From action to spoken and signed language through gesture

Some basic issues for a discussion on the evolution of the human language-ready brain

Virginia Volterra, Olga Capirci, Pasquale Rinaldi and Laura Sparaci
Institute of Cognitive Sciences and Technologies (ISTC), National Research Council (CNR) of Italy

~~We~~ review major developmental evidence on the continuity from action to gesture to word and sign in human children, highlighting the important role of caregivers in the development of multimodal communication. In particular, the basic issues considered here and contributing to the current debate on the origins and development of the language-ready brain are: (1) links between early actions, gestures and words and similarities in representational strategies; (2) importance of multimodal communication and the interplay between gestures and spoken words; (3) interconnections between early actions, gestures and signs. The innovation of this report is in ~~the bringing~~ these themes together ~~with~~ relevant findings from studies on children between 6 and 36 months of age and highlighting interesting parallels in studies on ape communicative behavior.

Keywords: child development, communication, multimodality, actions, gestures, words, signs

Introduction

~~We~~ review major developmental evidence on the continuity from action to gesture to both word and sign, reporting how their interrelationship extends beyond early childhood and across cultures. Our aim is to show the importance of pure motor acts (e.g., grasping) in the development of symbolic communication. We propose that ~~the clarifying of~~ mechanisms involved in language acquisition (spoken, or signed) and ~~the~~ overcoming ~~of~~ difficulties stemming from differing definitions of gestures ~~could~~ contribute to a more careful comparison of human vs. non-human

Interaction Studies 19:1-2 (2018), **216–238**. DOI 10.1075/is.17027.vol
ISSN 1572–0373 / E-ISSN 1572–0381

primate communication. ~~This terminology~~ is not only heterogeneous, but has often changed considerably over the years reflecting parallel changes in methodology, perspectives and questions addressed.

To gain better understanding of the role of gestures in the progression from actions to words, as well as their role in the evolution of the language-ready brain, we will present evidence from gesture studies conducted in different periods, with different methodologies and influenced by diverse theoretical perspectives. For example, in 1975 we observed how in young children specific goal-oriented actions (e.g., orienting, reaching, grasping), would gradually become separated from their concrete goals (i.e., the attempt to reach a specific object), and assume symbolic functions (Bates, Camaioni & Volterra, 1975). Many years later, novel neurophysiological findings on the link between actions and language and the theory of the mirror neuron system (MNS) led to new understanding of these data, grounding behavioral observations in neurophysiological evidence (Arbib & Rizzolatti, 1997) and leading us to rename those communicative acts 'action/gestures' in order to highlight their links with action schemas. These gestures ~~that mark~~ the emergence of symbolic communication together with the production of the first words, continue ~~their~~ interrelationship ~~with words~~ by giving the child the opportunity to combine elements (from one- to two-words stage) ~~and extending~~ beyond childhood ~~until~~ the gesture-speech integrat~~ing~~ system is consolidate in adulthood (McNeill, 1992). ~~These~~ findings were compatible with Corballis' theory (2002) on the gestural origin of language, proposing that gesture and speech have co-evolved in complex interrelationships throughout their long and changing partnership.

Moreover, we noted striking similarities between early gestures of hearing children learning to speak and early signs of deaf children learning to sign (Caselli & Volterra, 1994). We will show ~~that the~~ representational strategies ~~that we encounter~~ in everyday gestures and that we have studied in SLs are not only the same, but also grounded in basic embodied motor acts that we acquire in childhood. What emerges is not a clear-cut separation, but a continuity between co-speech gestures produced by hearing children and early signs produced by children exposed to a SL. For a schematic outline, see Supplemental Material #1.

The model of "language" ~~we think is most helpful consider that~~ human communication transcends the spoken medium, often exploiting embodied forms such as signs and gestures within a multimodal approach. In an elegant and provocative paper Slobin (2008) suggested and clarified that we cannot assume to uncover in ~~sign languages~~ the same linguistic categories and processes we find in some spoken languages (i.e., written English). Several studies o~~f~~ different ~~sign languages~~ have begun to consider signs as visible actions or dedicated gestures with linguistic properties (Cuxac, 2000; Liddell & Metzger, 1998). Studying the visible actions of speakers and signers leads to a revision of the traditional dichotomy

between *linguistic* and *enacted* (assumed to be "non-linguistic") and to the development of a new approach to embodied language (Kendon, 2014).

In the present paper we will describe different types of gestures and the diversity of mechanisms potentially involved, suggesting that some aspects of ontogeny might suggest hypotheses for the evolution of the Road Map, in particular ~~concerning the relation~~ between praxic actions and spoken and signed language acquisition. There is increasing evidence that all gestures ~~do not~~ share the same developmental process. To acknowledge that this diversity of mechanisms might also be involved in the ontogeny of nonhuman primate gestural communication might present an important ~~but~~ fruitful challenge for comparative neuroprimatology.

1. From action to gesture and word

1.1 Links between early motor skills and gestures

Traditional studies on the onset of intentional communication in humans highlighted the importance of considering communication as essentially multimodal by stressing the role of 'performatives' prior to speech (Bates et al., 1975). The term 'performatives' refers essentially to actions emerging prior to speech, often accompanied by vocalizations and used by young infants to signal different types of basic intentional states. Examples of performatives included: ritualized request, showing off, showing, giving and pointing (see Figure 1).

'Ritualized requests' are acts performed by the child to induce specific responses in caregivers to obtain specific objects or actions (e.g., at 8 months children often raise their arms to be picked up or extend ~~their~~ arm towards an object while opening and closing ~~the~~ hand to indicate that they wish to grasp the object ~~itself~~). Actions termed 'showing off' are funny or unexpected acts that induce laughter or positive reactions in caregivers and may initiate or maintain social interactions (e.g., around 9 months a child may blow raspberries, inducing laughter in caregivers). Showing acts are usually meant to display presence of an object/event or specific characteristics of an object/event to caregivers (e.g., by 10 months the child is able to extends her arms toward a caregiver while holding a toy and to open her hand to show it). Giving includes actions that imply the passage of an object from child to caregiver (e.g., by 13 months a child may pick up an object on the other side of the room, cross the room and drop it in the caregiver's lap). Pointing includes deictic actions that lead others' attention towards specific objects/events immediately present in the surrounding environment (e.g., between 12 and 13 months a child may point to a cat while looking at the adult) (Bates et al., 1975).

