How Pantomime Works: Implications for Theories of Language Origin

Steven Brown[1]*, Emma Mittermaier[1], Tanishka Kher[1] and Paul Arnold[2]

[1] Department of Psychology, Neuroscience & Behaviour, McMaster University, Hamilton, ON, Canada, [2] Department of Religious Studies, McMaster University, Hamilton, ON, Canada

Pantomime refers to iconic gesturing that is done for communicative purposes in the absence of speech. Gestural theories of the origins of language claim that a stage of pantomime preceded speech as an initial form of referential communication. However, gestural theories conceive of pantomime as a unitary process, and do not distinguish among the various means by which it can be produced. We attempt here to develop a scheme for classifying pantomime based on a proposal of two new sub-categories of pantomime, resulting in a final scheme comprised of five categories of iconic gesturing. We employ the scheme to establish associations between the category of pantomime used and the type of action and/or object being depicted. Based on these associations, we argue that there are two basic modes of pantomiming and that these apply to distinct semantic categories of referents. These modes of pantomiming lead to two alternative models for a gestural origin of language, one based on people and one based on the environment.

Keywords: pantomime, gesture, language, classification, iconic, narrative, evolution, mimicry

INTRODUCTION

Pantomime features prominently in most, if not all, gestural theories of the origins of language. However, none of the mainstream evolutionary models has provided a conception of pantomime that conveys the semantic and praxic diversity of pantomime's various modes of depiction. In this article, we attempt to show the importance of incorporating pantomimic diversity into gestural models of language origin. After providing a basic background regarding what pantomime is and is not, we present a new classification scheme for pantomime comprised of five categories. Based on this scheme, we attempt to establish associations between the category of pantomime used by the mimer and the type of action and/or object being depicted. In the final section, we use these associations to provide new insights into gestural models of language origin by arguing that the two major manners of pantomiming lead naturally to a contrast between two types of gestural models, what we will refer to as a People First model and an Environment First model.
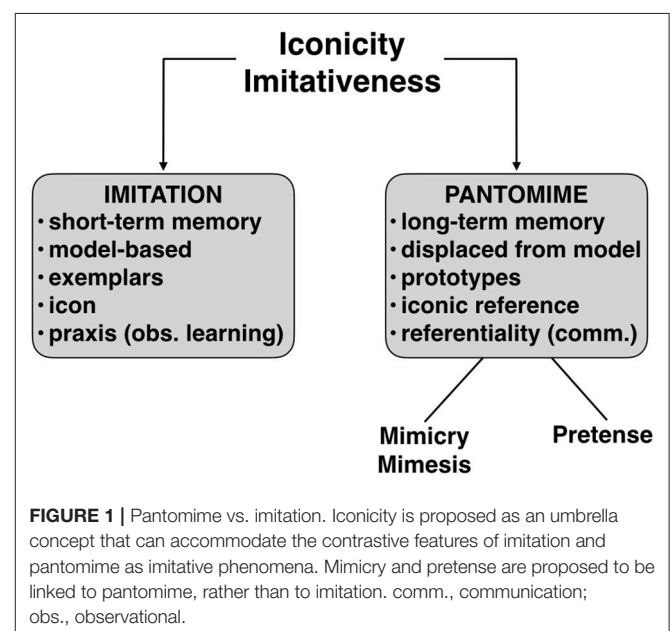
WHAT PANTOMIME IS AND IS NOT

Pantomime refers to iconic gesturing that is done for communicative purposes in the absence of speech. By "iconic" we mean a type of representational gesturing that shows a strong spatial resemblance to its referent(s) (Perniss and Vigliocco, 2014). According to Zywiczynski et al. (2018), other salient features of pantomime beyond its iconicity (Arbib, 2012) include that it is improvised, non-conventionalized, holistic, and open-ended, thus having a broad semantic potential. It is also referential, or triadic (Arbib, 2012; Zlatev, 2014). While Zywiczynski et al. (2018) argue that pantomime is a whole-body process (see also Zlatev, 2014), it is quite easy to think of

counter-examples to this, such as when a person uses their index and middle fingers to represent somebody walking. Hence, while pantomime can indeed engage the full body, it can also employ body parts alone. Pantomimes can depict both objects and actions, and the actions can be either transitive or intransitive. A defining feature of the miming of transitive actions is that the gestures are empty-handed, involving "imaginary objects" (Boyatzis and Watson, 1993; O'Reilly, 1995). The agents and recipients of mimed actions can be the self, but they can also be other people or non-human animals. The two major functional roles attributed to pantomime during everyday social interactions are gestural depiction during interpersonal communication (Clark, 1996) and demonstration during supervised learning (Gärdenfors, 2017). However, pantomime has a diverse profile of social uses (Zywiczynski et al., 2018) that spans from the performing art of mime theater (Hall, 2009) through to the neuropsychological testing that is done to diagnose syndromes, such as apraxia (Heilman et al., 1982; Hoeren et al., 2014), aphasia (Rose et al., 2017; van Nispen et al., 2018), and autism (Rogers et al., 1996; Smith and Bryson, 2007; Gizzonio et al., 2015).

How pantomime fits into the overall scheme of gesturing is still under dispute. The standard cognitive classification of gesturing is a tetrachotomy proposed by McNeill (1992, 2005)—as based on the work of Kendon (1988)—comprised of co-speech gesturing (gesticulation), emblems, pantomime, and sign language. This is presented as a continuum of speech involvement, such that gesticulation obligatorily requires the presence of speech, while pantomime and signing obligatorily require the absence of speech. McNeill's conception of pantomime is quite restricted, as he sees it mainly as a form of theatrical performance, or what he calls "dumb show" (McNeill, 2005). In this way, he separates the iconicity of pantomime from that which occurs in gesticulation and sign language. Gestural theories of language origin do not adopt this restricted view of pantomime, which they see as iconic gesturing in the general sense (Armstrong and Wilcox, 2007; Tomasello, 2008; Arbib, 2012). We too adopt this broad view of pantomime that the capacity for iconic gesturing is shared between pantomime and the iconic components of gesticulation and sign language.

Pantomime is often confounded with the concepts of imitation and mimicry, and so we would like to examine what pantomime shares and does not share with these concepts, as summarized in **Figure 1**. Pantomimic gesturing is iconic in that it visually resembles the object or action being depicted by the mimer. In this regard, pantomime is said to be "imitative" and "mimetic." Imitation, like pantomime, is iconic. However, there are significant cognitive and behavioral differences between pantomime and imitation that distinguish them as imitative phenomena. Imitation implies the online reproduction of some currently-observed sensory stimulus, be it visual or auditory, and be it dynamic or static. Imitation is generally an immediate, bottom-up process that works with *short*-term representations of *occurrent* stimuli, as is discussed extensively in the literature on vocal imitation in animals and humans (Petkov and Jarvis, 2012). (An important corollary to the immediacy of imitation for observational learning is the fact that imitated movements can be recalled long after the stimulus is gone as part of the

process of rehearsal, hence utilizing long-term representations, similar to pantomime). Imitation, almost by definition, implies that the stimulus is an exemplar, such as a specific dance step, as performed by someone who models the motor pattern for the imitator. Pantomimes, by contrast, generally operate using *long*-term representations of stimuli stored as object and action schemas in semantic memory. For this reason, they most often (but not always) deal with prototypes, compared to the exemplars that are used in imitation. These are not simply icons, but are categorical representations of actions or objects (Harnad, 1990). They are thus much more top-down and schematic in organization. A pantomime is generally an abstraction of multiple instances of some action or object. When pantomiming a tennis serve, a person does not typically have one particular person's serve in mind, but produces a depiction of a reasonable-looking serve. Pantomime is thus far more symbolic than imitation. For example, pantomime of a tennis serve involves the use of an imaginary racquet and ball, while the "call me" pantomime, in which a person uses a hand to symbolically represent a telephone receiver, involves body-part substitution (Suddendorf et al., 1999; Dick et al., 2005). In both cases, the pantomime is a stand-in for the real object and/or action, and is thus a referential gesture, something that will be called an iconic reference in this article. The term reference signifies a class or category of object or action, as in the case of functionally referential alarm calls in vervet monkeys that reference particular categories of predators (Seyfarth et al., 2005). Arbib (2012) has argued that imitation provided an evolutionary scaffold for pantomime, such that there was a transition in iconic motor behaviors from a stage of pure icon (imitation) to one of iconic reference (pantomime). We can imagine the related emergence of "displacement" in human communication (Hockett, 1960; Perniss and Vigliocco, 2014), whereby there was a transition from a stage of imitation dependent on the presence of models

Iconicity
Imitativeness

IMITATION	**PANTOMIME**
• short-term memory	• long-term memory
• model-based	• displaced from model
• exemplars	• prototypes
• icon	• iconic reference
• praxis (obs. learning)	• referentiality (comm.)

Mimicry
Mimesis

Pretense

FIGURE 1 | Pantomime vs. imitation. Iconicity is proposed as an umbrella concept that can accommodate the contrastive features of imitation and pantomime as imitative phenomena. Mimicry and pretense are proposed to be linked to pantomime, rather than to imitation. comm., communication; obs., observational.

to a stage of pantomime that was displaced from models. An intermediate evolutionary stage might have been the emergence of rehearsal for praxis (Donald, 1991; Gärdenfors, 2017), where imitative behaviors are displaced from their models, but where they do not yet function for communication or symbolization.

Beyond sensorimotor differences alone, it is important to consider the contrastive social functions of pantomime and imitation in human life. Pantomime is almost always linked to narrative communication, whereas imitation is most strongly connected with praxis and the learning of motor skills. A majority of motor learning in human life occurs via observational learning (Wulf and Mornell, 2008), where people imitate the movements or vocal patterns of role models, both in the moment and through later rehearsal. Imitative learning provides one of the foundations for the cultural evolution of social information through behavioral modeling (Rendell et al., 2010; Legare and Nielsen, 2015). In addition, the cognitive tendency to imitate the behaviors of others—so-called conformity bias—serves to increase within-group behavioral homogeneity and sharpen between-group cultural differences (Richerson et al., 2016).

A useful way to conceptualize the relationship between pantomime and imitation is to think of iconicity as the umbrella concept, and to see pantomime and imitation as two means of producing iconic gestures (**Figure 1**). Imitation works with exemplars in the moment, while pantomime works with prototypes that are stored in long-term memory. Imitation supports praxis through observational learning, while pantomime supports depiction during narrative communication. The concept of mimicry—and the related concept of mimesis— maps best onto the pantomime concept. To mimic someone (or something) is more to pantomime them than to imitate them in the moment, although some connotations of the term include this meaning as well. This will be discussed in more detail below in the subsection about mimicry. Another concept that fits into the pantomime category is pretense, as in the case of pretend play. Pretend play, like pantomime, is defined as "non-literal action" (Leslie, 1987; Carlson and Taylor, 2005; Lillard et al., 2014; Weisberg, 2015). It is a *representation* of an action, rather than a true (instrumental) action. Perhaps the biggest difference between pretend play and pantomime is that pretend play generally involves the use of props, while pantomime does not; if anything, pantomime is done in order to embody "imaginary objects," which become the props of pantomime. For example, while someone engaged in pretend play could use a banana as a prop to represent a telephone receiver, a mimer would have to use his or her hand to represent the same object. Another salient difference is that pretend play generally depicts events occurring in a fictional storyworld (Walton, 1990), while pantomime generally depicts actions and objects occurring in the real world of the mimer, since it functions in narrative communication.

