MURAJ
Minnesota Undergraduate Research & Academic Journal

Volume 1 • Issue 1
z.umn.edu/MURAJ

Human Language Evolution Points to Primate Gestures

Deniz Kihtir

Abstract:
The past century in particular has been concerned with the trajectory of human vocalizations and verbal language within an evolutionary framework. Gestural hypotheses, although immensely diverse, are only one subset of this expansive network of understandings and assertions. Gestural hypothesis revolves around the idea that human vocalizations are standard forms of communication derived from ape's physical gestures. While there is often overlap between gestural theories, researcher specialization within the theory has historically served to be more divisive than integrative. Our aim is to consolidate the multiple dimensions of the manual modality's relevance for the development of human verbal language. This will propel the theory into its strongest form, in which multiple versions share a common space, to better elucidate the origins of human language.

Only theories, not conclusive evidence, may illuminate any discussion of the potential evolutionary origins of human language. Scientists and professional researchers often observe ape characteristics and trace the hominin lineage to ground assertions justifying human language origin and its development. Humans are not the only species shown to implement vocal communication. Cross-species comparisons of the properties and function of vocalizations may provide insights into the environmental conditions or benefits that advantage this development. Observing the functions and nature of primate vocalizations may initially appear to potentially reveal the tangible advantages of verbal language as observed amongst behaviorally modern humans. But vocalizations or other modes of communication are not to be inherently considered language, as language might be recognized as an intentional and structured form of communication reflecting elaborate processing and symbolic representation. Gestural theorists argue that primate vocalizations are standardized displays of emotion (Fogassi, 2007), with human language developing from ape gestural communication and human manual manipulation (Stout, 2011). But the theory is expansive, with some proponents concerned with describing primate gestural expressions and others proposing that manual manipulation required for hominin material culture motivated protosign, eventually developing into protolanguage. Considering the various environmental conditions and physiological mechanisms responsible for human behaviors, segmented gestural theories collectively argue for a rationale of the manual modality.

The technological advancements of the hominin lineage observed over the past 3.5 million years may have co-developed alongside, or motivated, gestural communication. The often multi-step manual procedures required to create stone tools, extract safe and nutritious plant materials, start fire, or hunt big game would have required cooperation and information transmission for successful execution (Sterelny, 2012). Accurately constructing a handaxe is, for example, a hierarchical process requiring various stages of input and intention. As such, the neural mechanisms underlying the behavior involved in making a hand axe resemble the language recognition processes (Stout, 2011). The capacities of the human brain's Broca's area, associated with language processing and production, supposedly extend to extralinguistic behaviors such as tool use and production (Higuchi, 2009). Due to the fact that both linguistic expression and tool usage are both goal-directed motor activities, these cognitive overlaps evidence potential for convergent evolution (i.e. simultaneous or overlapping development) (Stout, 2011). Distinguished ability for pedagogy and social learning may have either encouraged or been a consequence of social foraging systems where individuals shared and recognized common goals.The deliberate and premeditated planning required for the success of such a foraging system would have required cooperation and attributing intentions of other group-members. Hierarchical processes that motivate understandings of complex structures of tool-making are related to purposeful communication systems (Stout, 2011). Enhanced theory of mind (i.e. ascribing the mental states

CC BY NC

of others) and presence of shared goals would have, according to Stout (2011), advanced intentional communicative gesture systems.

Gestural language theories often emphasize how monkey vocalizations are not associated with regions of the brain stimulated during human language and communication. Monkey vocalizations appear to trigger their limbic circuit, which is associated with both emotional displays and other standardized calls or remarks (Fogassi, 2007). Their inflexible vocalizations are limited by genetic determination (as primary vocal anatomy restricts human-like utterances) and are not thought to deliver complex information, serving instead as generalized warnings or exhibitions of emotion. While primate vocalizations can specify (and be understood by others as indicating) various types of predators in their calls (Price, 2015), the vocal anatomy restricts capacities for the range of sounds produced and limits the expression of language properties (Fogassi, 2007). These primate vocalizations are argued not to be combined in individually-constructed and infinitely varied complex sets that indicate detailed information as observed amongst human languages (Fogassi, 2007). Initial early hominin speech is unlikely to have expressed material with syntax or symbols, but rather indicate actions or objects through the manual modality. Those constructing gestural origins hypotheses are more inclined to reference those audio-mirror neurons which display potential bases for the evolution of gestural communication and action recognition (Kohler, 2002).