All performatives originate from basic motor actions that are already present in the toddler's motor repertoire. They are mainly attention-grabbers that

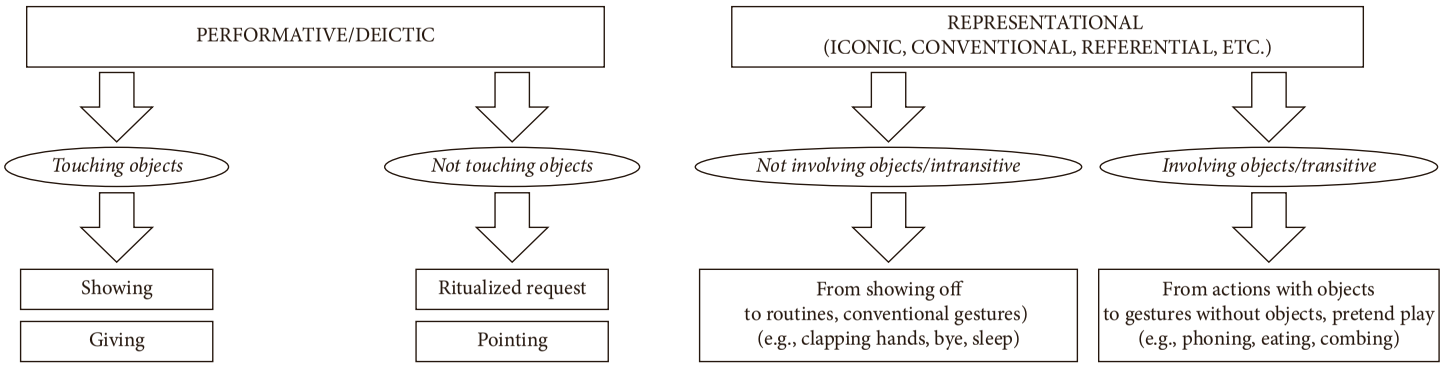


Figure 1. Types of gestures

gradually emerge during development ~~for attracting~~ caregiver's attention towards the child himself or towards objects/events. These gestures are often accompanied by vocalizations, which enhance their main scope as attention-grabbers, providing means of directing caregivers' attention toward things and events present in the environment, while also building relevant shared experiences. The primary cognitive prerequisite for performative intentions reported in these studies ~~to be Piaget's~~ sensorimotor stage 5 (Piaget, 1945)~~, in particular the ability of tool use~~. In other words, when they first appear on the scene of human communication, performatives don't necessarily require reference to intentional states; rather they constitute cases in which familiar actions performed in specific contexts attract the other's attention and induce specific reactions (Zukow-Goldring, 2012). These actions, through repetition, may lead to ontogenetic ritualization, a process of mutual anticipation in which particular social behaviors become ritualized to function as intentional communicative signals (Arbib, Ganesh, & Gasser, 2014).

Whole-body actions (e.g., lifting the arms to be picked up), initially induced by specific constants in the surrounding environment (i.e., the adult is usually in a higher position with respect to the child), may be repeated up to the point that they become ritualized requests produced to elicit an adult's response. Interestingly these types of performatives, which are already present in human infants by 8 months, have also been analyzed in apes. In particular, emergence of gestures implying ritualized request behaviors in bonobo mother-infant dyads preceding acts of carrying, highlight~~ing~~ relevant similarities with human child behaviors (Halina, Rossano and Tomasello, 2013). However, ontogenetic ritualization in human infants as well as in non-human primates is a highly debated topic (Liebal, 2016; Marentette & Nicoladis, 2012; Tomasello, 2008).

~~P~~erformative behaviors such as showing off are linked to motor behaviors ~~which involve~~ other parts ~~of the child's body:~~ torso, head, face, arms, hands or legs~~.~~ ~~O~~ther performative forms (i.e., request, show, give, and point) may be traced back to infants' early motor exploration of objects through different types of grasping. The development of this ability is linked to advances in postural control (e.g., independent sitting) and in reaching skills (for an extensive review see Sparaci & Volterra, 2017). Early performative signals, can also be found among other species (Gretscher et al., 2017 for a review). Similarly, the emergence of pointing may be traced back to early fine-motor actions (for an extensive review of pointing skills see Kita, 2003). While pointing is commonly described as a gesture in which the index finger and arm are extended in the direction of ~~the~~ interesting object, many other forms of pointing exist, such as pointing performed using other body parts (e.g. head and/or eye movements, lip-protruding, etc.) (Enfield, 2001). In this broader sense, pointing gestures have been identified by some authors also

in non-human apes in special circumstances (Leavens, 2004), ~~while others have argued against it~~ (Tomasello, 2006).

Most performative gestures described so far have later been called "deictic" as they express only a communicative intent ~~while~~ their content can only be interpreted referring to the extralinguistic context. Other performatives (e.g., "raising arms" to be picked up) have also been described in the past as "referential gestures," as they gradually denote a semantic content which may remain relatively stable across different contexts. What about representational gestures in which the task of representing an object or an action with an object is mainly carried out by the hands? A viable hypothesis is that these gestures are grounded in early actions with objects. In fact, various studies have observed that at the same age at which infants start mastering precision grips (around 12-months), and just before spoken naming onset, short action sequences with objects begin to emerge (Caselli, 1990). These action sequences, are usually related to objects' functions (e.g., using a spoon to eat). Soon afterwards, these same action sequences may be performed in absence of objects, while maintaining their meaning (e.g., the child could place an empty spoon in his/her mouth as if eating and subsequently reproduce the same handshape and movement used in eating with a spoon with an empty hand) (Capirci, Contaldo, Caselli & Volterra, 2005). In this sense, motor schemas and handshapes exploited by infants in grasping and functional acts may be linked to representational gestures, performed in absence of an object and denoting a specific referent while remaining relatively stable across different contexts (Sparaci & Volterra, 2017). Summing up, considering deictic and representational gestures allows us to trace back the origins of these forms to early ~~both~~ fine- and gross- motor skills exercised by infant-caregiver dyads within specific contexts.