TWO MODES OF PANTOMIMING

The psychological literature on gesture production has suggested that there are two major modes of pantomiming. More specifically, pantomimic gestures can be produced from either a

character viewpoint (CV) or an observer viewpoint (OV) (Cassell and McNeill, 1990; McNeill, 2005; Parrill, 2010; Cartmill et al., 2012; Beattie, 2016). When a gesture is done from the character viewpoint, it is carried out in a first-person manner such that there is a correspondence between the mimer's body parts and those mediating the actions being mimed. We will call this a situation of body-part use. For example, when miming a tennis serve, the mimer's arm is used to depict an arm movement. By contrast, when a gesture is done from the observer viewpoint, the mimer's body parts undergo a process of substitution to become *some other object*. We will call this a situation of body-part replacement. For example, when a person does the "call me" pantomime by placing her hand next to her ear, she uses her hand to represent a telephone receiver, not her hand. As a result of this body substitution, her hand becomes a symbol for the telephone receiver. However, this is an iconic symbol due to its conventionalization (Deacon, 1997; Taub, 2001; Arbib, 2012), providing evidence against the hard distinction between icon and symbol found in classic semiotic accounts of communication (Peirce, 1903/1987).

One of the major proposals of this article is that the CV/OV distinction should be replaced by terminology from the study of spatial cognition and navigation. In particular, we propose replacing CV with the term "egocentric" and replacing OV with the term "allocentric." Egocentric indicates a conception of space with reference to the position and orientation of a person's body, whereas allocentric indicates a map-like conception of space with reference to objects and their surrounding environment (Klatzky, 1998; Mellinger and Vosgerau, 2010; Gramann, 2013). Egocentric coordinate systems are defined with respect to a viewer, whereas allocentric coordinate systems are not, since they are purely object-related (i.e., there is no "ego"). In the classification of pantomime that we develop here, we will make a binary distinction between egocentric and allocentric forms of iconic gesturing, mirroring that between CV and OV, respectively; it also has similarities to the distinction between figure and ground in cognitive linguistics (Langacker, 1986). The egocentric type is exemplified by the "imaginary object" (IO) pantomime, in which a person performs an empty-handed transitive action in the first person with an imaginary object (Suddendorf et al., 1999), such as when someone mimes a tennis serve with an imaginary racquet in one hand and an imaginary ball in the other. The allocentric type is exemplified by the "body-part-as-object" (BPO) pantomime, in which a person uses a body part (typically the hand or its parts) to represent some external object, such as when a person does the "call me" pantomime by using their hand to represent a telephone receiver, or when a person mimes someone walking by using their index and middle fingers to represent their two legs. Egocentric and allocentric pantomimes require very different conceptions of how the body performs actions on objects.

Egocentric pantomimes, being full-body actions, are specialized for representing actions occurring in *peripersonal* space, while allocentric pantomimes are specialized for representing actions taking place in *extrapersonal* space. For example, if someone wanted to pantomime a peripersonal action like eating soup, he would most likely IO the action such that his dominant hand holds an invisible spoon that

brings the invisible soup to his mouth. However, if he wanted to pantomime an extrapersonal action like two cars colliding head-on, he would most likely form each hand into the shape of a fist to BPO each of the two cars, and then strike his two fists together to represent the head-on collision between the two cars. In general, egocentric gestures are more transparent to viewers than are allocentric gestures (Beattie and Shovelton, 2002), since the latter require the symbolic act of body-part replacement, whereas the former do not. **Figure 2** shows the classic scheme of pantomime classification, in which IO is the standard pantomime from the character viewpoint, and in which BPO is the standard pantomime from the observer viewpoint.

THE STATUS OF PANTOMIME IN GESTURAL THEORIES OF LANGUAGE ORIGIN

Theories of language origin can be dichotomized into "vocal" and "gestural/visual" models (MacNeilage, 1998; Corballis, 2002, 2009; MacNeilage and Davis, 2005; Armstrong and Wilcox, 2007; Arbib, 2012; McGinn, 2015; Ferretti et al., 2017; Zlatev et al., 2017), as well as multimodal models that posit a synergistic relationship between vocalization and gesture (McNeill, 2012; Kendon, 2017). Gestural models, broadly understood, posit that visually-conveyed symbols evolved earlier than those produced vocally, and that speech was a replacement for a pre-established symbolic system that was mediated by gestures alone. Importantly, the kind of gesturing that most gestural models allude to is pantomime, or iconic gesturing (Hewes, 1973; Armstrong and Wilcox, 2007; Tomasello, 2008; Corballis, 2009; Arbib, 2012; McGinn, 2015; Zlatev et al., 2017), which is generally considered to be manual, although some gestural models also include non-pantomimic manual gestures (pointing: Tomasello, 2008) and even non-manual gesturing (orofacial gesturing: Corballis, 2002; Orzechowski et al., 2014). Iconic gesturing through pantomime is thought to have predated symbolic

FIGURE 2 | The standard classification of pantomime. The figure presents the standard classification of pantomime. The imaginary object (IO) pantomime is the standard type of pantomime from the character viewpoint, while the body-part-as-object (BPO) pantomime is the standard type of pantomime from the observer viewpoint.

gesturing in the emergence of referential communication, passing through an intermediate stage that Arbib (2012) refers to as "protosign." Gestural models are supported by the fact that non-human primates show only the most limited use of iconicity in their gestural communicating in the wild, despite their strong proclivity for performing manual actions (Cartmill et al., 2012; see Perlman and Gibbs, 2013, for gorilla gestures labeled as pantomimes but that employed objects).

To our mind, a drawback of virtually all gestural theories of language origin is that they view pantomime as a unitary process, and do not make distinctions between various manners of producing it. Our point here is to not criticize such models but to highlight an opportunity to expand the scope of gestural theories so as to develop more-nuanced evolutionary models. These theories would benefit from an analysis of the different means by which pantomiming occurs, since this could help explain how pantomimic communication could have emerged during human evolution. To the extent that different types of pantomimes are specialized for depicting specific kinds of referents, then evolutionary proposals regarding the types of objects or actions that pre-linguistic humans might have first communicated about would provide insight into the type of pantomiming that they might have used to convey such objects or actions.

In an influential model, Donald (1991) proposed that language was preceded by a stage of human evolution that he referred to as "mimetic culture." However, Donald's sense of "mimetic" covers both imitation and pantomime. For him, a major part of mimetic skill is imitative learning as a road to complex praxis, as related principally to tool technology. It is about the ability to store exemplar representations of motor actions in long-term memory and then voluntarily bring them up during bouts of rehearsal or pedagogy (Donald, 1993, 2013). As he writes (Donald, 1991:169): "When there is an audience to interpret the action, mimesis also serves the purpose of social communication. However, mimesis may simply represent the event to oneself, for the purpose of rehearsing and refining the skill…that is, represented to oneself. Mimesis is not absolutely tied to external communication." Hence, Donald's concept of mimetic culture straddles the divide between imitative processes underlying rehearsal for skill learning and referential processes for social communication. Arbib (2012) too talks about a stage of "complex imitation" that preceded the emergence of language, not unlike Donald's concept of mimetic culture. He argues that this stage was advantageous for "sharing acquired praxis knowledge…independently of any implication for communication" (p. 176). His model posits that this non-communicative stage fed directly into a stage of pantomime production and thus a stage of referential communication based on iconicity. Gärdenfors (2017) argued, in agreement with Donald (1991), that demonstration for praxis is the primary function of pantomime, and that the communicative function is secondary. However, it is not clear why demonstration would depend upon what Gärdenfors calls a "caricature" of an action, rather than the real action. Given that tool-use pantomimes are empty-handed gestures, it is unclear why the evolution of demonstration in humans would favor an empty-handed version over the actual action.
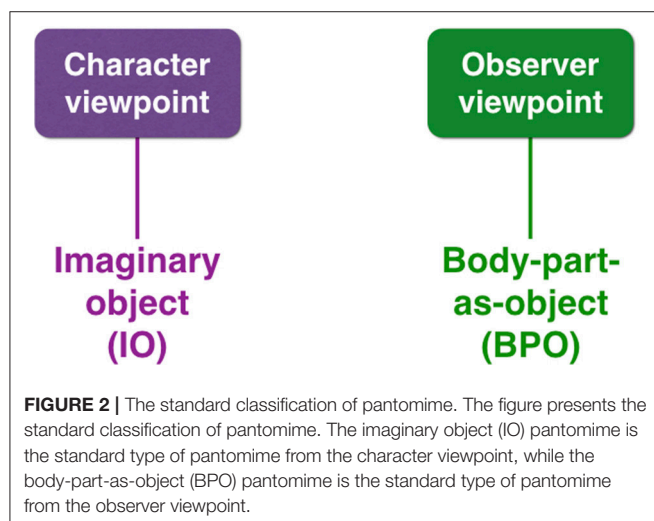
A CLASSIFICATION SCHEME FOR PANTOMIME

The current section provides a detailed presentation of our proposal for a new classification scheme for pantomime, which is comprised of five major categories. As shown in **Figure 3**, there are two principal dimensions to the scheme, although the dimensions are not neatly crossed with one another. The first one is whether the gesture occurs in egocentric space or allocentric space, and the second one is whether the gesture involves body-part use or body-part replacement. Note that both categories of body-part use are egocentric, while forms of body-part replacement can be either egocentric or allocentric depending on the space in which they occur. **Table 1** provides a detailed comparison among the five types of pantomime with regard to "what is mimed" and "how the miming is done." The examples listed at the top of each column are displayed in photographs in the **Appendix**.

Intransitive Action (IA)

We propose a new type of pantomime, called the Intransitive Action (IA) that is the intransitive counterpart to the IO pantomime. IA is a body-part use pantomime, but one that is devoted to intransitive actions, compared to the transitive actions depicted in IO's. Examples of IA's include walking, swimming, and thinking. Hence, these might be the closest things to the full-body pantomimes that Zywiczynski et al. (2018) talk about. We also include in this category gestures of emotional expression, including both facial expressions and expressive body movements. Finally, we include standard emblems that are used in interpersonal communication, such as the thumbs-up gesture or the peace sign. In the sub-section below entitled "Mimicry as a problematic case," we explain how the IA category leads to analytical complications when it comes to someone mimicking the gestures of another person or an animal.

FIGURE 3 | Proposed new classification scheme for pantomime. Five categories of pantomime are shown, divided into egocentric and allocentric varieties. Body-part use pantomimes are egocentric, whereas body-part replacement pantomimes can be either egocentric or allocentric, depending on the space depicted. Not shown in the figure is pointing, which acts as an adjunct to tracing when specifying object trajectories and locations. A, allocentric; BPO, body-part-as-object; E, egocentric.