The F5 region of the monkey brain is considered homologous to Broca's area in the human brain in its physical arrangement, and the actions that stimulate it indicate recognition of gestural communication. The same Macaque motor neurons are fired when witnessing an action and performing the action themselves, including when an experimenter hits or breaks an object and when personally performing the same actions themselves (Kohler, 2002). Mirror neurons were also stimulated when the Macaque executed a particular hand-related action as when it heard the sounds (auditory cues) associated with performing the action (Kohler, 2002). These audio-visual mirror neurons represent tangible evidence for primate capacities to store an action and its associated sounds as information, either when witnessing or performing the action. Gestural theory indicates that the neurons of certain primates are responsible for extracting meaning from particular actions and even recognizing the auditory indicators of such activities. Recognizing the inter-individual dependencies of language, the ability to perceive the meaning, intent, or goal that a fellow communicator is expressing (theory of mind) is a potential foundation for human language emergence (Kohler, 2002). Gestural theorists specify that the properties of inter-individual dependencies of language extended to developing and extending material culture, such as tool usage and production. However, as noted, the F5 region is considered homologous to human Broca's area 44, and the presence of macaque goal detection and meaning evaluation indicate the primate preadaptations for human language potential.

Arbib (2010) proposes seven stages that distinctly trace the primate evolution of human language systems from grasping adaptations in the mirror-systems hypothesis. The stages are as follows: stage 1 (S1) broadly references manual grasping adaptations, S2 includes the mirror system recognition of actions, S3 involves the development of a simple imitation system to repeat observed manual action somewhere within the early hominin lineage, S4 would have allowed the capacity to engage in a complex imitation system, which is often proposed to involve tool production or generalized manual expertise, ultimately motivated by human recognition of both sequential and hierarchical structures (Arbib, 2010). Arbib (2010) argues that this complex imitation would have eventually encouraged the deliberately communicative protosign characteristic of S5. Protosign is defined as conventionalized manual communicative gestures (Arbib, 2005). Protosign is imagined to have emerged before structured vocal protolanguage, not only because of potential biological constraints, but because the brain mechanisms specializing in human language systems are in fact inclined to utilize facial and manual mediums (Arbib, 2010). A biologically increased ability to utilize or control the vocal tract would have facilitated the development of proto-language in S6, which overlapped with integrated protosign in order to ultimately represent modern human language in S7 (Arbib, 2010). Gestures in the mirror-systems hypothesis serve as the platform of symbolic representation ability, allowing protospeech to evolve to serve as a symbolic indicator of a designated meaning (Arbib, 2010). Arbib (2010) implies that exploiting this capacity to represent objects or concepts symbolically encouraged further encephalization [or at least specialization in relevant brain areas] to develop languages that employed various communicative modes, including facial, vocal, and gestural.