1.2 Early action and gesture "Vocabulary" and its relationship to word comprehension and production

The early studies described above had identified different types of gestures produced by very young children. In a further longitudinal study on three children followed between 10 and 23 months of age, Capirci and colleagues (2005) found that the 'practiced meanings' that infants initially exercise in communicative actions with caregivers are likely to enter their communicative repertoires as representational (i.e., empty-handed) gestures and/or words. These longitudinal findings have been confirmed in a cross-sectional study on a broader sample of 492 Italian infants between 8 and 18 months of age using the Italian MB-CDI assessment tool (Caselli, Rinaldi, Stefanini & Volterra, 2012). The production of action-gestures was strongly correlated with word comprehension, probably because meanings of these gestures are shared with caregivers who often produce the related word

before or after the child's gesture production, reinforcing the link between action/gestures (A/G) and words. The order of appearance of early A/G has also been analyzed considering gestures' motor execution, outlining a developmental pattern for different gesture types and introducing the distinction between A/G involving or not object manipulation (for age of A/G appearance see Supplemental Material #2). This distinction was introduced to underline different origins, times of emergence and further outcomes of these gestures. Gestures not involving objects derive from what we called "showing off behaviors" performed very early in dyadic child-caregiver interactions (Reddy, 2003), while gestures involving objects derive from object manipulation behaviors (the traditional schemes with objects described by Piaget, 1945). Age of production of action/gesture and comprehension/production of words with related meaning have been analyzed in Caselli et al., (2012) (see Supplemental Material #3).

The emergence of intentional and symbolic communication relies on both ~~vocal and gestural~~ modalities ~~but also~~ speech and gesture interact in the development of the ability to convey two pieces of information within a single communicative utterance. Various studies (Capirci et al., 2005; Capirci, Iverson, Pizzuto & Volterra, 1996; Capobianco et al., 2017; Goldin-Meadow & Butcher, 2003) were able to show that cross-modal utterances precede in all children the emergence of two-word utterances (e.g., a child may point towards a chair while saying "mommy" to ask her mother to sit with her; or a child may perform an eating gesture and then point towards a specific food on the table in order to request being fed that food). At this stage a large portion of children's prototypical nomination and predication structures are expressed via two-element cross-modal utterances. In particular Capirci et al. (1996, p. 663–664) showed how children's supplementary combinations expressed a variety of basic semantic relations similar to those expressed by children ~~at~~ the early two-word stage as shown in Supplemental Material #4.

Further support for the fact that early communicative gestures may be used as a form of naming comes from recent studies conducted on ~~two-three year old~~ children, showing that when children expand and consolidate their spoken vocabulary, gesture production, far from declining, continues to accompany spoken words. Our research group carried out several studies on this topic using a new task assessing vocabulary comprehension and production (Bello, Giannantoni, Pettenati, Stefanini & Caselli, 2012), which showed that children, when requested to produce spoken word labels for pictures, also performed pointing and representational gestures (Stefanini, Bello, Caselli, Iverson, & Volterra, 2009).

It is striking that children who are already able to name a picture in speech still often resort to gestures. Representational gestures produced were stylized and conventionalized versions of manipulative actions (e.g., bringing a hand to the

head as if combing, holding the bar of a playground merry-go-round as if spinning it). Evidence of a link between gestures and actions performed directly on objects are indicated by examples in which children executed the gesture while holding the picture (e.g., combing themselves with the picture of a comb, spinning the picture of a playground merry-go-round). This suggests that words may not be fully decontextualized yet, and that gesture production may allow recreating the action as well as the motor context in which the word was initially acquired. Also, the need to point to the referent depicted in the picture may be understood as an attempt to participate in a communicative interaction based on a joint attention scheme. Similar results on the production of pointing and representational gestures in performing the same naming task have been found in Japanese and Canadian, children. Results from these studies confirm that motor representations may be needed to support linguistic representations in speech, irrespective of the cultural environment in which the child is raised, even though the rate and the way gestures are produced may be influenced by the culture from even early stages in development (Marentette et al., 2016). Analyses of spontaneous gesture production in a naming task also provide empirical support for the idea that gestures and speech share a common conceptual space as well as an activation of hand-mouth motor programs associated with specific objects or actions. The symbolic strategies adopted are also indications of a strong continuity from actions to both spoken and signed language (Pettenati et al., 2012).

2. Representational techniques across elicited pantomime in children, communicative gestures and sign languages

In recent research, considerable efforts have been made in attempting to consider together representational techniques in the manual modality as described by different traditions of studies, for example elicited pantomimes in symbolic development studies, co-speech and silent gesture in child and adult spontaneous communication, and signs in Sign Language research.

Crucial for any attempt to develop a unified taxonomy is to clarify whether during gesture execution the body/hands represent real actions in the physical world (i.e., how an action is performed or how an object is held/used) or something other (e.g., the object itself or its size/shape). Table 1 shows some of the different labels used in different ~~kinds of~~ studies for four main representational strategies.

For example, technique (a) that involves the person's own-body can depict movements of an agent in which case the gesturing body parts engage in a pattern of action that has many kinematic features in common with the action that is being referred to. The origin of this strategy can be traced back to early performatives

Table 1. Labels used for representational techniques or strategies in different traditions

	In symbolic development studies	In gesture studies on adults and children	In sign language studies	Examples
a		**own-body** enactment mime/pantomime character viewpoint action gesture	constructed action body classifier person transfer	– *swimming:* paddling the arms/hands in the air – *turning:* rotating the torso/swiveling
b	imagined-object	**hand-as-hand** handling manipulation action gesture character viewpoint function gesture	constructed action handling classifier person transfer	– *fork:* bringing a closed fist towards the mouth – *combing:* fingers wrapped around an imaginary handle of a comb
c	body-part-as-object	**hand-as-object** observer viewpoint modeling action gesture form gesture	entity/instrument classifier situation transfer	– *umbrella:* bringing an open hand above the head – *fork:* extended fingers as tines, move from table to mouth
d		**size-and-shape** depiction delimitation observer viewpoint	size /shape specifiers tracing form transfer	– *small:* finger and thumb held close together – *tree:* tracing the form of the tree in the air

involving the whole body (e.g., showing off) described above. In a similar way, in technique b, when actions can be represented using only the hands rather than the entire body, hands can be used to enact how an object is held or manipulated or to show motor acts or grasps associated to object use. The origin of this representational strategy (also termed 'hand-as-hand') can be traced back to functional actions with objects (e.g., using a fork to eat). Studies on symbolic development have called these gesture types "imagined-object gestures". Both techniques 1 and 2, involving ones own body or only the ones own hand-as-hand, rely on representational techniques involving some form of enactment. However, our hands can also be used to represent an object itself, as in technique c. The origin of this representational strategy can be traced back to actions and gestures involving objects during early stages of development.