Imaginary Object (IO)

IO is the classic pantomime that is performed from the character viewpoint. It encodes transitive actions using empty-handed gestures acting on invisible objects. IO's are basically tool-use gestures for handling objects. The hands assume the shape of the grip that would be used to interact with the mimed objects. In addition, hand dominance is generally abided by. Hence, in miming a tennis serve, the mimer's dominant hand would form a closed grip appropriate for a tennis racquet, while her non-dominant hand would form an open-handed grip appropriate for holding a tennis ball. Hence, an IO is a *grasp-shaping* gesture. IO and IA collectively define the body-part use pantomimes. Because IO's are egocentric gestures, they reference the full body. Hence, even a manual gesture like pantomiming a tennis serve is ultimately anchored to the body core, as suggested by Zlatev (2014) and Zywiczynski et al. (2018).

Body-Part-as-Object (BPO)

In addition to IA, we propose the existence of a second new class of pantomime by dividing the BPO category into two subtypes: BPO-E is performed in egocentric space, while BPO-A is performed in allocentric space (see **Figure 3**). Both types are object-*embodying* gestures in which the mimer's body part comes to embody an object through body substitution. A BPO-E is anchored to egocentric space by virtue of its combination with an IO done by the other hand, or as a BPO (one- or two-handed) performed in peripersonal space next to the body, such as holding binoculars or cutting hair. Note that the classic scheme (**Figure 2**) categorizes this as an observer-viewpoint pantomime since the gesture is not performed from the first-person perspective, which is one reason why we opt for the egocentric/allocentric terminology in place of CV/OV. The "call me" pantomime is perhaps the most common everyday example of a BPO-E. BPO-E's are defined by their peripersonal location. This is in contrast to BPO's that are clearly in extrapersonal space (e.g., two cars colliding) or that occur in some indeterminate space, which collectively define our BPO-A category. For example, when someone mimes a bowl by shaping their hand as a bowl, it is not clear if that bowl is supposed to exist in peripersonal or extrapersonal space. To the extent that the space is uncertain, then we classify it as a BPO-A. BPO-E's are almost exclusively hand-held objects, such as tools or vessels. They therefore support the depiction of transitive actions and thus tool use and object handling. BPO-A's, by contrast, can be any type of object, including people and large objects like airplanes and planets that are non-manipulable. The same object can straddle both categories. If a person shapes their hand into a bowl while miming ladling soup into a bowl, then the bowl is a BPO-E, since the IO of ladling anchors the BPO to peripersonal space and makes the mime into a transitive action. However, if the mime is simply of a bowl by itself, with no reference to peripersonal space, then the bowl is now a BPO-A because of the uncertainty of the spatial reference frame. BPO-E and BPO-A collectively define body-part replacement pantomimes. The result of this new distinction is that BPO's are no longer purely allocentric gestures, as in the standard observer-viewpoint classification (**Figure 2**). They can be either egocentric or allocentric gestures
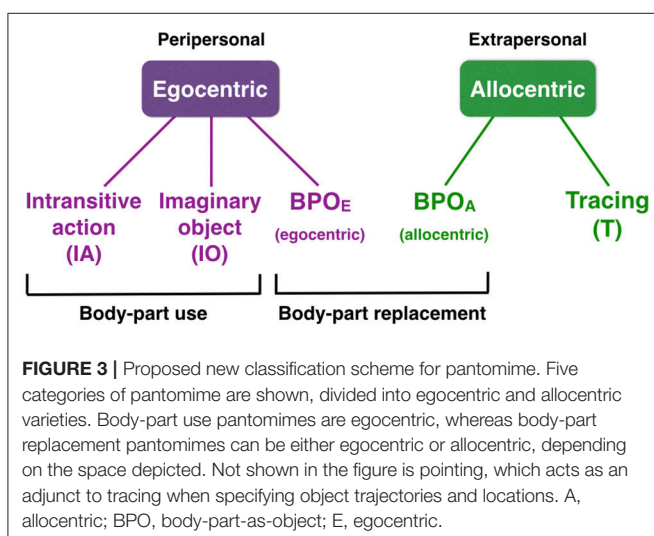
TABLE 1 | Comparison of the five types of pantomime.

	IA	IO	BPO-E	BPO-A	Tracing
	Thinking	**A tennis serve**	**Cutting hair**	**An airplane**	**Outlining a box**
WHAT IS MIMED					
What is mimed	Action (intrans.)	Action (trans.)	Object/action	Object/action	Object
Human actions	Yes	Yes	No	Yes	Yes
Transitive actions	No	Agent of	Recipient of	Generally no	No
Manipulable objects	No	Agent of	Yes	Yes	Often yes
Non-manipulable objects	No	No	No	Yes	Yes
Static or moving objects	Neither	Both	Both	Both	Both
Actionless objects	No	No	No	Yes or no	Yes or no
Non-body actions	No	No	Yes	Yes	No
Self-as-other	Mimicry	Mimicry	Object	Object/person	No
Affective expression	Yes	No	No	No	No
HOW MIMING IS DONE					
Reference frame	Egocentric	Egocentric	Egocentric	Allocentric	Allocentric
Space	Peripersonal	Peripersonal	Peripersonal	Extrap./indet.	Extrapersonal
Body-part use	Yes	Yes	No	No	No
Body-part replacement	No	No	Yes	Yes	No
Empty-handed action	No	Yes	No	No	No
Shaping	No	Grasp shape	Object shape	Object shape	Obj./traject.
Spatial reduction	No	No	Yes or no	Yes or no	Yes or no
Absolute/relative size	Absolute	Absolute	Relative	Relative	Relative
Tense, time frame	Not-present	Not-present	All tenses	All tenses	All tenses

The table shows general trends, not absolutes. For example, tracing is generally extrapersonal, but it can be peripersonal on rare occasions. Manipulable objects are hand-held objects, while non-manipulable objects are generally large objects. A, allocentric; BPO, body-part-as-object; E, egocentric; extrap., extrapersonal; IA, intransitive action; indet., indeterminate; intrans., intransivie; IO, imaginary object; obj., object; traject., trajectory; trans., transitive.

depending on the space in which they occur. They therefore straddle what would be CV and OV in the classic scheme. Notice that while IO's, IA's, and BPO's can be used to represent human actions, only BPO's can be used to represent non-body (i.e., non-human) objects, such as inanimate objects. In addition, BPO's can represent either static objects or moving objects; the same is true of tracing gestures, as described next. Compared to the grasp-shaping nature of IO's, BPO's are *object-embodying* gestures. Some cases of BPO-E's straddle this distinction. For example, when someone mimes using binoculars, it is difficult to distinguish if they are IO'ing the act of holding binoculars or if they are BPO'ing the embodiment of binoculars themselves, since the two can look quite similar to an observer. Finally, while most BPO-A's are hand or arm gestures, it is also possible to use the full body to produce them. For example, a standing posture in which both arms are stretched outright perpendicular to the body could be used to depict a tree or an airplane as full-body replacements.

Tracing (T)

Tracing involves using part or all of the hand—most commonly an index finger—to trace the outline of an object's shape or its movement trajectory. While tracing on rare occasions occurs in peripersonal space, such as when someone traces the spatial extent of a bruise that used to be on her arm, it is generally done in a frontal manner to connote extrapersonal space, such as when outlining the size of a box that is needed in order to

hold an object. This can be done in a one-handed or two-handed (generally symmetric) manner, and it is often done as a planar action in two dimensions (i.e., the coronal plane). Tracing is an iconic object-*outlining* gesture in the same way that BPO's are object-embodying gestures. This creates a temporal difference between tracing and BPO's with regard to their depictions of objects. Whereas, a BPO typically represents a persistent visible object by means of body-part replacement, a tracing gesture can create a transient outline of the object that does not persist. Like BPO's, tracing gestures can be static when depicting an object's shape and/or size, or they can be dynamic when mapping out the trajectory of a moving object, such as when showing an object spiraling down to the ground. While most tracing employs a finger as the effector (most often the index finger), tracing can also involve the whole hand, such as when someone traces the shape of a sphere with symmetric use of two curved hands, or uses their two hands to trace out a large semicircle in front of their abdomen to depict pregnancy. Tracings are object-outlining gestures, and whether a full hand or a finger is used for tracing depends on the nature of the outline being shaped.

Pointing

Pointing, also called deixis in the gesture literature, is one example of what Ekman and Friesen (1969) refer to as an "illustrator" gesture when it occurs in the context of speech. Pointing occurs either as an imperative gesture, where the

pointer commands an action on the part of the addressee, or as a declarative gesture, where the pointer draws attention to an inter-subjective object or action (Brinck, 2004). Declarative pointing is a means of drawing attention to the *occurrence* or *location* of an object or action, rather than to its spatial structure (Haviland, 2000; Bangerter, 2004). Pointing can be used to indicate the location of an object in a depicted scene, whereas tracing can be used to demonstrate the trajectory of an object's movements. In addition, pointing can be used to demonstrate the *endpoints* of a movement trajectory, thus making it a substitute for tracing. For example, a tourist traveling on a train in a foreign country might gesturally beseech someone for help in getting a heavy suitcase onto a luggage rack by first declaratively pointing to the suitcase and then imperatively pointing to the luggage rack above, with the implication being that he is tracing out a movement trajectory between the ground and the luggage rack. In this example, pointing indicates an object and a location, while tracing indicates a trajectory to get there. In a fundamental sense, tracing is related to pointing in that, while pointing is a static gesture, tracing can be its dynamic counterpart. The combination of pointing and tracing might depict objects that are out of reach or, as in the case of the luggage rack on a train, target locations that are inaccessible to the sender. With respect to human evolution, imperative pointing may have emerged out of the act of failed reaching (Cochet and Vauclair, 2010), whereas declarative pointing might have required broader cognitive skills related to a sharing of social context in referential communication.

To summarize, our scheme has several novel features compared to the classic scheme. (1) The concepts of egocentric and allocentric are borrowed from the field of spatial cognition as replacements for character viewpoint and observer viewpoint, respectively. (2) IA is proposed as the intransitive counterpart to IO. It includes the depiction of emotional expression and the production of emblems. (3) A new binary distinction is proposed for BPO, where BPO-E is the egocentric and peripersonal counterpart to BPO-A, which maps onto the standard conception of BPO. (4) Tracing is proposed as a type of pantomime since it is an iconic gesture that maps out both spatial structure and movement trajectory. Tracing is similar to BPO-A, except that it outlines an object, rather than embodying it. We are agnostic with regard to whether pointing should be classified as a sixth category of pantomime, but we do highlight the fact that pointing can be used as an adjunct to tracing in order to signify movement trajectories. It is also used extensively in metaphorical gestures, for example signifying the past by pointing in the backward direction (Cartmill et al., 2012; e.g., Casasanto and Jasmin, 2012).