Characteristics that define primate gestures often overlap with the same functions served by human gesture, but lack other properties associated with human language that likely developed after the human - chimpanzee split. Human gestures are thought to mentally represent actions or objects, often requiring abstract perception and

thought). While this is partially true for apes, observed primate gestures often indicate specified later actions (Cartmill, 2011). These gestures are not entirely manual, but as with humans, are often expressed through facial or even whole-body position for communication. Social considerations are clear, as gestures depend on the sender's visibility to others, and communicators often modify, repeat, or clarify their gestures if they do not receive their preferred response (Cartmill, 2011). Recognizing intention is proposed to form a foundation for human language evolution (Cartmill, 2011). The differences between primate gesture and human languages are primarily recognized as universal human sentence properties (syntax or verb/noun interactions), the expansive range of subjects humans functionally communicate about, and deictic expressions (Cartmill, 2011). Yet the ability for great ape species to learn and imitate gestures suggests that they can be motivated by the responses and reactions of others and modify their behavior perhaps to advance a particular goal or respond specifically to these social cues. Though these gestures are not identically representational as they are for humans, primates have been found to rely on symbols to determine the relationship between individuals and the particular actions they perform (Cartmill, 2012). Primates often observe gestures and body language to predict or conceptualize behaviors, and this form of social learning and observation is proposed to be evidence of a preadaptation for human language (Cartmill, 2012). Human cognitive development may have expanded on this social learning and extension beyond the self, allowing for representational gestures to relay complex events.

Human infant gestural expressions are often expected for successful language development as the primary exhibition of intention to communicate and share attention (Iverson, 2005). Gestures generally predate all verbal language, and articulations and are expected within the first year of life. These gestures often involve clapping or pointing at some third party object to personally recognize it or perhaps to share attention, which is considered an infant developmental milestone before verbal language can be employed (Iverson, 2005). If these gestures allow infants to express particular information or meaning before the physical possibility of articulating a spoken language, they might be considered to transition or guide verbal language learning (Iverson, 2005). Infants who supplement simple one-word utterances with different but relevant gestures will develop two-word combinations faster than those who point to a complementary image or object with a one-word description (Iverson, 2005). And the children who supplement their speech with supplementary images are actually recognizing and communicating two separate pieces of information, one example describing a child saying "nap" and pointing at a picture of a bird who was sleeping, encouraging two-word combinations and so on (Iverson, 2005). Iverson (2005) ultimately proposes that gestures may serve to signal a child's cognitive maturity to their caretaker, displaying a preparedness for verbal stimuli. Gestural communication is most influential for children in earlier stages of development and loses its relevance with age, though gesture continues to accentuate or elaborate adult verbal language communication around the world.

While human infant gestures are often primary indicators of intention to communicate, hearing individuals will continue to gesture through adulthood. Surely deaf humans use gestural language as a primary form of communication, but [perhaps inadvertent] co-speech gestures are argued to emphasize the most crucial or important parts of a spoken sentence. These co-speech gestures, including the facial and manual modalities, might indicate emotion and are often serve inter-cultural or global functions. Gestures more generally are argued to serve vivid and efficient communicative functions. Deictic gestures are considered basic indicators of human infant desire to communicate or reference a third-party, perhaps even advancing a joint-attention interpretation of language evolution and function (Cartmill, 2012). According to Cartmill (2012), gestures are employed to ground abstract internal understandings of concepts or items into a tangible, representative form that can be shared. While conventional gestures are culturally variable, representational gestures communicate actions, concepts, or items either through iconic symbols or figurative moments (Cartmill, 2012). Increased encephalization potentially allotted recognition of the correspondence between individual actions, perhaps neurologically allowing for imitative behaviors (Arbib, 2005). Considering primate neural mechanisms, these imitations would eventually prompt pantomiming, which is crucial for proto-sign (Arbib, 2005). Abstract conceptual understandings and desire to tangibly express such thoughts may have motivated early hominin communication. For example, [add details] indicate direction or clarify procedures through deictic gestures. Proponents, such as Arbib (2010), do not tend to argue that gestural communication was ever the exclusive mode of communication, but that combining vocalizations (not speech) with gestural indicators would have encouraged the development of proto-language.