Finally, with technique d, the hand can represent the size-and-shape of an object (i.e., tracing the contours of an object or the outcome path of a movement). This representational strategy should be kept distinct from the hand-as-object

strategy. In fact, in this case the hand is used to describe the object, but does not physically stand-for the object itself. The origin of this representational technique may be found in A/G, irrespective of whether they involve objects or not. As shown in Table 1, these four techniques have been described by different disciplines using different labels.

Results from elicited pantomime studies reported that three-year-olds, requested to label an object such as "comb" and asked to pretend to use that item, were more likely to produce hand-as-object gestures, while children older than 6 years produced also imagined-objects gestures in which they pretended to hold and use an object, depicting its function (e.g., fingers wrapped around an imaginary handle of a comb) Boyatzis & Watson, 1993).

Iverson, Capirci and Caselli (1994) provided the first detailed analysis of representational techniques present in gestures produced by 16- and 20-month-old Italian children. This study examined children's use of communicative gesture showing how they relied on different strategies depicting: (a) characteristics or qualities of an object (e.g., holding a hand high over the head to indicate big, similar to the size-and-shape technique); (b) actions with an object (e.g., combing with or without a comb in the hand to indicate "comb" or "combing", similar to the hand-as-hand technique); (c) the form or movement of the object itself (e.g., flapping hands and arms for "bird", similar to the own-body technique). Iverson et al. (1994) found that the proportion of gestures reproducing actions with objects (e.g., the function of the object) or the form or movement of the object itself, tended to increase between 16- and 20-months-of-age.

More recent studies have analyzed gesture production in older children and in different tasks (naming and narrative tasks). Marentette and colleagues (2016) considered cross-cultural differences comparing spoken and gestural productions of 2-year-old children in a picture naming task, showing that Italian children growing up in a 'high gesture culture' produced twice as many gestures as Canadian children from a 'low gesture culture', but gestures from both groups involve a similar range of the four main representational techniques discussed above. In fact, two-year-olds were equally likely to produce gestures depicting function (technique b) as form (technique c). Despite cultural and linguistic differences in frequency of use of the individual techniques, strategies for depicting information about objects and events make visible different types of embodied practices and suggest a shared cognitive basis, which is recruited by both language and gestural systems.

Finally, Capirci, Cristilli, de Angelis, and Graziano (2011) in examining the development of co-speech gestures in four- to ten-year-olds' narratives, found that the most widely used strategy for all age groups was hand-as-hand, but older children produced more hand-as-object with respect to younger ones. These authors

suggest that co-speech gestures become more abstract as their form becomes representationally more flexible and that children aim at depicting specific aspects of objects, rather than using their hands as if acting in the physical world.

These results by showing that use of hand-as-object representational technique increases with age growth are in clear opposition to data collected using elicited pantomimes and described above as showing that only older children tended to use the hand-as-hand strategy. Possibly these contrasting data are due to the small number of items used in tasks ~~for~~ eliciting pantomime. In fact, as reported in studies relying on spontaneous gesture production, children tend to use all four representational techniques at younger ages, their performance reflecting responses to particular items or communicative situations, rather than indicating a limited symbolic capacity. Furthermore, it appears that chosen techniques may depend to a greater extent than formerly recognized on properties of the object or ~~of~~ object affordances (Marentette et al., 2016).

In addition to studies on gestures in young children acquiring a spoken language, these representational techniques have been reported upon in signed languages. The ~~sign language~~ forms relevant here have often been called 'classifiers' and involve specific handshapes and orientations that function as morphemes and also indicate the semantic class to which the referent belongs by denoting some of its salient, perceived or imputed characteristic. These forms fall into categories which parallel the four representational techniques described above in Table 1: (a) constructed actions or transfer of person (similar to own-body gestures), in which the whole body represents a character and/or his/her actions; (b) handling classifiers, in which the hands represent the manipulation of the object (similar to hand-as-hand gestures); (c) entity classifiers, in which the hands represent the object as a whole or a class of objects (similar to hand-as-object gestures); (d) size-and-shape classifiers (SASSs), in which the hands represent the size or overall shape of an object.

It is only recently that sign language researchers have explicitly attempted to connect their linguistic analyses to similar analyses conducted in research on symbolic development. For example, Brentari, Di Renzo, Keane, & Volterra (2014) found that, across two cultures (American and Italian), two signed languages (Italian and American) and two spoken languages (Italian and English), signers and speakers using gestures (both adults and children) were more likely to represent agentive situations (i.e., people acting on objects) using handling strategies (i.e., hand-as-hand) rather than entity strategies (i.e., hand-as-object).

Similarly, Padden, Meir, Hwang, Lepic, Seegers and Sampson (2013) considering patterned iconicity in the American Sign Language (ASL) lexicon showed that the particular distribution of representational techniques can vary depending on context. In other words, adults seem to use different signs in ASL in naming an

object (i.e., hand-as-object strategy) or describing its use (i.e., hand-as-hand strategy). This evidence suggests that similar differences may emerge in children as well.

To sum up this rather brief overview of representational strategies used in both gestures and signs, it is clear that all four representational strategies observed both in studies on adult gesturing and signing appear to ~~already~~ be present in the representational gestures of hearing children from high as well as low gesture culture.