Mimicry as a Problematic Case

The critical distinction between body-part use and body-part replacement pantomimes is that, in a body-part use pantomime, the mimer's hand is depicting a hand, whereas in a body-part replacement pantomime, the mimer's hand is depicting *some other object*. This "hand identity criterion" allows us to categorically distinguish body-part use and body-part replacement pantomimes in almost all cases. However, mimicry provides strong challenges for this distinction. If Mary does a

tennis-serve IO during a conversation, she might be depicting herself doing it or she might be depicting her tennis partner doing it. If she is representing her own serve, then her hand is depicting her own hand. However, if she is miming her partner's serve, then her hand is *not* depicting her own hand. Hence, this violates the "hand identity criterion." One could argue that her hand is still depicting a hand, even if it is not her own hand. However, is the mime occurring in the mimer's peripersonal space or in the allocentric space of another person, since the mimer's hand is being replaced by her partner's hand in the pantomime? An even more complicated example is when a person flaps their arms to represent a bird flying. In this case, the mimer's arms are not even depicting human arms, and so the gesture could not possibly be in the mimer's peripersonal space. But bird wings are evolutionarily homologous to mammalian arms, and so arms are biologically equivalent stand-ins for wings when it comes to pantomiming. By contrast, miming the mouth movements of an animal would be seen as an allocentric gesture, such as using two horizontally-positioned hands to create an opening and closing movement. The bottom line is that when a person does a body-part use pantomime of their own actions, there is no question that the mime is occurring in an egocentric manner. However, when someone impersonates another person or an animal by doing a body-part use or similar pantomime, it is no longer clear whether this is occurring in the mimer's egocentric space. Mimicry thus makes the egocentric/allocentric distinction indeterminate or mixed. This places it in a class by itself, a fusion category. Beyond this "first-order" mimicry of the mimer impersonating someone else, we can also imagine situations of "second-order" mimicry in which the mimed person is impersonating yet some other person. For example, a male mime actor could pantomime a man who was momentarily mimicking the gestures of a woman he had seen. We will return to this general issue in the Discussion section, since mimicry has important implications for the origins of role playing in human life (Brown, 2017) as well as for the origins of language. It is important to point out that the concept of a "character viewpoint" in the gesture literature is based on studies in which participants are explicitly asked to convey the actions of characters, for example cartoon characters (Beattie and Shovelton, 2002; McNeill, 2005). This first-order mimicry and its inherent blurring of the distinction between character and self are never discussed in these studies. We believe that it is an important issue that needs to be addressed.

Two-Handed Analysis

In addition to describing the pantomime types as specific categories, we attempt to provide principles for how the hands combine in creating two-handed pantomimes. The two-handed analysis is based on the scheme outlined in **Table 2**. Terminology is introduced here to aid in understanding the table. For two-handed IO's and BPO's, a pantomime is said to be "double" if the two hands represent *two different objects* (e.g., a tennis racquet serving a tennis ball [double IO]; a pen writing on a pad [double BPO-E]). A pantomime is referred to as "joint" if the two hands represent or contribute to *a single object* (e.g., lifting a large box [joint IO]; rain falling [joint BPO-A]; outlining a sphere [joint tracing]). Combinations of different types of pantomimes

by the two hands are referred to "intra-category mixes" if both hands perform either egocentric pantomimes alone or allocentric pantomimes alone (e.g., ladling soup into a bowl [IO/BPO-E, where both are egocentric]). Combinations of different categories of pantomimes by the two hands are referred to as "inter-category mixes" if one hand performs an egocentric pantomime while the other hand performs an allocentric pantomime (e.g., pressing a launch button [IO, egocentric] to make a rocket take off [BPO-A, allocentric]).

SOME PRINCIPLES OF HOW PANTOMIME WORKS

A major objective of our analysis is to establish principles for associating pantomime types with the kinds of action or objects being depicted. The current section discusses some general considerations, while the next section focuses on hand use during two-handed pantomime production.

Body-Part Use Pantomimes (IO and IA) Can Only Be Used for Body Actions

These can be either transitive (IO) or intransitive (IA) actions. However, they can only represent actions *per se* (including object-directed actions), and not objects alone. Using a cupped hand to IO the act of holding a bowl in one's hand is the mime of the transitive action of holding. A nearly identical hand position could be used to BPO a bowl itself, rather than the act of holding one. However, the BPO is no longer a body-part use pantomime. Body-part use pantomimes can only be done for human agents and their actions, but not for the depiction of inanimate objects or non-human animate objects on their own (although see the caveat above about bird-flying mimicry).

Non-body (Non-human) Actions Can Only Be BPO's or Tracings

The corollary to the last point is that non-body objects can only be represented with BPO's (either BPO-E or BPO-A) or tracings, depending on the space in which they occur. However, BPO-A's can also represent human bodies and their movements, just like IO's and IA's, although they do so using body-part replacement gestures. As will be discussed below, BPO-A's and tracings are the preferred manner of depicting human actions when the body's core is dynamic, such as in jumping over a tennis net or doing a back flip. Compared to IO's and IA's, BPO's and tracings are able to represent objects themselves as well as their movement patterns. An important observation about this is that *large objects have to be BPO'd*. IO's can only depict manipulable objects, and so BPO's are needed in order to depict large, non-manipulable objects, like airplanes, office buildings, and planets. The same is true of extremely small objects, like atoms. Next, dynamic actions for non-human objects have to be represented allocentrically using BPO's and tracings, sometimes with the assistance of pointing. For example, two cars colliding cannot be mimed egocentrically. Egocentric pantomimes can only be done for human actors, but not for inanimate or non-human objects, the major exception being BPO-E's, which can depict

TABLE 2 | Major types of combinations for two-handed pantomimes.

	Two-handed IO		Two-handed BPO		Two-handed tracing	Intra-category mixes (ego/ego or allo/allo)			Inter-category mix (ego/allo)
Name	(1) Double IO	(2) Joint IO	(3) Double BPO	(4) Joint BPO	(5) Joint tracing	(6) IO/BPO-E	(7) IA/ BPO-E	(8) IO/ IA	(9) IO/BPO-A
Definition	2 IO's mimed with 2 hands	1 IO mimed with 2 hands	2 BPO's mimed with 2 hands	1 BPO mimed with 2 hands	1 Tracing mimed with 2 hands	1 IO + BPO-E mimed with 2 hands	1 IA + BPO-A mimed with 2 hands	1 IO + 1 IA mimed with 2 hands	1 IO + BPO-A mimed with 2 hands
No. objects	2 objects	1 object	2 objects	1 object	1 object	2 objects	1 object But not joint	1 object But not joint	2 objects
Example	Serving a tennis ball (IO) with a tennis racquet (IO)	Picking up a large box with both hands (IO's)	A skier (BPO-A) skiing down a ski slope (BPO-A)	Looking through a pair of binoculars (BPO-E's) Fire; rain (BPO-A's)	Outlining the shape of a box	Ladling soup (IO) into a bowl (BPO-E)	Looking through binoculars (BPO-E), a lifeguard recoils (IA) in seeing someone drowning	A lifeguard swims (IA, left hand) while pulling a victim (IO, right hand) back to shore	Pressing a launch button (IO) to make a rocket take off (BPO- A)

The table is not meant to be exhaustive, but aims to describe a variety of combination types. In the Example row, objects are written in italics. A, allocentric; allo, allocentric; BPO, body-part-as-object; E, egocentric; ego, egocentric; IA, intransitive action; IO, imaginary object.

inanimate, manipulable objects in peripersonal space during the representation of transitive actions on them (e.g., using the fingers to BPO scissors cutting one's hair).

Extrapersonal Actions Can Only Be BPO's or Tracings

Extrapersonal space defines the BPO-A subtype within our BPO category. BPO-A's can be static or dynamic. Tracing can be used for extrapersonal objects to convey their outlined shape and movement trajectory.

Many Pantomimes Are Shaping Gestures

IO's are grasp-shaping gestures, BPO's are object-embodying gestures that conform with an object's shape, and tracings are object-outlining or trajectory-outlining gestures. IO's have a grip quality to them since they are tool-use gestures. When miming a tennis serve, there will be a closed grip for the dominant hand (the racquet) and an open grip for non-dominant hand (the tennis ball).

Scaling

Egocentric pantomimes are usually to-scale mimes that represent spatial dimensions in the same way that they would appear in real-life actions. Allocentric pantomimes, by contrast, have the potential to be spatial reductions or expansions of the objects and actions being depicted, resulting in spatial dimensions that are completely out-of-scale. Non-manipulable objects have to be BPO'd, and this creates a point of distinction between BPO-A and BPO-E. BPO-E's are basically manipulable objects that are used in transitive actions. BPO-A's are generally larger or smaller objects than those that are manipulable. A BPO-A (perhaps along with tracing) is the only way to mime a car, ski slope, or galaxy.

PRINCIPLES OF TWO-HANDED USE

To the best of our knowledge, the current analysis is the first one to characterize two-handed pantomime use. Some major patterns are summarized graphically in **Figure 4**.

Hand Dominance

When pantomiming transitive body actions—using either double IO's, double BPO-E's, or IO/BPO-E mixes—the dominant hand carries out the action that it would if it were actually carrying out the action. Hence, when miming a tennis serve (double IO), the dominant hand will mime holding the tennis racquet, while the non-dominant will mime throwing up the ball. For an IO/BPO-E mix like ladling soup into a bowl, the dominant hand will mime holding the ladle, while the non-dominant hand will BPO-E the bowl. For a double BPO-E like a pen writing on a pad, the dominant hand will embody the pen (i.e., the dynamic element), while the non-dominant hand will embody the pad (i.e., the passive element).

One or Two Objects

Two-handed mimes can represent one object jointly or two distinct objects. (1) A two-handed IO can mime one imaginary object (holding a large box with both hands) or two imaginary objects (doing a tennis serve, with the non-dominant hand throwing the ball upward). If two objects are presented, then the dominant hand will be the hand that mimes holding the tool. (2) A two-handed BPO can mime one object (fire, rain, binoculars) or two objects (a plane landing on a runway; a pen writing on a pad). This applies equally to two-handed BPO-A's and two-handed BPO-E's. (3) Two-handed tracings generally mime one object jointly (a sphere, a pregnant woman's belly).

