021) Disentangling pantomime from early sign in a new sign language: window into language evolution research. *Front Psychol* 12:1130
- Motamedi Y, Schouwstra M, Smith K, Culbertson J, Kirby S (2019) Evolving artificial sign languages in the lab: From improvised gesture to systematic sign. *Cognition* 192:103964
- Mulder JWF, Hervey SGJ (1972) *Theory of Linguistic Sign*. Mouton, The Hague
- Newton M (2002) *Savage Boys and Wild Girls: A History of Feral Children*. Faber and Faber, London
- Ortega G, Özyürek A (2020) Systematic mappings between semantic categories and types of iconic representations in the manual modality: a normed database of silent gesture. *Behav Res Methods* 52(1):51–67
- Orzechowski S, Wacewicz S, Żywicznyński P (2016) Vocal-auditory feedback and the modality transition problem in language evolution. *Reti Saperi Linguaggi* 3(1):157–178
- Papineau D (2017) Naturalism. *The Stanford Encyclopedia of Philosophy*. <https://plato.stanford.edu/archives/win2016/entries/naturalism> (DOA: 12 Dec 2020)
- Peirce CS (1982) *The Writings of Charles S. Peirce: A Chronological Edition, Volumes 1–6*, Peirce. Indiana University Press, Bloomington IN, Project
- Rousseau J-J (1979) *Emile, or On Education*. Basic Books, New York
- Rousseau J-J (1998) *Essay on the Origin of Languages and Writings Related to Music*. University Press of New England, Hanover
- Sandler W (2012) Dedicated gestures and the emergence of sign language. *Gesture* 12:265–307. <https://doi.org/10.1075/gest.12.3.01san>
- Schouwstra M, de Swart H, Thompson B (2019) Interpreting Silent Gesture: cognitive biases and rational inference in emerging language systems. *Cognitive Science* 43(7):e12732
- Senghas A, Kita S, Özyürek A (2004) Children creating core properties of language: evidence from an emerging sign language in Nicaragua. *Science* 305:1779–1782. <https://doi.org/10.1126/science.1100199>
- Seyfarth R, Cheney D (1990) The assessment by vervet monkeys of their own and another species' alarm calls. *Anim Behav* 40(4):754–764
- Tomasello M (2008) *Origins of Human Communication*. MIT Press, Cambridge, MA

- Waciewicz S, Orzechowski S, Żywiczyński P (2021) Visible movements of the orofacial area: evidence for gestural or multimodal theories of language evolution? *Gesture* 15(2):250–282
- Zlatev J (2019) Mimesis theory, learning and polysemiotic communication. In: Peters M (ed) *Encyclopedia of Educational Philosophy and Theory*. Springer, Singapore
- Zlatev J, Waciewicz S, Żywiczyński P, van de Weijer J (2017) Multimodal-first or pantomime-first? Communicating events through pantomime with and without vocalization. *Interact Stud* 18:465–488. <https://doi.org/10.1075/is.18.3.08zla>
- Zlatev J, Żywiczyński P, Waciewicz S (2020) Pantomime as the original human-specific communicative system. *J Lang Evol* 5(2):156–174
- Zlatev J, Persson T, Gärdenfors P (2005) Bodily mimesis “the missing link” in human cognitive evolution. LUCS: Lund University Cognitive Studies 121 <http://www.lucs.lu.se/LUCS/121/LUCS.121.pdf>
- Żywiczyński P (2018) *Language origins. From mythology to science*. Peter Lang, Berlin
- Żywiczyński P, Waciewicz S, Lister C (2021) Pantomimic fossils in modern human communication. *Philos Trans R Soc B* 376:20200204. <https://doi.org/10.1098/rstb.2020.0204>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.