3. Similarities between gestures and signs

Initial studies on sign languages tended to focus on the discrete, arbitrary and categorical nature of signs, which seem more like spoken languages, but thereby overlooks the pervasive iconic nature of many sign language structures as well as similarities of signs with co-speech and silent gestures. In these early days, the priority of sign researchers was to demonstrate just how *unlike* "loose gesturing and pantomime" sign language really was (Kendon, 2014). Only in subsequent years did several researchers working on different ~~signed languages~~ begin to focus on what Cuxac (2000) has termed 'Highly Iconic Structures' (HIS), and began considering signs as visible actions or dedicated gestures with linguistic properties. Cuxac (2000) claimed that all ~~sign languages~~ are grounded upon, and exploit a basic capacity signers have in iconizing their perceptual/practical experience of the physical world. The iconization processes in ~~signed languages~~ endow them with two ways of signifying: in an illustrative way, by 'telling with showing' and by a non-illustrative 'telling without showing'. The operations signers perform when choosing an illustrative intent are defined by Cuxac as "Transfers", and conceived as traces of cognitive operations whereby signers transfer their conceptualization of the real world into the four-dimensional world of signed discourse (i.e. the three dimensions of space plus the dimension of time). These transfers are techniques for depicting information about objects, events and their relationship. These techniques are also labeled 'constructed actions', 'classifiers' and 'productive forms' by other researchers, as detailed in Table 1.

A particular striking feature of these ~~HIS~~ transfers is that they can be combined with each other or with one lexical unit to encode several kinds of different information on one (or more) referents in a simultaneous manner that has no parallel in speech. Several authors, overlooking terminological differences, have found the widespread use of Highly Iconic Structures in different genres of signed discourse and in different ~~sign languages~~ (e.g., for a comparative study on the signed languages of Italy, the United Stated and France, see Antinoro Pizzuto, Rossini, Sallandre, & Wilkinson, 2006). While the fundamental and pervasive semiotic dimension of Highly Iconic Structures in ~~sign languages~~ is hardly questionable, they are almost absent in linguistic descriptions of ~~sign languages~~ and often

relegated to a *gestural non-linguistic status*. The same fate is suffered by depictions in spoken languages that, as recently suggested by Clark (2016), although representing integral parts of everyday utterances, ~~they~~ are absent from standard models of language processing.

A structuralist/formalist approach to language assumes a representation of language as composed of discrete and listable symbols with strict boundaries and well-defined discrete categories. On the other hand, a usage-based approach, which moves beyond the linguistic vs. gestural dichotomy, leads to a more cognitive view of language in which linguistic units can exhibit variability and gradience (Occhino & Wilcox, 2017).

Some recent studies on gestures in children and adults have adopted many strategies for analysis borrowed from ~~sign languages~~ in order to investigate the continuity between gestures and signs. For example, in a study aimed at exploring motor characteristics of representational gestures produced by Italian hearing children (two/three years old) using the Picture Naming Game (PiNG) task, Pettenati, Stefanini and Volterra (2010), showed interesting similarities and consistencies in the manual parameters produced by individual children requested to label the same visual stimulus. Furthermore, some motor characteristics found in the production of these gestures were also found in the first signs produced by signing toddlers. Gestures and early signs were both produced using similar locations (e.g. the face/head, neutral space) and the same restricted set of six *basic* handshapes. In particular, the same handshapes, which appear in the gestures produced by Italian hearing children, correspond also to the same handshapes described by Boyes Braem (1981) as part of Stage 1 and 2 in her model describing the acquisition of handshapes in ASL. These findings support the view that motoric factors involved in the production of handshapes are seen both for the production of gestures and of signs and could be largely explained by the anatomy and physiology of the hand and arm.

Finally, it must be noted that the elements that we classified as gestures in spoken language don't play the same function as signs/gestures do in sign languages, is also suggested ~~in~~ recent brain studies (Newman et al., 2015).

4. Toward a new road map

In the present paper we have described different types of gestures – deictic and representational – produced by human children ~~in the~~ early stages of language acquisition. ~~The~~ four basic strategies involved in representational gestures, showing how the same strategies are found not only in children communication but also in Sign languages. ~~Furthermore, t~~he ontogenetic relation we have described between praxic actions, gestural communication and signed and spoken language

development may provide relevant suggestions and hypotheses ~~for designing~~ a new road map as briefly summarized in the four major points ~~made below~~.

1. It appears particularly fruitful to have a more differentiated view that ~~looks~~ different types of gestures (deictic and representational; involving and not involving objects) as possible building blocks of early communicative signals. In the previous road map (Arbib, 2016), performatives or deictic gestures were scarcely described although according to data from human children, they are the first gestures to emerge. The meaning of this type of gestures is given by the extralinguistic context in which they are performed. As reported by different studies (Gretscher et al., 2017) apes can produce ritualized requests mainly for food as well as showing off behaviors to get attention. ~~They~~ learn to use declarative pointing only in a human environment, suggesting a role for social responsiveness in the caregiver ~~lacking~~ in the last common ancestor (LCA-c) as described by Arbib (same issue).
2. Another important aspect is that gestural and vocal modalities are exploited together for communication by human children from ~~their~~ very early stages of ~~communication~~: hand and mouth synergies ~~(Iverson & Thelen, 1999)~~ are present since the very beginning also in the production of performatives gestures which are often produced together with vocalizations. The strong interrelation between the two modalities, gestural and vocal, is also evident in the case of the child's first two element combinations: Basic semantic relations are expressed though a combination of deictic and symbolic elements before being expressed by two representational words or signs. Despite the differences between evolution and ontogeny, the transition from gesture-word combinations to two words-combinations could suggest a possible scenario for the transition from protolanguage to language. In addition it appears that children are capable of expressing multimodal complex combinations where the distinction between gestures or words representing nouns and verbs is not needed.
3. The distinction between pantomimes and protosigns could be further explored according to results from developmental studies especially in relation to representational strategies described not only in children and adults gestural communication but also in sign language descriptions which give more attention to different kinds of 'productive'/'classifier' forms. The definition provided by Russon of great apes pantomimes as well the examples provided by Arbib in the case of LCA-m as a form of iconic gesture in which actors intentionally enact their meaning and referents and simulate engagement in activities (Arbib this same issue; Russon & Andrews, 2011) appear to be very similar to behaviors we have described in the present paper as representational gestures

adopting an "own body" strategy. Early child communication development shows that gestures adopting an own body strategy can become conventionalized within a community through ontogenetic ritualization. But a key distinction between ontogenetic ritualization in great apes and in the transition from pantomimes to "protosigns" in humans is that the former stays within dyads while the latter can migrate ~~through~~ the community.