Miming Two Objects

There are four ways to mime two different objects. They comprise combinations of imaginary hand-held objects (IO) and embodied objects (BPO). (1) Double IO = hand-held + hand-held object. (2) Double BPO = embodied + embodied object (within the same space). (3) IO/BPO-E = hand-held + embodied objects in the same space. (4) IO/BPO-A = hand-held + embodied objects in different spaces, which is extremely uncommon.
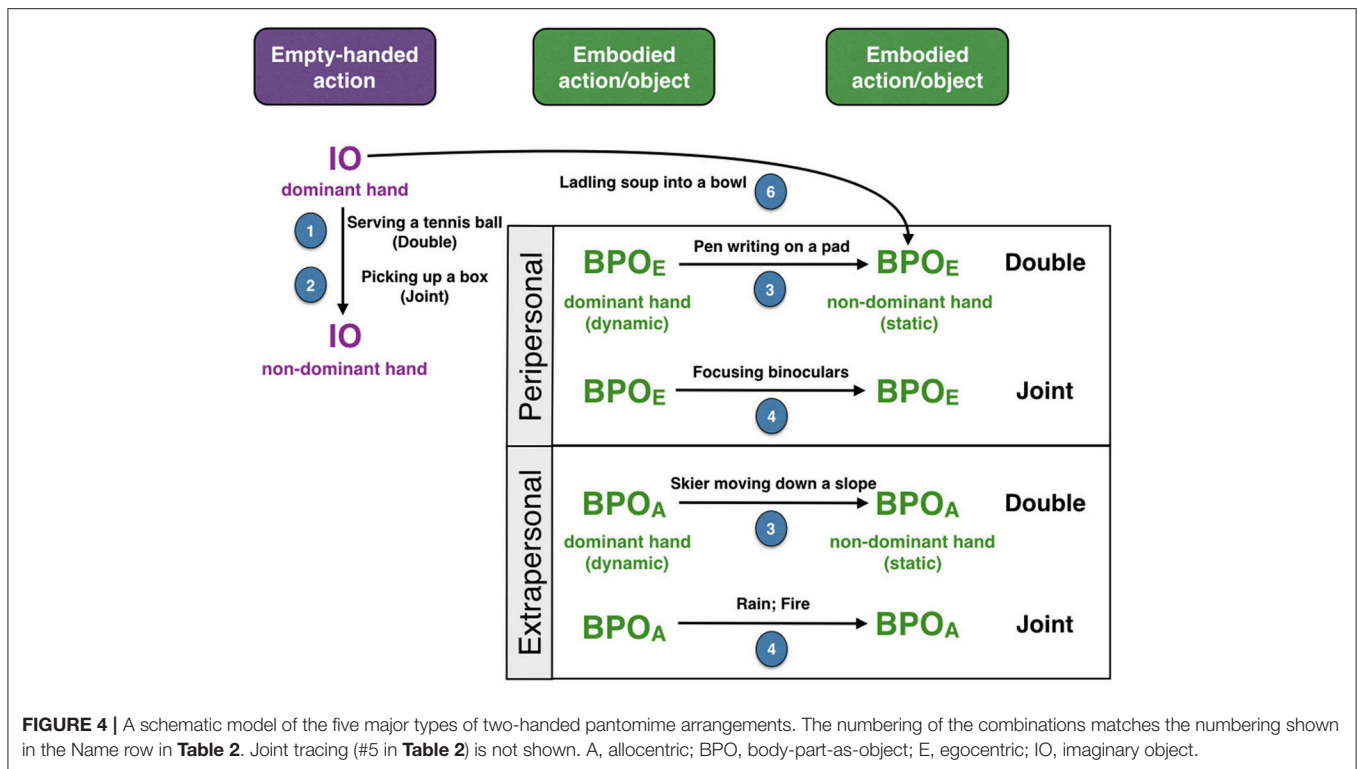
Dynamic/Static Pairing

For an IO/BPO-E mix, the IO is a dynamic action with an imaginary tool, while the BPO is often a static object. In other words, the BPO is generally the static recipient of the transitive action. As mentioned above, the dynamic IO gesture will be done with the dominant hand, while the static BPO-E will be done with the non-dominant hand.

The Agent/Patient Principle

The previous observations can be consolidated into a principle of relevance to language and syntax. For transitive actions, the dominant hand tends to IO the agent's action, whereas the non-dominant hand tends to BPO the patient (i.e., the recipient of the transitive action), as in ladling soup (IO) into a bowl (BPO-E). This defines a BPO-E as the patient of an IO's action. Transitivity is mainly conveyed using a double IO, joint IO, or an IO/BPO-E mix. Less commonly, it can be conveyed using a double BPO-A, such as using a flattened hand to represent a tennis racquet hitting the fist of the other hand representing a tennis ball. However, we would strongly expect a person to mime the act of hitting a tennis ball egocentrically using a double IO (with the non-dominant hand throwing up an imaginary tennis ball, combined with the dominant hand gripping an imaginary tennis racquet), rather than as a double BPO-A.

Shared Space Principle

Two-handed pantomimes will preferentially use the *same* spatial category of gestures—either both egocentric or both allocentric—rather than different categories (ego/allo mixes), since *mimes within a category share the same physical space* (either peripersonal or extrapersonal) and can therefore be coupled in a natural manner within that shared space. This applies comparably to two-handed IO's, two-handed BPO's, two-handed tracings, and intra-category mixes. This applies as well to the subdivisions of BPO: BPO-E will more naturally combine with BPO-E, and BPO-A will more naturally combine with BPO-A, since they occur in the same physical space. The most natural mixture is an IO/BPO-E, which might underlie the subject-verb-object word order in gestural models of syntax discussed

FIGURE 4 | A schematic model of the five major types of two-handed pantomime arrangements. The numbering of the combinations matches the numbering shown in the Name row in **Table 2**. Joint tracing (#5 in **Table 2**) is not shown. A, allocentric; BPO, body-part-as-object; E, egocentric; IO, imaginary object.

below. IO can couple with BPO-A as well, but the mimed objects occur in very different physical realms, such as when pushing a launch button with one hand (IO) makes a rocket launch with the other hand (BPO-A). It is telling that this example depicts sequential gestures, since it would probably be unreasonable to have simultaneous miming occurring in two distinct spaces.

Static vs. Dynamic Core Principle of Egocentric vs. Allocentric Pantomiming

When it comes to depicting human actions, there is a tendency to IO or IA actions that involve a static body core, whereas there is a tendency to BPO-A or trace out actions that require body-core movements. When talking about playing tennis, it is common to mime this using an IO, with an imaginary tennis racquet in the dominant hand. However, when talking about jumping over the net upon winning a match, few people would be inclined to IA a jumping-over action with their full body. They would be far more likely to BPO-A this action, perhaps using the non-dominant hand as the net and then two fingers of the dominant hand as their legs engaged in a jumping-over action, or alternatively using a finger to trace out the body's trajectory in jumping over the net. Likewise, if a person were to mime having done a back flip while visiting a trampoline park, they would unquestionably mime that action by tracing out a circular path with their finger, rather than performing a back flip. Therefore, a major determinant of the type of bodily pantomime that a person does is the space occupied by the action, and most especially *whether the body core is static or dynamic*. It is far easier for someone to trace out a circle with a finger than it is to simulate a flip with one's full body.

While IO's are ideal for full-body actions (like playing tennis), they only really work well for actions in which the mimer's core is static. When the action requires a dynamic core, such as in jumping over a tennis net or doing a back flip, it is far more economical to do a spatial reduction of the action using an allocentric gesture. That probably makes sense in terms of pantomime's function in conversation. When we converse with people, we tend to be close to them and are often times seated. It doesn't make sense to run away from them in order to mime jumping over the net at a tennis match. Likewise, it doesn't make sense to move down an inclined surface to pantomime skiing down a ski slope, since moving a finger down an inclined forearm can do this so much more efficiently. Therefore, a consideration that comes strongly into play when deciding between egocentric and allocentric pantomimes of human actions is how static or dynamic the core of the body tends to be in the depicted action.

IMPLICATIONS FOR THE ORIGINS OF LANGUAGE

We would now like to consider the implications of the classification scheme for models of language origin, since numerous thinkers have made the case for a pantomimic origins of language (e.g., Hewes, 1973; Armstrong and Wilcox, 2007; Tomasello, 2008; Corballis, 2009; Arbib, 2012; McGinn, 2015; Ferretti et al., 2017; Zlatev et al., 2017), but without discussing the diversity of manners by which pantomimes can be produced. Before proceeding, we would like to present the caveat that this discussion will not be considering either the

origins of pantomime itself or how pantomime becomes later conventionalized into a sign system for language (e.g., Arbib's "protosign"), but will simply explore alternative accounts of a presumed pantomimic stage of language evolution. Abramova (2018) argued that pantomime is too complex of a precursor for language since it begs the question of the origins of both symbolization and intentional communication that are present in pantomime-based interactions. She proposed instead that ontogenetic ritualization is a better precursor for both pantomime and language (see responses in Zlatev, 2018 and Arbib, 2018). While this issue is an important one, we will not address it further here.

Our pantomime scheme highlights two distinct modes of creating pantomimes: egocentrically in peripersonal space, and allocentrically in extrapersonal (or indeterminate) space. This egocentric/allocentric distinction not only implies different ways of miming, but also different *semantic categories* of depiction, with important ramifications for gestural models of language origin. As discussed previously, egocentric mimes generally only depict humans (in particular human actions), although see the caveats above about animal mimicry. Conferring evolutionary precedence to egocentric pantomiming would mean that the first referential communications of ancient humans were about *people*, including their tool-use actions. Allocentric mimes mainly depict non-human entities (e.g., animals, plants, inanimate objects), including their shape, size, movement trajectory, and location (with the help of pointing). Conferring evolutionary precedence to allocentric pantomiming would mean that the first referential communications of ancient humans were about *environmental* factors, such as predators, prey, plant food sources, and the natural landscape. This could relate to hunting, foraging, defense, and the like. Allocentric pantomiming can be about people as well, although only for actions occurring in extrapersonal space. This could be used, for example, to represent a rival group of humans situated in a nearby location.
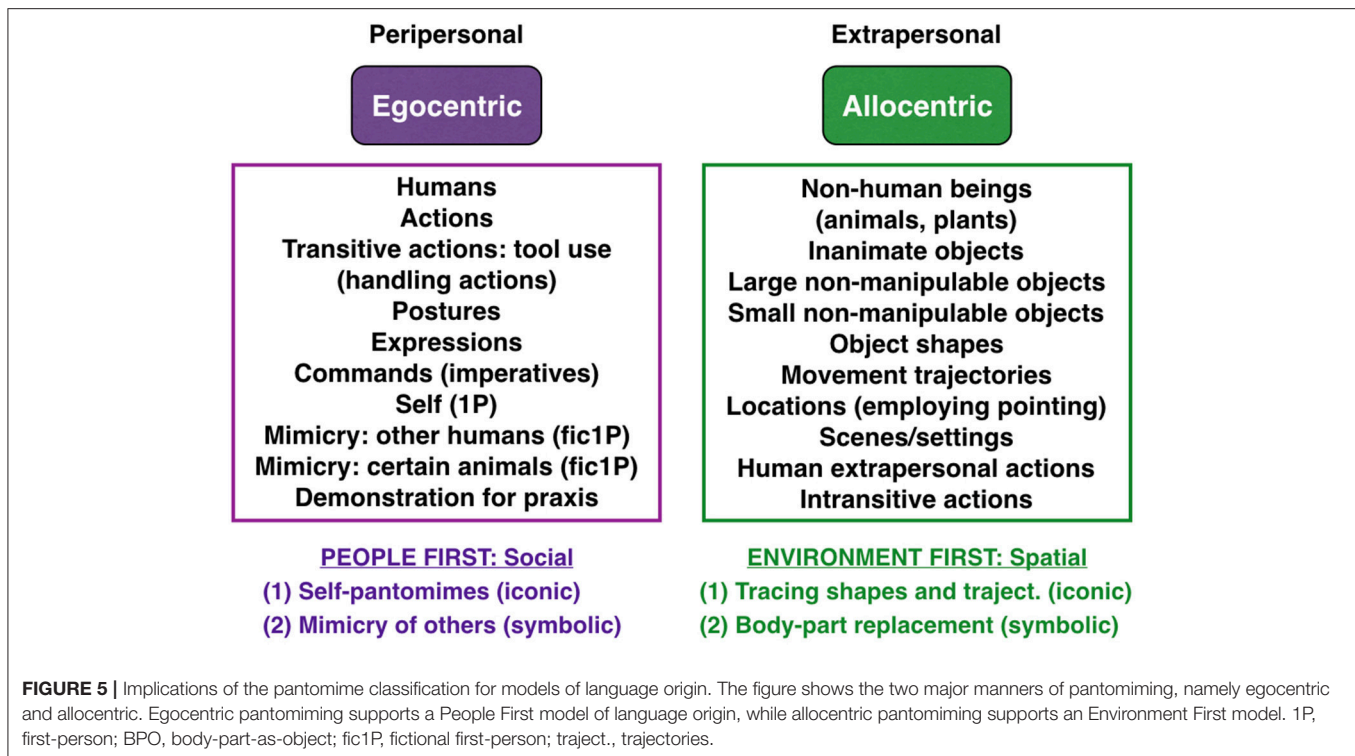
While it is quite doubtful that we could know whether early humans first communicated about people or the environment, the egocentric/allocentric distinction makes it clear that these two manners of pantomiming tend to emphasize different subject matters, at least to a first approximation. An egocentric model would favor a People First model, whereas an allocentric model would favor an Environment First model, which could perhaps be an Animal First model, since animals might be the most relevant non-human items to represent. A People First model would be consistent with Donald's (1991) model of mimetic culture, in which pantomimes could serve as a tool for demonstration during supervised learning for praxis (Gärdenfors, 2017).