A homesign, which possesses both combinatorial and segmented displays of information, is often inde-

pendently developed and constructed by deaf children. Even without interactions with a formal sign language, individually developed homesign often tends to express expected properties of language, such as placing objects or nouns before verbs (Goldin-Meadow, 2016). This sentence structure is clearly not mere imitation of hearing individuals' co-speech gestures, perhaps implying an inclination for human linguistic structures even when expressed manually (Goldin-Meadow, 1995). Segmented gestures indicating particular objects or actions are independently developed to function as individual references in an intentional sequence. Combining parts to form a particular expression are often observed amongst homesigners, where certain hand positions represent objects and applied motions indicates an action to form a complete (as intended) gesture (Goldin-Meadow, 1995). This particular account recognizes the diverging functions of both gestural communication and speech, arguing that the manual modality allows for mimetic expression while verbal language can articulate analytical responses (Goldin-Meadow, 1995). Coevolution of gestural and vocal communication systems is suggested by the complementary interactions between these spheres that allow humans to efficiently express content and encourage clarity.

While vocal capacities of nonhuman primates are relatively inflexible, enculturated primate studies have demonstrated a proclivity towards gestural communication acquisition. Human-like speech is nearly impossible for primates to perform, or imitate, considering general vocal anatomy constraints and positioning of the larynx, and this immense difficulty has been revealed through primate research (Fogassi, 2007). Yet both captive wild and reared-from-birth apes have been demonstrated to learn (or perhaps imitate) gestures to aid in communication and even advance one's goals. One of the especially popular examples of this ability to not only conceptualize the content that gestures can represent but also implement such communication manually with relative ease is a chimpanzee named Washoe. Washoe learned hundreds of signs (around 350) throughout her life, and began to both combine signs and imitate signs and behaviors that the researchers were not actively teaching her (Gardner, 1969). Washoe would not only limit her use of these gestures to reference particular items or concepts (such as sorry or funny), but would also implement them when asking for things (Gardner, 1969). Signs were used to indicate physically present objects and extended to photographs or illustrations, clearly displaying the conceptual capacity to recognize icons and perhaps indexes. And many of Washoe's word combinations were not merely observed or imitated (though imitation is surely informative of the cognitive overlaps between humans and other primates) but personally invented. Examples of personally motivated communication might include Washoe's "gimme drink please" or "listen dog" (Gardner, 1969). Gestural theory advocates often cite such enculturated ape studies to display shared primate imitative behaviors and potential for symbolic or iconic representations. Primate communicative displays in the wild, often manual gestures or facial indicators, are behaviors regularly employed to enrich social interactions or induce particular conspecific responses. Vocal communications do not serve the same social purpose, and enculturated ape studies highlight the behavioral capacities of primates that motivated human evolution of language.

Fay (2014) clearly establishes the evaluation that gesture possesses greater opportunity to successfully communicate emotions, actions, and objects than vocalizations. These iconic gestures that might be described as motivated signs were purportedly expressed by hominins starting at some point in the early hominin lineage (Fay, 2014). The demands of social grouping and mutual cooperation (to ensure optimal foraging success) is proposed to necessitate relaying certain knowledge between generations, pressuring structured language emergence. Fay (2014) broadly proposes that vocalizations are a less effective means of communication than are gestures. Vocalizations, in the absence of a shared verbal language system, are limited in their capacity to express actions or objects as thoroughly as gestures might be inclined to (Fay, 2014). While the authors predicted that combining both gestures and vocalizations might produce more effective and accurate information transmission than just vocalizations alone, results negate even this expectation. Currently, most humans around the world who speak a verbal language will, according generalized assumptions, either deliberately or more subconsciously, utilize gestures to enhance a particular message or statement. Yet when these modes of linguistic expression are stripped of their shared communicative value, vocalizations were not especially effective for information transmission, even when combined with gestures (Fay, 2014). Spontaneously constructed gestures (highly iconic and indexical signs) were instead most efficient for relaying messages, implying an intrinsic inclination to recognize certain properties of human language (Fay, 2014). The manual modality is ultimately supposed to indicate emotions, objects, or action more successfully than vocalizations have the capacity to indicate without the shared knowledge of a linguistic system.