Our hypothesis ~~could be~~ that the so called protosigns usually adopt one of the two symbolic strategies we have described above in which the hand becomes the object itself or becomes the form of the hand performing the action being referred to.

4. Given the link we have described between word comprehension and action/gesture production in infants, it is essential to consider both communicative productions in non human primates and their comprehension of audible and visual signals in order to better understand the link between actions (transitive and intransitive) and gestural and vocal production. This relationship is an actual research topic in human development, but studies investigating 'comprehension' of multimodal signals by non-human species are only ~~in a beginning~~ stage.

Finally, studying the relationship between action, gesture, sign and speech offers a valuable tool ~~for investigating~~ the overarching question of how language emerges from a non-linguistic state. In doing this, the traditional dichotomy between gestures as *enacted (gradient, variable, iconic)* and signs/word as *linguistic (categorical, invariable, arbitrary)* should be replaced with a multimodal approach to the study of both spoken and signed languages.

References

Antinoro Pizzuto, E., Rossini, P., Sallandre, M. A. & Wilkinson, E. (2006). Deixis, anaphora and highly iconic structures: Cross-linguistic evidence on American (ASL), French (LSF) and Italian (LIS) signed languages. *TISLR9* (pp. 475–495). Editora Arara Azul, Brazil.

Arbib, M. A., Ganesh, V. & Gasser, B. (2014). Dyadic Brain Modeling, Ontogenetic Ritualization of Gesture in Apes, and the Contributions of Primate Mirror Neuron Systems. *Phil Trans Roy Soc B*, 369 (1644), 20130414. https://doi.org/10.1098/rstb.2013.0414

Arbib, M. A. (2016). Towards a Computational Comparative Neuroprimatology: Framing the language-ready brain. *Physics of Life Reviews*, 16, 1–54.

Arbib, M. A. & Rizzolatti, G. (1997). Neural expectations: A possible evolutionary path from manual skills to language. *Communication and Cognition*, 29, 393–424.

Bates, E., Camaioni, L. & Volterra, V. (1975). The acquisition of performatives prior to speech. *Merril Palmer Quarterly*, 21, 205–226.

Bello, A., Giannantoni, P., Pettenati, P., Stefanini, S. & Caselli, M. C. (2012). Assessing lexicon: validation and developmental data of the Picture Naming Game (PiNG), a new picture

naming task for toddlers. *International Journal of Language and Communication Disorders*, 47, 589–602. https://doi.org/10.1111/j.1460-6984.2012.00168.x

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

Boyes Braem, P. (1981). *Features of the handshape in American Sign Language*. Unpublished doctoral dissertation, University of California, San Diego.

Brentari, D., Di Renzo, A., Keane, J. & Volterra, V. (2014). Cognitive, Cultural, and Linguistic Sources of a Handshape Distinction Expressing Agentivity. *Topics in Cognitive Science* 1–29.

Capirci, O., Contaldo, A., Caselli, M. C. & Volterra, V. (2005). From Action to language through gesture: a longitudinal perspective, *Gesture*, 5, 155–177. https://doi.org/10.1075/gest.5.1-2.12cap

Capirci, O., Cristilli, C., De Angelis, V. & Graziano, M. (2011). Learning to use gesture in narratives: developmental trends in formal and semantic gesture competence. In G. Stam & M. Ishino (Eds.) *Integrating Gestures* (pp. 189–200), Amsterdam/Philadelphia, John Benjamins.

Capirci, O., Iverson, J. M., Pizzuto, E. & Volterra, V. (1996). Gestures and words during the transition to two-word speech. *Journal of child language*, 23, 645–673. https://doi.org/10.1017/S0305000900008989

Capobianco, E., Antinoro Pizzuto, E. & Devescovi, A. (2017). Gesture–speech combinations and early verbal abilities. New longitudinal data during the second year of age. *Interaction Studies* 18(1), 55–76. doi: https://doi.org/10.1075/is.18.1.03cap.

Caselli, M. C. & Volterra, V. (1994). From communication to language in hearing and deaf children. In V. Volterra & C. J. Erting (Eds.) *From Gesture to Language in Hearing and Deaf Children* (pp. 263–277). 2nd ed., Washington, DC: Gallaudet University Press.

Caselli, M. C. (1990). Communicative gestures and first words. In V. Volterra, & C. J. Erting (Eds.), *From gesture to Language in Hearing and Deaf Children* (pp. 56–67). Berlin: Springer-Verlag. https://doi.org/10.1007/978-3-642-74859-2_6

Caselli, M. C., Rinaldi, P., Stefanini, S. & Volterra, V. (2012). Early Action and Gesture "Vocabulary" and Its Relation With Word Comprehension and Production. *Child Development*, 83, 526–542.

Clark, H. H. (2016). Depicting as a method of communication. *Psychological review*, 123(3), 324–347.

Corballis, M. C. (2002). *From hand to mouth. The origins of language*. Princeton, Princeton University Press.

Cuxac, C. (2000). La langue des signes française (LSF): les voies de l'iconocité (No. 15–16). *Ophrys*.

Enfield, N. J. (2001). 'Lip-pointing': A discussion of form and function with reference to data from Laos. *Gesture*, 1(2), 185–211. https://doi.org/10.1075/gest.1.2.06enf

Goldin-Meadow, S. & Butcher, C. (2003). Pointing toward two-word speech in children. In S. Kita (Ed.), *Pointing: where language, culture and cognition meet* (pp. 85–107). Mahwah, N.J.: Erlbaum.

Gretscher, H., Tempelmann, S., Haun, D. B. M., Liebal, K. & Kaminski, J. (2017). Prelinguistic human infants and great apes show different communicative strategies in a triadic request situation. *PLoS ONE* 12, 4, e0175227.