We propose that each of the two evolutionary pantomime models can be broken down into (1) an earlier stage of miming that would be iconic and relatively transparent to perceivers, followed by (2) a later stage that would be more complex due to its symbolic nature (bottom part of **Figure 5**). For the People First model, the iconic stage would be the process of self-pantomiming, while the symbolic stage would be a process of personal mimicry, in other words egocentric miming while

depicting some other person. It seems reasonable to propose that self-miming from the first-person perspective would be more transparent to perceivers than would be an act of personal mimicry—the "fictional first-person" perspective of a portrayed individual—in which perceivers would have to come to terms with the full-body replacement of the mimer. Self-pointing could be a support for self-miming by reinforcing the fact that the mime is about the self. As mentioned above, personal mimicry complicated our analysis of pantomime, not least since mimicry can incorporate an element of recursion, such as when a person mimes somebody miming yet a third person. Despite these complications, personal mimicry is a highly significant cognitive and behavioral process that is ignored in all gestural models of language. Brown (2017) argued that personal mimicry via egocentric pantomiming provided the foundations for role playing and character portrayal in human life, leading ultimately to various forms of proto-acting and theatrical acting, including their expression as pretend play (Carlson and Taylor, 2005). Hence, we propose that egocentric pantomimes emerged initially as self-pantomimes and only later evolved as other-pantomimes through personal mimicry and character portrayal.

We propose a similar two-stage progression for allocentric pantomimes. For the Environment First model, the iconic stage would be tracing, while the symbolic stage would be BPO-based pantomiming, in other words body-part replacement. Tracing is a relatively transparent type of spatial gesturing since it aims to iconically depict object dimensions and movement trajectories through outlining. Pointing can be used as an adjunct to tracing in order to convey spatial locations or the endpoints of movement trajectories. This would be the most iconic type of allocentric gesturing, occurring in extrapersonal space. Some scholars have argued for the importance of pointing in both the evolution and development of communication (Brinck, 2004; Tomasello, 2008; Colonnesi et al., 2010), since pointing establishes a point of reference and stimulates joint attention. BPO's, by contrast, would be far more symbolic, since they would require the perceiver to comprehend that an act of body-part replacement had occurred. In addition, for large objects, perceivers would have to comprehend the spatial reduction inherent in the mime, since a BPO gesture can in theory be no larger than the body parts used to produce it, typically limited to a forearm, but potentially employing the full body. For example, using BPO's, a ski slope can be no longer than a forearm, and an airplane can be no larger than a full body.

A comparison of these two models leads to an important point of overlap. It is notable that the proposed symbolic stage in both the egocentric and allocentric models involves *body replacement*. It involves full-body replacement in the case of mimicry, while it is involves more-focal body-part replacement in the case of BPO's. Of these two, the full-body replacement of mimicry is far more transparent than the body-part replacement for BPO's, since the former can only reference a single type of object—a body, just as in the original—whereas a BPO has an unlimited number of possible referents, but has to be exclusive of the original body part. For this reason, BPO-A is the most complex pantomime of the five categories contained in the classification scheme.

FIGURE 5 | Implications of the pantomime classification for models of language origin. The figure shows the two major manners of pantomiming, namely egocentric and allocentric. Egocentric pantomiming supports a People First model of language origin, while allocentric pantomiming supports an Environment First model. 1P, first-person; BPO, body-part-as-object; fic1P, fictional first-person; traject., trajectories.

A final plank of our evolutionary thinking regarding the relative timing of the emergence of egocentric vs. allocentric pantomimes is that *BPO-E, being egocentric, would have evolved before BPO-A.* To the extent that the symbolic nature of BPO's limits their interpretability, having BPO's first appear as BPO-E's, perhaps in combination with IO's, would have narrowed down the possible scenic extent of the pantomime to peripersonal space, thus improving their interpretive transparency. Using an IO in combination with a BPO-E to indicate that the latter is the recipient of a transitive action would have greatly aided in interpreting the BPO'd object and thus increased the iconicity of the gesture by specifying its space of action. This proposal thus suggests an interactive origin of BPO's, arguing that the first BPOs were BPO-E's done in combination with IO's. Whether this occurred before or after the emergence of tracing as an allocentric pantomime is unclear.

While we lack the grounds for prioritizing the egocentric and allocentric evolutionary models between themselves, we can describe the unique specializations of each of the two modes of pantomiming, and consider what they reveal about the pantomimic precursors of language. Going further than this will depend upon theoretical arguments or empirical evidence regarding what ancient peoples initially needed to communicate about referentially, in other words the driving forces for the evolution of symbolic gesturing. Our analysis, if nothing else, establishes two alternative models for what the earliest referential communications might have been about.

Egocentric Advantages

Egocentric pantomimes are the principal ones used for communicating about human actions. This can be done more directly and with more detail using egocentric pantomimes than in using BPO-A's to depict humans, except in the case of extrapersonal actions or actions having a dynamic core. Egocentric pantomimes, such as IO and BPO-E, are specialized for depicting transitive actions. This creates an important link with language, since syntax communicates information about transitivity, and this often relates to tool use from a functional standpoint. The standard model of syntax represents various orderings of a subject (S), a verb (V), and an object (O) to create sentences (Tallerman, 2015). Many egocentric pantomimes—not least double IO's and IO/BPO-E mixes—convey transitive actions and thus SVO sequences. (Intransitive actions convey SV sequences.) A distinct advantage of egocentric over allocentric pantomimes when it comes to evolutionary models of syntax is that *the mimer is assumed to be the subject of the action.* The vast majority of languages abide by an Agent First structure in which S precedes either V or O in a sentence (which vary among themselves between SVO and SOV varieties; Jackendoff, 1999, 2002). Egocentric mimes place S front-and-center in the gesture and so conform with an Agent First structure. In order for an allocentric mime to convey transitivity, not only do two distinct object identities need to be established through body-part replacements, but conditions need to be specified to know who is acting upon whom (or what is acting upon what). This greatly reduces the transparency in conveying transitivity, as compared to an egocentric pantomime. In general, egocentric

pantomimes are perceived as being more transparent than allocentric pantomimes (Beattie and Shovelton, 2002), since no object substitution or scaling is required. This direct iconicity of egocentric gestures supports their communicative efficiency. In addition, transitivity for egocentric pantomimes can be easily conveyed with one hand, whereas it requires two hands for an allocentric pantomime, since one hand needs to mime S (or more accurately SV) and the other hand needs to mime O. We argued above that a reasonable evolutionary progression for egocentric pantomimes was to assume that self-pantomimes preceded acts of personal mimicry. If that were the case, then S would automatically default to the mimer, rather than some other person. The capacity to mimic other people would emerge as the next stage in the progression. Overall, transitivity has an important place in models of syntax (e.g., SVO ordering) as well as in models of tool use, such as Donald's (1991) concept of mimetic culture. Egocentric pantomimes are efficient communicators of transitive human actions, and thus seem better poised to serve as gestural precursors of syntax and of demonstrational learning for praxis, as compared to allocentric pantomimes. They are also good at communicating information about the self and others in order to support social cognition, as would be consistent with a social-grooming model of the origins of language (Dunbar, 1996). One disadvantage, however, is that "they require that the gesturer have a strong mental representation of the tool object involved in the action because there is no physical placeholder standing in for the tool" (Cartmill et al., 2012: 132).

Allocentric Advantages

Allocentric pantomimes offer a host of reciprocal advantages and specializations compared to those just mentioned for egocentric pantomimes (see **Figure 5**). Because egocentric pantomimes are limited to actions in the mimer's peripersonal space, allocentric pantomimes open up a huge scenic potential for different types of actions. In addition, because egocentric pantomimes are mainly limited to depicting human actions, allocentric pantomimes open up a new world of semantic possibilities by being able to represent objects and actions that have no direct connection with humans. Hence, to communicate information about a bird in a tree or a predator perched on a rock or a prey animal grazing on a savannah, only allocentric miming is available. The information that is conveyed can be about an object's size, shape, location, and/or movement patterns, making it very rich spatially. In addition, because of their scaling potential, allocentric pantomimes can display large objects, large spaces, or a combination of the two, creating the capacity for complex scene production using gesture (which would later evolve into the capacity for scene production using drawing, Yuan and Brown, 2014). Even though we argued that egocentric pantomimes are the richest and most transparent manner of depicting human actions, a desire to communicate the idea that a group of people from a hostile tribe is situated just over the hill from where we are located could only be conveyed scenically through allocentric pantomimes, including BPO's for the people and the landscape elements, and perhaps tracings to show the paths that the people could use to get to us or the paths that we could use to get away from them. Tracing gestures would contribute to this scenic detail since they are efficient as showing human actions that require a dynamic core, such as walking along a path or climbing up a tree. Overall, allocentric pantomiming greatly expands the semantic potential of spatial gesturing, and supports the cognitive process of "displacement" by which people are able to communicate about objects not immediately present (Hockett, 1960; Perniss and Vigliocco, 2014). It strongly diversifies the categories of objects that can be depicted (both large and small objects, manipulable or not), and it greatly expands the spatial realm of object-mediated actions by depicting extrapersonal spaces that can span from the spatial domain of two ants to that of two planets. Therefore, the combination of object diversity and scenic expansion makes allocentric pantomiming very flexible. Tracing is the most transparent version of this. By contrast, creating a BPO-A is the most opaque and complex of all the pantomime categories, since it requires both a specification of an object's identity through body-part replacement and often times a situating of that object in some extrapersonal space. This requires many more assumptions on the part of the perceiver than any other type of pantomime discussed in this article. Therefore, a certain amount of conventionalization would be required for BPO-A's to be adopted if they indeed preceded the emergence of speech, since BPO-A's are quite symbolic. In contemporary discourse, we use speech as a means of clarifying the content of our BPO's. We announce to our conversation partner that our right hand represents our car, and that our left hand represents the barrier that the car ran into on the highway. This may even be preceded by some scenic BPO's depicting the stretch of highway that we were driving on, as well as where the barrier was located in relation to the road. But all of this is cognitively complex without speech. BPO's offer a great deal of descriptive flexibility for gestural communication, but they come at a clear cost in terms of the transparency.

