

From babbling to articulatory echo neurons and unsolved questions of syntax

Friedemann Pulvermüller, MRC Cognition and Brain Sciences Unit, Cambridge, UK,  
[friedemann.pulvermuller@mrc-cbu.cam.ac.uk](mailto:friedemann.pulvermuller@mrc-cbu.cam.ac.uk)  
<http://www.mrc-cbu.cam.ac.uk/People/Friedemann.Pulvermuller.html>

Summary: Mirror neurons demonstrate close links between the cortical mechanisms of action and action-perception, which are best explained by associative learning in long distance cortico-cortical connections. The resulting *action-perception networks, APNs*, are the substrate for repetition and imitation. Articulatory echo neurons linking vocalisation sounds and syllables to their respective motor programs may have evolved together with, not subsequent to, mirror neurons for manual activities, because both exploit the same neuroscientific principles. The mystery of the huge vocabulary size of human languages, which contrasts with the limited set of monkey calls, may be grounded in behavior: Only humans go through a stage of vocal and manual babbling during which abundant arbitrary motor patterns are practiced excessively. The babbling stage, which must have a specifically human genetic basis, sets up a great variety of APNs, including echo and mirror neurons and serving as a brain basis of large vocabulary build-up and selective reciprocal links between the neuronal ensembles embodying words and nonlinguistic actions and perceptions. The evolution or emergence of syntax and recursion may require additional specific neural mechanisms.

Arbib has done a great job in summarizing research on mirror neurons and their putative role in language evolution and emergence. It is possible that language evolved indirectly, from a mirror neuron (MN) system for grasping, through intermediate stages of grasp imitation, pantomime, proto-sign language, and, finally, spoken proto- and full-fledged language. However, it is not clear that the detour through the manual mode is necessary.

The mirror neuron: explanans or explanandum?

The discovery of MNs was a great breakthrough in cognitive neuroscience. Rizzolatti, Fogassi, Gallese, Fadiga and their colleagues (for review, see Rizzolatti & Craighero, 2004) proved a link between actions and their perceptions using grasping as a flagship example. The primacy of grasping and similar actions in neuroscience research may have resulted from accidents and practical considerations in the laboratory: For a human experimenter, it is, for example, much easier to imitate a grasp than a monkey call. The biological principles that lead to the emergence of MNs for grasping may as well yield other action-related sensory-motor links, especially between representations of articulatory gestures and the accompanying sound sequences. The prediction is that, when experiments are done to search for these "echo neurons", they will eventually be found. In summary, the MNs for grasping and other manual actions are important *because they reveal functional principles immanent to the nervous system*. They do not explain language, but are explained by the same principles that are also indispensable for language networks

Imitation and the babbling hypothesis

MNs *per se* are not sufficient for the ability to repeat a great range of actions performed by others, as demonstrated by the absence of such abilities in apes and monkeys. The

emergence of the imitation ability of human infants, both in the manual and the vocal domains, thus appears mysterious even after one realises the importance of a mirror/echo mechanism for manual/articulatory actions. What is the brain basis of the huge repertoire of repeatable actions in infants?

To bind actions and the correlated sensory (visual and auditory) patterns, neuroanatomical connections between the respective brain areas are necessary, along with a learning mechanism that maps correlation of neuronal activity onto connection strength. When an action is performed repeatedly, the neurons controlling the action (in motor, premotor and prefrontal areas) become active together with neurons in posterior sensory cortex that respond to the auditory and visual aspects of the action, the shape and trajectory of the grasping hand or the noise burst and formant transition of the spoken syllable. For auditory input, the primary auditory cortex is as important as the areas surrounding it (belt and parabelt areas in macaca, Brodmann areas 42 and 22 in humans). In monkeys, the action-perception networks, APNs, linking actions to their sensory aspects can develop whenever an action is performed repeatedly, but since the action repertoire is limited, a limited set of APNs results (Pulvermüller & Preissl, 1991; Pulvermüller, 1999).

A fundamental distinctive feature of human ontogenic development is the so-called babbling phase during which syllables, but also body movements, are repeated excessively. The babbling stage precedes, and may be necessary for, imitation of articulations and the first spoken words. The repeated performance of virtually random but actually well-structured actions provide an ideal basis of the setup of APNs and may be crucial for the human repertoire of repeatable gestures, both in the articulatory and manual domains. This view known as the Babbling Hypothesis (Braitenberg & Pulvermüller, 1992; Fry, 1966; Pulvermüller, 1999; Pulvermüller & Preissl, 1991) links the vocabulary explosion to babbling, a specifically human behavior, which is genetically pre-programmed. It implies that the APNs set up by babbling include manual mirror and articulatory echo neurons in premotor, but possibly also in prefrontal and temporal cortex that become active when an articulation or manual gesture is being performed or perceived.