Halina, M., Rossano, F. & Tomasello, M. (2013). The ontogenetic ritualization of bonobo gestures. *Animal cognition*, 16(4), 653–666. https://doi.org/10.1007/s10071-013-0601-7

Iverson, J. M., Capirci, O. & Caselli, M. C. (1994). From communication to language in two modalities. *Cognitive Development*, 9, 23–43. https://doi.org/10.1016/0885-2014(94)90018-3

Iverson, J. M. & Thelen, E. (1999). Hand, mouth, and brain: The dynamic emergence of speech and gestures. *Journal of Consciousness Studies*, 6, 19–40.

Kendon, A. (2014). Semiotic diversity in utterance production and the concept of 'language'. *Philosophical Transactions of the Royal Society B*, 369, 20130293. https://doi.org/10.1098/rstb.2013.0293

Kita, S. (Ed.) (2003). *Pointing: Where language, culture, and cognition meet*. Mahwah, NJ: Erlbaum.

Leavens, D. A. (2004). Manual deixis in apes and humans. *Interaction Studies*, 5(3), 387–408. https://doi.org/10.1075/is.5.3.05lea

Liddell, S. K. & Metzger, M. (1998). Gesture in sign language discourse. *Journal of Pragmatics*, 30, 657–697. https://doi.org/10.1016/S0378-2166(98)00061-7

Liebal, K. (2016). The ontogeny of great ape gesture – not a simple story: comment on "Towards a computational comparative neuroprimatology: framing the language-ready brain" by M.A. Arbib. *Physics of Life Reviews*, 16, 85–87.

Marentette, P. & Nicoladis, E. (2012). Does ontogenetic ritualization explain early communicative gestures in human infants? In S. Pika & K. Liebal (Eds.). *Developments in Primate Gesture Research* (pp. 33–53). John Benjamins Publishing Company

Marentette, P., Pettenati, P., Bello, A. & Volterra, V. (2016). Gesture and symbolic representation in Italian and English-speaking Canadian two-year-olds. *Child Development*, 87(3), 944–961.

McNeill, D. (1992). *Hand and mind. What the hands reveal about thought*. Chicago: University of Chicago Press.

Newman, A. J., Supalla, T., Fernandez, N., Newport, E. L. & Bavelier, D. (2015). Neural systems supporting linguistic structure, linguistic experience, and symbolic communication in sign language and gesture. *Proceedings of the National Academy of Sciences of the United States of America* 112(37), 11684–89. https://doi.org/10.1073/pnas.1510527112

Occhino, C. & Wilcox, S. (2017). Gesture or sign? A categorization problem. *Behavioral and Brain Sciences*, 40, e66. doi: https://doi.org/10.1017/S0140525X15003015

Padden, C. A., Meir, I., Hwang, S. O., Lepic, R., Seegers, S. & Sampson, T. (2013). Patterned iconicity in sign language lexicons. *Gesture*, 13(3), 287–308. https://doi.org/10.1075/gest.13.3.03pad

Pettenati, P., Stefanini, S. & Volterra, V. (2010). Motoric characteristics of representational gestures produced by young children in a naming task. *Journal of Child Language*, 37(4), 887–911. https://doi.org/10.1017/S0305000909990092

Piaget, J. (1945). *La formation du symbole chez l'enfant: Imitation, jeu et rêve, image et representation*. Neuchatel; Paris: Delachaux et Niestlé.

Reddy, V. (2003). On being the object of attention: implications for self–other consciousness. *TRENDS in Cognitive Sciences*, 7(9), 397–402. https://doi.org/10.1016/S1364-6613(03)00191-8

Russon, A. E. & Andrews, K. A. (2011). Pantomime in great apes: evidence and implications. *Communicative and Integrative Biology*, 4(3), 315–317. https://doi.org/10.4161/cib.4.3.14809

Pettenati, P., Sekine, K., Congestrì, E. & Volterra, V. (2012). A comparative study on representational gestures in Italian and Japanese children. *Journal of Non Verbal Behaviour*, 36, 149–164. doi: https://doi.org/10.1007/s10919-011-0127-0

Slobin, D. (2008). Breaking the Molds: Signed language and the nature of human language. *Sign Language Studies*, 8(2), 114–130. https://doi.org/10.1353/sls.2008.0004

Sparaci, L. & Volterra, V. (2017). Hands shaping communication: From gestures to signs. In: M. Bertolaso & N. Di Stefano (Eds.), *The Hand. Studies in Applied Philosophy, Epistemology and Rational Ethics*, vol 38, (pp. 29–54). Springer, Cham.

Stefanini, S., Bello, A., Caselli, M. C., Iverson, J. M. & Volterra, V. (2009). Co-Speech gestures in a naming task: Developmental data. *Language and Cognitive Processes*, 24(2), 168–189. https://doi.org/10.1080/01690960802187755

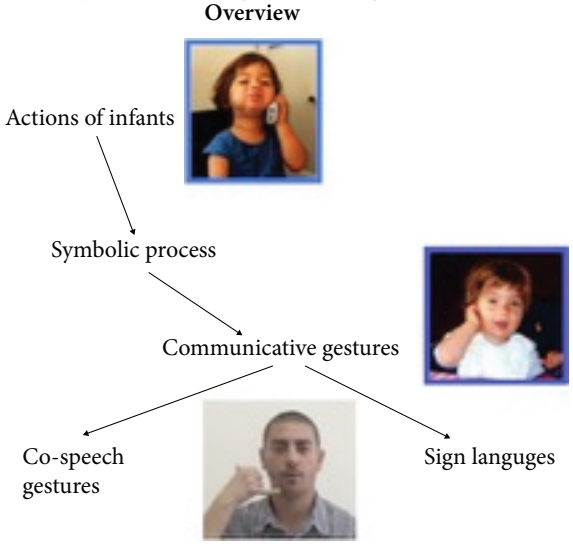
Tomasello, M. (2006). Why don't apes point? In N. J. Nicholadis & S. C. Levinson (Eds.), *Roots of human sociality: Culture, Cognition, and Interaction* (pp. 506–524). Oxford: Berg.

Tomasello, M. (2008). *The origin of human communication*. Cambridge, MA: MIT Press.

Zukow-Goldring, P. (2012). Assisted imitation: First steps in the seed model of language development. *Language Sciences*, 34, 569–582. https://doi.org/10.1016/j.langsci.2012.03.012

Supplemental material #1

Overview of the continuity from action to gesture and sign.