To summarize this section about language evolution, egocentric and allocentric manners of pantomiming offer different types of advantages in depicting different types of objects, actions, and functional spaces (Cartmill et al., 2012). While we are not in a position to have knowledge of what ancient humans first communicated about, we could reasonably predict what types of pantomimes they would have used to depict whatever they were communicating about. If we consider Donald's proposal of a pre-linguistic stage of mimetic culture, then the iconic forms of both egocentric and allocentric pantomiming (namely, self-pantomimes and tracing) could have initially served a demonstrational role in supervised learning, not least due to the transitivity that IO's bring to tool use in a culture presumably oriented toward tool technology. In addition, transitive forms of egocentric miming may have laid the groundwork for language syntax and its ability to convey transitivity through SVO sentence structures (Armstrong and Wilcox, 2007; McGinn, 2015).

The symbolic forms of pantomime would have led to distinct trajectories for egocentric and allocentric pantomiming. The symbolic form of egocentric miming (personal mimicry) would have expanded the depictional complexity of social cognition, emphasizing narrative communication (e.g., who

did what to whom), gossip, interpersonal relationships, social hierarchies, and ethnocentrism. The symbolic form of allocentric miming (BPO-A) would have expanded the representational breadth of pantomiming beyond depicting people, as well as expanded the richness of spatial depiction beyond peripersonal actions. It would be able to represent complex scenes through spatial reduction, which might have aided in group planning for hunting, foraging, and warfare. Overall, the People First model emphasizes both social cognition and tool use, whereas the Environment First model emphasizes spatial cognition, permitting a wide diversity of objects to be scenically represented in large and complex spaces. While we cannot prioritize one model over the other evolutionarily, we do argue that the more iconic pantomime types within each model preceded the more symbolic pantomime types, in part because both of the symbolic types require body replacement. We also argue that, for BPO's themselves, BPO-E's may have preceded BPO-A's because of the former's connection with peripersonal pantomimes. The use of pantomime as a first stage in language evolution may have been underlain by a "platform of trust" for honest communication between mimer and recipient (Dor et al., 2014; Wacewicz and Zywiczynski, 2018), not least since pantomime employs invisible objects, making deception possible.

We would like to point out a difference between egocentric and allocentric miming when it comes to the cognitive encoding of time. The basic idea is that allocentric miming is specialized for depicting the present, while egocentric miming is specialized for depicting the non-present, hence requiring mental time travel. When someone pantomimes a human action egocentrically through an IO, they are not saying "I am doing this action right now," since a pantomime is a substitute for the real action. Egocentric pantomimes essentially *exclude* the present tense. Hence, the mimer could either be conveying "I did this action in the past" (narrative) or "I plan to do this action in the future" (planning, simulation). By contrast, BPO's are more of a description of the way things currently are, e.g., the shapes, sizes, and/or locations of objects. Hence, they can convey the present tense in a way that egocentric pantomimes cannot. While they can certainly depict the past and future as well, the default mode of interpretation of an allocentric pantomime is the present tense. In contrast, an egocentric pantomime has to be any tense other than the present tense, although the proposed role of pantomime in demonstration (Gärdenfors, 2017) might be the closest thing to a present-tense usage. This might suggest a trade-off between egocentric and allocentric miming with respect to the contrast between object transparency and time transparency: egocentric pantomimes are more transparent with respect to identity but require mental time travel, whereas BPO's require object substitution but offer temporal transparency by defaulting to the present tense.

As a final thought, one area where the People First and Environment First models might effectively come together is seen in the "confrontational scavenging" model of Bickerton and Szathmáry (2011). This model argues that scavenging provided a natural-selection pressure for the evolution of both language and cooperation in hominins, since scavenging requires the recruitment of group members in order to both perform the scavenging and bring the food materials back to the group. "Band members who had located a carcass would have had to use sounds, gestures or mimicry to inform potential recruits of what they had found" (p. 4). From our perspective, allocentric mimes might have been used to iconically convey information about the nature of the carcass and its distant location, while egocentric mimes might have been used to communicate information about the recruitment of group members by miming the tool-use actions that would be required for the scavenging. Hence, the confrontational scavenging model of ancestral communication unites the environment and people into a single suite of cooperative survival behaviors. This would combine temporal information about the present (the carcass and its location) with that about the future (recruitment for planned scavenging).

CONCLUSIONS

We have presented a new classification scheme for pantomime that adds two new categories, in addition to tracing, to the standard scheme of IO and BPO. The categories vary with regard to whether they occur in peripersonal space or extrapersonal space. They also vary with regard to whether they employ body-part use or body-part replacement. Using the scheme, we have attempted to create associations between the type of pantomime used and the type of action or object being depicted. We have applied this reasoning to theorizing about the evolutionary origins of language in order to propose a semantic-content-based distinction between a People First model based on egocentric pantomimes about human actions, and an Environment First model based on allocentric pantomimes about non-human objects and their landscapes. Future research can capitalize on this distinction to further explore models of language origins, as based both on gesture and vocalization.

AUTHOR CONTRIBUTIONS

All authors contributed to the conceptualization of the ideas of the paper, to the classification scheme for pantomime, and to the writing of the manuscript.

FUNDING

ACKNOWLEDGMENTS

REFERENCES

Abramova, E. (2018). The role of pantomime in gestural language evolution, its cognitive bases and an alternative. *Lang. Evol.* 3, 1–15. doi: 10.1093/jole/lzx021

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

Arbib, M. A. (2018). In support of the role of pantomime in language evolution. *Lang. Evol.* 3, 41–44. doi: 10.1093/jole/lzx023

Armstrong, D. F., and Wilcox, S. E. (2007). *The Gestural Origins of Language.* Oxford: Oxford University Press.

Bangerter, A. (2004). Using pointing and describing to achieve joint focus of attention in dialogue. *Psychol. Sci.* 15, 415–419. doi: 10.1111/j.0956-7976.2004.00694.x

Beattie, G. (2016). *Rethinking Body Language: How Hand Movements Reveal Hidden Thoughts.* London: Routledge.

Beattie, G., and Shovelton, H. (2002). An experimental investigation of some properties of individual iconic gestures that mediate their communicative power. *Br. J. Psychol.* 93, 179–192. doi: 10.1348/000712602162526

Bickerton, D., and Szathmáry, E. (2011). Confrontational scavenging as a possible source for language and cooperation. *BMC Evol. Biol.* 11:261. doi: 10.1186/1471-2148-11-261

Boyatzis, C. J., and Watson, M. W. (1993). Preschool children's symbolic representation of objects through gestures. *Child Dev.* 64, 729–735. doi: 10.2307/1131214

Brinck, I. (2004). The pragmatics of imperative and declarative pointing. *Cogn. Sci. Q.* 3, 429–446.

Brown, S. (2017). Proto-acting as a new concept: personal mimicry and the origins of role playing. *Humanities* 6:43. doi: 10.3390/h6020043

Carlson, S. M., and Taylor, M. (2005). Imaginary companions and impersonated characters: sex differences in children's fantasy play. *Merrill Palmer Q.* 51, 93–118. doi: 10.1353/mpq.2005.0003

Cartmill, E. A., Beilock, S., and Goldin-Meadow, S. (2012). A word in the hand: action, gesture and mental representation in humans and non-human primates. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 367, 129–143. doi: 10.1098/rstb.2011.0162

Casasanto, D., and Jasmin, K. (2012). The hands of time: temporal gesture in English speakers. *Cogn. Linguist.* 23, 643–674. doi: 10.1515/cog-2012-0020

Cassell, J., and McNeill, D. (1990). "Gesture and ground," in *Proceedings of the Sixteenth Annual Meeting of the Berkeley Linguistics,* eds K. Hall, J.-P. Koenig, M. Meacham, S. Reinman, and L. A. Sutton (Berkeley, CA: Berkeley Linguistics Society), 57–68.

Clark, H. H. (1996). Depicting as a method of communication. *Psychol. Rev.* 123, 324–347. doi: 10.1037/rev0000026

Cochet, H., and Vauclair, J. (2010). Pointing gesture in young children: hand preference and language development. *Gesture* 10, 129–149. doi: 10.1075/gest.10.2-3.02coc

Colonnesi, C., Stams, G. J. J. M., Koster, I., and Noom, M. J. (2010). The relation between pointing and language development: a meta-analysis. *Dev. Rev.* 30, 352–366. doi: 10.1016/j.dr.2010.10.001

Corballis, M. C. (2002). *From Hand to Mouth: The Origins of Language.* Princeton, NJ: Princeton University Press.

Corballis, M. C. (2009). The evolution of language. *Ann. N. Y. Acad. Sci.* 1156: 19–43. doi: 10.1111/j.1749-6632.2009.04423.x

Deacon, T. W. (1997). *The Symbolic Species: The Co-Evolution of Language and the Brain.* New York, NY: Norton.

Dick, A. S., Overton, W. F., and Kovacs, S. L. (2005). The development of symbolic coordination: representation of imagined objects, executive function, and theory of mind. *J. Cogn. Dev.* 6, 133–161. doi: 10.1207/s15327647jcd0601_8

Donald, M. (1991). *Origins of the Modern Mind: Three Stages in the Evolution of Culture and Cognition.* Cambridge, MA: Harvard University Press.

Donald, M. (1993). Précis of origins of the modern mind: three stages in the evolution of culture and cognition. *Behav. Brain Sci.* 16, 737–791. doi: 10.1017/S0140525X00032647

Donald, M. (2013). "Mimesis theory re-examined, twenty years after the fact," in *Evolution of Mind, Brain, and Culture,* eds G. Hatfield, and H. Pittman (Philadelphia, PA: University of Pennsylvania Press), 169–192.

Dor, D., Knight, C., and Lewis, J. (2014). "Introduction: a social perspective on how language began," in *The Social Origins of Language,* eds. D. Dor, C. Knight, and J. Lewis (Oxford: Oxford University Press), 1–14.