#### Linking words to categories of knowledge

The syllable representations set up by babbling are necessary for repeating and for learning word forms used by others. These word forms are available for conceptual binding. If a word form is used together with a nonlinguistic action, for example when the infant regularly hears it when picking up a fruit or kicking a ball, APNs in perisylvian language areas link with action representations in other parts of premotor and premotor cortex, so that after such learning, the perception of an action word alone can "light up" specific parts of motor and premotor cortex (Hauk, Johnsrude, & Pulvermüller, 2004; Pulvermüller, Härle, & Hummel, 2000; Shtyrov, Hauk, & Pulvermüller, 2004). It is noteworthy that this mechanism is not restricted to Broca's area but that it may even involve the centro-dorsal leg motor and premotor cortex, arguing for a wide cortical distribution of mirror neurons. On the other hand, when words are frequently used together with objects, canonical neurons in the ventral visual processing stream along with neurons in inferior frontal cortex responding to the visual features of the object may become active with perisylvian word related APNs, thus binding word and object representations to each other. These considerations, which are substantiated by neurophysiological and metabolic imaging results obtained in humans (Pulvermüller,

2001), suggest that multimodal neurons of the mirror and echo types, along with canonical object related neurons, are not imprisoned in a small part of premotor cortex or Broca's area, but are spread out over large parts of the human cortex, where they serve the binding of information about language, actions and perceptions.

From words to syntax and recursion: Genes or geniuses?

It is parsimonious to claim that much of human language is a consequence of elementary neuroscientific principles and it appears reasonable that syntactic structure building is a feature emerging from protolanguages shared with non-human species. However, in Arbib's proposal, the gradual invention of syntax is attributed to an ancestry of geniuses each of whom is thought to have invented a fraction of grammar. Put this way, the actual mechanisms necessary for the serial ordering and hierarchical, or heterarchical, structuring of words in sentences remain unspecified. The genius postulate may therefore run into the same problems as the earlier postulate of a big genetic mutation, or series of unnoticed small mutations, which was thought to have led to the growth of language in the human brain. To make such and similar postulates more concrete, I suggest discussing the putative circuits of syntax and recursion in terms of neurons and their links (Pulvermüller, 2003). Syntactic circuits, however hypothetical, can then provide clues about which features of syntax are explained by known neuroscience principles and which require evolutionary change and neuroanatomically hard-wired information.

## References

- Braitenberg, V., & Pulvermüller, F. (1992). Entwurf einer neurologischen Theorie der Sprache. *Naturwissenschaften*, 79, 103-117.
- Fry, D. B. (1966). The development of the phonological system in the normal and deaf child. In F. Smith & G. A. Miller (Eds.), *The genesis of language* (pp. 187-206). Cambridge, MA: MIT Press.
- Hauk, O., Johnsrude, I., & Pulvermüller, F. (2004). Somatotopic representation of action words in the motor and premotor cortex. *Neuron*, 41, 301-307.
- Pulvermüller, F. (1999). Words in the brain's language. *Behavioral and Brain Sciences*, 22, 253-336.
- Pulvermüller, F. (2001). Brain reflections of words and their meaning. *Trends in Cognitive Sciences*, 5(12), 517-524.
- Pulvermüller, F. (2003). *The neuroscience of language*. Cambridge: Cambridge University Press.
- Pulvermüller, F., Härle, M., & Hummel, F. (2000). Neurophysiological distinction of verb categories. *Neuroreport*, 11(12), 2789-2793.
- Pulvermüller, F., & Preissl, H. (1991). A cell assembly model of language. *Network: Computation in Neural Systems*, 2, 455-468.
- Rizzolatti, G., & Craighero, L. (2004). The Mirror-Neuron System. *Annu Rev Neurosci*, 27, 169-192.
- Shtyrov, Y., Hauk, O., & Pulvermüller, F. (2004). Distributed neuronal networks for encoding category-specific semantic information: the mismatch negativity to action words. *European Journal of Neuroscience*, 19(4), 1083-1092.