Supplemental material #2

Repertoire of action-gestures produced by Italian infants. Data collected through the parent questionnaire (Gestures and Words MacArthur-Bates Communicative Development Inventories (MB-CDI)

Action/Gesture	Age (in months) at which the Action/ Gesture is produced by at least 50% of children of the sample
Extends his/her arm upward to signal a wish to be picked up (*)	8
Drink from a cup or bottle containing liquid	8
Request object extending his/her hand (*)	8
Play peekaboo (*)	8
Waves the hand for 'bye bye' (*)	9
Clap hands (*)	9
Dance (*)	10
Open and close door and closet	10
Shows to adult what has in his/her hand	10
Caress (*)	10
Gives to adult what has in his/her hand	11
Points to an object or event (*)	11
Throw a ball	12
'Sleep' (leaning head on hand or pillow and closing eyes) (*)	12
'Read' (opens book, turns page)	12
Push toy car or truck	12
Shakes head or finger for 'no' (*)	12
Give slap (*)	12
Comb or brush own hair	12
Eat with a spoon or fork	12
Use remote control	12
Blow kisses (*)	12
Stir with spoon	13
Pound with hammer or mallet	13
Write with a pen, pencil, or marker	13
Put on hat	13
Kiss or hug (dolls or stuffed animals)	13
Play musical instrument	14
Clean with cloth or duster	14
Drive car by turning steering wheel	14
Put key in door or lock	14

Action/Gesture	Age (in months) at which the Action/ Gesture is produced by at least 50% of children of the sample
Sweep with broom or mop	14
Sniff flowers	14
Blow to indicate something is hot (*)	14
Put on a shoe or sock	14
Wipe face or hands with a towel or cloth	15
Shrugs to indicate 'all gone' or 'where'd it go' (*)	15
Nods head 'yes' (*)	15
Index to cheek for something tasting good (*)	16
Dig with shovel	16
Pour pretend liquid from one container to another	16
Put on a necklace, bracelet, or watch	17
Brush teeth	17
Caress (dolls or stuffed animals)	17
Push in stroller/buggy (dolls or stuffed animals)	17
Feed with spoon (dolls or stuffed animals)	17
Index to lips for 'sch' ('be silent') (*)	18
Vacuum	18
Feed with bottle (dolls or stuffed animals)	18
Rock (dolls or stuffed animals)	18
Put to bed (dolls or stuffed animals)	18
Try to put shoe or sock (on dolls or stuffed animals)	> 18
Water plants	> 18
Brush/comb hair (dolls or stuffed animals)	> 18
Pat or burp (dolls or stuffed animals)	> 18
'Type' at a typewriter or computer keyboard	> 18
Talk to (dolls or stuffed animals) (*)	> 18
Cover with blanket (dolls or stuffed animals)	> 18
Hold plane and make it 'fly'	> 18
Wash dishes	> 18
Wipe face or hands (dolls or stuffed animals)	> 18

Note:
* Not involving objects

Adapted from:
Volterra, V., Capirci, O., Caselli, M.C., Rinaldi, P. & Sparaci, L. (2017). Developmental evidence for continuity from action to gesture to sign/word. *Language, Interaction and Acquisition, 8(1)*, 13–41.

Supplemental material #3

Age of appearance of A/G and words with related meaning. The age of appearance corresponds to the age at which the 50% of children perform the action or comprehend and produce the word

		Age of appearance (in months): Age at which 50% of children do it		
Actions/Gestures	**Corresponding Italian word (and English translation)**	**Action/ Gesture**	**Word Comprehension**	**Word Production**
* Play peekaboo	Cucù (peekaboo)	8	9	18
* Waves the hand for bye bye	Ciao (Hallo; Bye bye)	9	9	17
* Clap hands	Bravo (good boy)	9	11	20
* Dance	Ballare (to dance)	10	11	24
* Give slap	Dare le tottò (give slap)	12	12	24
* Blow kisses	Baciare (to kiss)	12	12	25
* Shakes head or finger for "no"	No (no)	12	10	17
* "Sleep" (leaning head on hand or pillow and closing eyes)	Nanna (sleep)	12	11	17
Drink from a cup or bottle containing liquid	Acqua (water)	8	10	15

		Age of appearance (in months): Age at which 50% of children do it		
Actions/Gestures	**Corresponding Italian word (and English translation)**	**Action/ Gesture**	**Word Comprehension**	**Word Production**
Eat with a spoon or fork	Pappa (eat/food)	12	8	14
Comb or brush own hair	Capelli (Hear)	12	12	20
Throw a ball	Palla (Ball)	12	11	17
Push toy car or truck	Brum brum (sound for car)	12	13	17
"Read" (opens book, turns page)	Libro (book)	12	14	22
Use remote control	Televisione (television)	12	13	23
Open and close door	Apre (to open)	10	12	20

Note:
* not involving object manipulation

Adapted from:

Caselli, M.C., Rinaldi, P., Stefanini, S., Volterra, V. (2012). Early Action and Gesture "Vocabulary" and Its Relation With Word Comprehension and Production. *Child Development, 83,* 526–542. doi: 10.1111/j.1467-8624.2011.01727.x

Supplemental material #4

Semantic relations in crossmodal and unimodal combinations

Semantic relations	**Crossmodal Gesture-Word combinations**	**Unimodal spoken Two words combinations**
Notice	HI + mommy	hi + mommy
Recurrence	EXTEND ARM + other	more + other
Non-existence	ALL_GONE + food	no/all gone + water
Agent + action	POINT(to owen) + burns	mommy + puts

Semantic relations	Crossmodal Gesture-Word combinations	Unimodal spoken Two words combinations
Action + object	POINT(to the toy to be taken away) + takes away	takes it + away
Action or entity + locations	POINT(to toy) + down	go + inside
Possession	POINT(to daddy's cup) + daddy	mommy's + pen
Attribute + entity	POINT(to balloon) + big	green + that

Note: Gestures are showed in capitals English glosses. Words are given in lower case letters

Address for correspondence

Virginia Volterra
ISTC, CNR
Via Nomentana 56
00161 Roma
Italy

virgi.volterra@gmail.com