Dunbar, R. I. M. (1996). *Grooming, Gossip and the Evolution of Language.* London: Faber and Faber.

Ekman, P., and Friesen, W. V. (1969). The repertoire of non-verbal behavior: categories, origins, usage, and coding. *Semiotica* 1, 49–98. doi: 10.1515/semi.1969.1.1.49

Ferretti, F., Adornetti, I., Chiera, A., Nicchiarelli, S., Magni, R., Valeri, G., et al. (2017). Mental time travel and language evolution: a narrative account of the origins of human communication. *Lang. Sci.* 63, 105–118. doi: 10.1016/j.langsci.2017.01.002

Gärdenfors, P. (2017). Demonstration and pantomime in the evolution of teaching. *Front. Psychol.* 8:415. doi: 10.3389/fpsyg.2017.00415

Gizzonio, V., Avanzini, P., Campi, C., Orivoli, S., Piccolo, B., Cantalupo, G., et al. (2015). Failure in pantomime action execution correlates with the severity of social behavior deficits in children with autism: a praxis study. *J. Autism Dev. Disord.* 45, 3085–3097. doi: 10.1007/s10803-015-2461-2

Gramann, K. (2013). Embodiment of spatial reference frames and individual differences in reference frame proclivity. *Spat. Cogn. Comput.* 13, 1–25. doi: 10.1080/13875868.2011.589038

Hall, E. (2009). "Introduction: pantomime, a lost chord of ancient culture," in *New Directions in Ancient Pantomime,* eds E. Hall, and R. Wyles (Oxford: Oxford University Press), 1–40.

Harnad, S. (1990). The symbol grounding problem. *Physica D* 42, 335–346. doi: 10.1016/0167-2789(90)90087-6

Haviland, J. (2000). "Pointing, gesture spaces, and mental maps," in *Language and Gesture,* ed D. McNeill (Cambridge: Cambridge University Press), 13–46.

Heilman, K. M., Rothi, L. J., and Valenstein, E. (1982). Two forms of ideomotor apraxia. *Neurology* 32, 342–346. doi: 10.1212/WNL.32.4.342

Hewes, G. W. (1973). Primate communication and the gestural origins of language. *Curr. Anthropol.* 14, 5–24. doi: 10.1086/201401

Hockett, C. (1960). The origin of speech. *Sci. Am.* 203. 88–111. doi: 10.1038/scientificamerican0960-88

Hoeren, M., Kümmerer, D., Bormann, T., Beume, L., Ludwig, V. M., Vry, M. S., et al. (2014). Neural bases of imitation and pantomime in acute stroke patients: distinct streams for praxis. *Brain* 137, 2796–2810. doi: 10.1093/brain/awu203

Jackendoff, R. (1999). Possible stages in the evolution of the language capacity. *Trends Cogn. Sci.* 3, 272–279. doi: 10.1016/S1364-6613(99)01333-9

Jackendoff, R. (2002). *Foundations of Language: Brain, Meaning, Grammar, Evolution.* Oxford: Oxford University Press.

Kendon, A. (1988). "How gestures can become like words," in *Cross-Cultural Perspectives on Non-Verbal Communication,* ed F. Poyatos (Toronto, ON: Hogrefe), 131–141.

Kendon, A. (2017). Reflections on the "gesture-first" hypothesis of language origins. *Psychon. Bull. Rev.* 24, 163–170. doi: 10.3758/s13423-016-1117-3

Klatzky, R. (1998). "Allocentric and egocentric spatial representations: definitions, distinctions, and interconnections," in *Spatial Cognition,* eds C. Freksa, C. Habel, and K. F. Wender (Heidelberg: Springer), 1–17.

Langacker, R. W. (1986). An introduction to cognitive grammar. *Cogn. Sci.* 10, 1–40. doi: 10.1207/s15516709cog1001_1

Legare, C. H., and Nielsen, M. (2015). Imitation and innovation: the dual engines of cultural learning. *Trends Cogn. Sci.* 19, 689–699. doi: 10.1016/j.tics.2015.08.005

Leslie, A. M. (1987). Pretense and representation: origins of "theory of mind". *Psychol. Rev.* 94, 412–426.

Lillard, A., Pinkham, A. M., and Smith, E. (2014). "Pretend play and cognitive development," in *Wiley-Blackwell Handbook of Childhood Cognitive Development, 2nd Edn.,* ed U. Goswami (West Sussex: John Wiley & Sons), 285–311.

MacNeilage, P. F. (1998). The frame/content theory of evolution of speech production. *Behav. Brain Sci.* 21, 499–546. doi: 10.1017/S0140525X98001265

MacNeilage, P. F., and Davis, B. L. (2005). The frame/content theory of evolution of speech: a comparison with a gestural-origins alternative. *Interact. Stud.* 6, 173–199. doi: 10.1075/is.6.2.03mac

McGinn, C. (2015). *Prehension: The Hand and the Emergence of Humanity.* Cambridge, MA: MIT Press.

McNeill, D. (1992). *Hand and Mind: What Gestures Reveal About Thought.* Chicago, IL: University of Chicago Press.

McNeill, D. (2005). *Gesture and Thought.* Chicago, IL: University of Chicago Press.

McNeill, D. (2012). *How Language Began: Gesture and Speech in Human Evolution.* Cambridge, UK: Cambridge University Press.

Mellinger, T., and Vosgerau, G. (2010). "Putting egocentric and allocentric into perspective," in *Spatial Cognition VII*, eds C. Hölscher, T. F. Shipley, M. O. Belardinelli, J. A. Bateman, and N. S. Newcombe (Berlin: Springer-Verlag), 207–221.

O'Reilly, A. W. (1995). Using representations: comprehension and production of actions with imagined objects. *Child Dev.* 44, 309–314.

Orzechowski, S., Sławomir, W., and Zywiczynski, P. (2014). "Orofacial gestures in language evolution. The auditory feedback hypothesis," in *The Evolution of Language. Proceedings of the 10th International Conference (EVOLANG 10)*, eds E. Cartmill, S. Roberts, H. Lyn, and H. Cornish (Singapore: World Scientific), 221–227.

Parrill, F. (2010). Viewpoint in speech-gesture integration: linguistic structure, discourse structure, and event structure. *Lang. Cogn. Process.* 25, 650–668. doi: 10.1080/01690960903424248

Peirce, C. S. (1903/1987). "Logic as semiotic: the theory of signs," in *The Philosophical Writings of Peirce*, ed J. Buchler (New York, NY: Dover), 98–119.

Perlman, M., and Gibbs, R. W. Jr. (2013). Pantomimic gestures reveal the sensorimotor imagery of a human-fostered gorilla. *J. Mental Imagery* 37, 73–96.

Perniss, P., and Vigliocco, G. (2014). The bridge of iconicity: from a world of experience to the experience of language. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 369:20130300. doi: 10.1098/rstb.2013.0300

Petkov, C. I., and Jarvis, E. D. (2012). Birds, primates, and spoken language origins: behavioral phenotypes and neurobiological substrates. *Front. Evol. Neurosci.* 4:12. doi: 10.3389/fnevo.2012.00012

Rendell, L., Fogarty, L., Hoppitt, W. J. E., Morgan, T. J. H., Webster, M. M., and Laland, K. N. (2010). Cognitive culture: theoretical and empirical insights into social learning strategies. *Trends Cogn. Sci.* 15, 68–76. doi: 10.1016/j.tics.2010.12.002

Richerson, P., Baldini, R., Bell, A. V., Demps, K., Frost, K., Hillis, V., et al. (2016). Cultural group selection plays an essential role in explaining human cooperation: a sketch of the evidence. *Behav. Brain. Sci.* 39:e30. doi: 10.1017/S0140525X1400106X

Rogers, S. J., Mcevoy, R., and Pennington, B. F. (1996). Imitation and pantomime in high-functioning adolescents with autism spectrum disorders. *Child Dev.* 67, 2060–2073. doi: 10.2307/1131609

Rose, M. L., Mok, Z., and Sekine, K. (2017). Communicative effectiveness of pantomime gesture in people with aphasia. *Int. J. Lang. Commun. Disord.* 52, 227–237. doi: 10.1111/1460-6984.12268

Seyfarth, R. M., Cheney, D. L., and Bergman, T. J. (2005). Primate social cognition and the origins of language. *Trends Cogn. Sci.* 9, 264–266. doi: 10.1016/j.tics.2005.04.001

Smith, I. M., and Bryson, S. E. (2007). Gesture imitation in autism: II. symbolic gestures and pantomimed object use. *Cogn. Neuropsychol.* 24, 679–700. doi: 10.1080/02643290701669703

Suddendorf, T., Fletcher-Flinn, C., and Johnston, L. (1999). Pantomime and theory of mind. *J. Gen. Psychol.* 160, 31–45. doi: 10.1080/00221329909595378

Tallerman, M. (2015). *Understanding Syntax. 4th Edn.* London: Routledge.

Taub, S. F. (2001). *Iconicity and Metaphor in American Sign Language.* Cambridge: Cambridge University Press.

Tomasello, M. (2008). *Origins of Human Communication.* Cambridge, MA: MIT Press.

van Nispen, K., Mieke, W. M. E., van de Sandt-Koenderman, E., and Krahmer, E. (2018). The comprehensibility of pantomimes produced by people with aphasia. *Int. J. Lang. Commun. Disord.* 53, 85–100. doi: 10.1111/1460-6984.12328

Wacewicz, S., and Zywiczynski, P. (2018). Language origins: fitness consequences, platform of trust, cooperation, and turn-taking. *Interact. Stud.* 19, 167–182. doi: 10.1075/is.17031.wac

Walton, K. L. (1990). *Mimesis as Make-Believe: On the Foundations of the Representational Arts.* Cambridge, MA: Harvard University Press.

Weisberg, D. S. (2015). Pretend play. *Wiley Interdiscip. Rev. Cogn. Sci.* 6, 249–261. doi: 10.1002/wcs.1341

Wulf, G., and Mornell, A. (2008). Insights about practice from the perspective of motor learning: a review. *Music Performance Research* 2, 1–25.

Yuan, Y., and Brown, S. (2014). The neural basis of mark making: a functional MRI study of drawing. *PLoS ONE* 9:e108628. doi: 10.1371/journal.pone.0108628

Zlatev, J. (2014). Human uniqueness, bodily mimesis and the evolution of language. *Humana Mente* 27, 197–219.

Zlatev, J. (2018). Different kinds of parsimony: association-learning versus bodily mimesis. *J. Lang. Evol.* 3, 45–46. doi: 10.1093/jole/lzy003

Zlatev, J., Wacewicz, S., Zywiczynski, P., and van de Weijer, J. (2017). Multimodal-first or pantomime-first? *Interact. Stud.* 18, 452–475. doi: 10.1075/is.18.3.08zla

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

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

APPENDIX: EXAMPLES OF THE 5 CATEGORIES OF PANTOMIMING

FIGURE A1 | The model in these photos provided written informed consent for the publication of these images.