Observing the gestural theory of language reveals an expansive range of proposals that detail the contribu-

tions of the manual modality in respect to language evolution. Theories range from concretely grounded comparative neurological indications to more generalized comments about human inclination. Gestural theorists are often specialized in their focus, whether observing the functions of gestures within nonhuman primates in the wild, speculating early hominin activities and pressures, or studying anatomically modern human linguistic behaviors and tendencies. Individual contributions, however, collaborate to construct a cohesive and integrative theory on the relevance of the manual modality to human verbal language development. While perhaps not all researchers come to the same conclusions or necessarily advance their arguments using similar evidence, gestural communication continues to captivate those curious about the origins of human language.

References:

Arbib, Michael A. 2010. *Action to language via the mirror neuron system.* Cambridge: Cambridge University Press.

Arbib, Michael A. 2005. "From Monkey-Like Action Recognition to Human Language: An Evolutionary Framework For Neurolinguistics." *Behavioral and Brain Sciences* 28 (02).

Cartmill, Erica A., Sian Beilock, and Susan Goldin-Meadow. 2011. "A Word in the Hand: Action, Gesture and Mental Representation in Humans and Non-Human Primates." *Philosophical Transactions of the Royal Society B: Biological Sciences* 367 (1585): 129-43.

Fay, Nicolas, Casey J. Lister, T. Mark Ellison, and Susan Goldin-Meadow. 2014. "Creating A Communication System From Scratch: Gesture Beats Vocalization Hands Down." *Frontiers in Psychology 5*. Web.

Fogassi, Leonardo, and Pier Francesco Ferrari. 2007. "Mirror Neurons and the Evolution of Embodied Language." *Current Directions in Psychological Science* 16 (3): 136-41.

Gardner, R. A., and B. T. Gardner. 1969. "Teaching Sign Language to a Chimpanzee." *Science* 165 (3894): 664-72.

Goldin-Meadow, Susan. 2016. "What the Hands Can Tell Us About Language Emergence." *Psychonomic Bulletin & Review* 24 (1): 213-18.

Goldin-Meadow, Susan, Carolyn Mylander, and Cynthia Butcher. 1995. "The resilience of combinatorial structure at the word level: morphology in self-styled gesture systems." *Cognition* 56 (3): 195-262.

Higuchi, Satomi, Thierry Chaminade, Hiroshi Imamizu, and Mitsuo Kawato. 2009. "Shared neural correlates for language and tool use in Broca's area." *NeuroReport* 20 (15): 1376-381.

Iverson, J. M., and S. Goldin-Meadow. 2005. "Gesture Paves the Way for Language Development." *Psychological Science* 16 (5): 367-71

Kohler, Evelyne, Christian Keysers, M. Alessandra Umiltà, Leonardo Fogassi, Vittorio Gallese, Giacomo Rizzolatti. 2002. "Hearing Sounds, Understanding Actions: Action Representation in Mirror Neurons." *Science* 297 (5582) : 846-48.

Liebal, Katja, and Josep Call. 2011. "The Origins of Non-Human Primates' Manual Gestures." *Philosophical Transactions of the Royal Society B: Biological Sciences* 367 (1585): 118-28.

Price, Tabitha, Philip Wadewitz, Dorothy Cheney, Robert Seyfarth, Kurt Hammerschmidt, and Julia Fischer. 2015. "Vervets Revisited: A Quantitative Analysis of Alarm Call Structure and Context Specificity." *Scientific Reports* 5 (1): 1-11.

Sterelny, Kim. 2012. "Language, Gesture, Skill: the Co-Evolutionary Foundations of Language." *Philosophical Transactions: Biological Sciences* 367 (1599): 2141–2151.

Stout, D., and T. Chaminade. 2011. "Stone Tools, Language and the Brain in Human Evolution." *Philosophical Transactions of the Royal Society B: Biological Sciences* 367 (1585): 75-87.