

Leonardo Fogassi, Gino Coudé and Pier Francesco Ferrari

The extended features of mirror neurons and the voluntary control of vocalization in the pathway to language

Abstract: In this book it has been proposed that the mirror system can be a scaffold for building a language-ready brain, because of its property of matching action observation with action execution, a feature that can correspond to the “parity” requirement for communication. In this commentary we will first emphasize two properties of mirror neurons and motor cortex that may have contributed to language: the generalization of the property of understanding action goals and the capacity to decode the goal of action sequences. Then we will propose, based on recent behavioural and neurophysiological data in monkeys, that the vocalization in non-human primates could have reached a partial voluntary control, thus contributing to the emergence of a communicative system relying on the coordination of gestures and utterances.

Keywords: monkey, goal generalization, action, brachiomanual gestures

Leonardo Fogassi: Dipartimento di Neuroscienze, and Istituto Italiano di Tecnologia (RTM), Via Volturno 39, 43125 Parma. E-mail: leonardo.fogassi@unipr.it

Gino Coudé: Dipartimento di Neuroscienze di Parma, V. Volturno 39 and National Institutes of Health, Poolesville, MD

Pier Francesco Ferrari: Dipartimento di Neuroscienze di Parma and Istituto Italiano di Tecnologia, V. Volturno 39 and National Institutes of Health, Poolesville, MD

Premise

The question of the origin of language is still a stimulating matter of debate. One of the discoveries that had a major impact among scholars came from the field of neurophysiology two decades ago: The mirror neurons. The property of matching action observation with action execution offered a parsimonious explanation of how a speaker and a listener can share roughly the same meaning about an utterance, a requirement for communication known as “*parity*”. In this book, Michael Arbib gives an extensive account of how the mirror system could have interplayed

with other brain mechanisms through a lengthy process of biological and cultural evolution to allow a “*language-ready*” brain to be built.

In this commentary, we will focus on the central claim of the book, that relates to the role of mirror neurons in language evolution. Then we will also offer a critical view on how non-human primates vocal communication, in conjunction with gestures, could have had an active role in the emergence of the first voluntary forms of utterances (protospeech).

1 How extended features of mirror neurons have been exploited in the pathway to language. The case of goal generalization and action sequencing

Arbib clearly states as the central claim of his book: “The mechanisms which support language in the human brain *evolved atop* a basic mechanism not originally related to communication. Instead, the mirror system for grasping with its capacity to generate and recognize a set of actions, provides the evolutionary basis for *language parity* – the property that an utterance means roughly the same for both sender and receiver.” (p. 120).

The striking aspect of *The Mirror System Hypothesis* (MSH) is that it offers an explanation of a phenomenon that would be otherwise a conundrum, that is, how two communicating individuals can share common knowledge. Motor neurons of premotor cortex are activated during goal-directed motor acts such as grasping, reaching or tearing (Rizzolatti and Craighero 2004) and constitute a kind of internal storage of motor knowledge (“motor vocabulary”). Sensory information can access this internal motor knowledge, allowing its translation into action. The core of this proposal is that of stressing the role of the motor system in providing a plausible neurological explanation for language parity and for drawing an evolutionary scenario in which gestures are important for the transition to a vocal-based communication system. In considering the properties of the mirror system described by Arbib we would like to emphasize some aspects that, in our opinion, have not been sufficiently addressed, and that could help clarify the importance of this mechanism for language evolution.

The first aspect relates to the demonstration that mirror neurons can generalize motor goals that are usually outside the individual’s natural motor repertoire,

thus demonstrating a certain degree of plasticity. In fact, it has been shown that sensory-motor experience with tools can affect mirror neurons' discharge during observation of goal-related motor acts performed with these tools. Two types of studies support such plasticity. In one of them (Rochat et al. 2010; Umiltà et al. 2008), monkeys were trained to take possession of food with pliers. After training was completed, single neuron recording revealed that grasping neurons of area F5 fired not only when the monkey grasped a piece of food with the hand, but also when it grasped it with the tool. More interestingly, some of these neurons responded when the monkey observed an experimenter grasping food with his hand or with the same tool used by the monkey. Thus, mirror neurons can generalize their response to motor acts that, after training, were incorporated in the monkey motor repertoire.

In another study (Ferrari et al. 2005), monkeys were not trained to use a tool for getting food, but were exposed for a long experimental period to the sight of tools (sticks, pliers, etc.) that the experimenter used to pick up food that, in some cases, was given to the monkey. The monkey therefore had opportunities to physically interact with the tool, even though it could not use it according to its function. Also in this study it has been shown that a percentage of mirror neurons discharged during observation of motor acts performed by the experimenter with the tool. This result has been interpreted in terms of the possibility of the motor system to extend the capacity of understanding goals to observed actions that have not been motorically experienced *Sensu stricto*.

The two described types of goal generalization coded by mirror neurons are very important, in our opinion, to elucidate the flexibility of the motor system, which could be critical for language evolution. In fact, if, according to Arbib's proposal, a mirror system for pantomimes (in the pathway to protosigns) follows, in evolutionary terms, that for goal-directed actions, it is plausible that the above described neural plasticity represents a substrate from which a system matching action observation with action execution may have expanded, incorporating also several types of intransitive gestures endowed with new meanings, as those described by Arbib as protosigns.

A second, important aspect of Arbib's book that, we believe, should be integrated and further expanded, is the relation between the capacity of the cortical motor system to organize action sequences and the need of vocal and gestural communication to combine and control different effectors in order to produce complex social signals.

For example, studies on gestural communication in apes have shown that chimpanzees can use hand gesture sequences for producing signals in one or more modalities (visual, auditory or tactile) (Liebal et al. 2004). Although the sequences could involve the repetition of the same gesture, it has been noted

that several of them were composed by three or more different gestures. Similar properties emerge also when we look at the vocal-based communicative systems. In fact, speech production requires a sequential activation of the phono-articulatory tract, in which the movement of the jaw must be coordinated with larynx muscles contraction and lips and tongue movements. This sequential organization is controlled according to specific rules that take into account mechanical constraints. In addition, when we consider language in its complexity, also the structure of what is emitted needs to be regulated by combining and coordinating elements, i.e. words, hierarchically, into meaningful phrases (Pinker and Jackendoff 2005). This latter aspect constitutes the basis for syntax.

In this respect, we recently proposed that some elements of the sequential structure involved in gestural and vocal communication could have been exploited starting from the neural organization of the motor cortex (Fogassi and Ferrari 2012). The evidence of this comes from monkey neurophysiological investigation. A series of studies have shown that neurons in mesial cortices (pre-supplementary motor area, pre-SMA/F6 and supplementary motor area, SMA proper/F3) and prefrontal cortex activated specifically while monkeys executed sequences of movements (Tanji, 2001; Tanji and Hoshi, 2008). These neurons could code either the sequence, the order of a movement inside a sequence or the final location of a trajectory. In another series of studies, carried out in our laboratory (Fogassi et al. 2005; Bonini et al. 2011), the responses of parietal (area PFG) and premotor (area F5) grasping neurons during execution and observation of natural action sequences were assessed (Figure 1). The results showed that during execution most of these neurons discharge differently depending on the specific action sequence in which the grasping act is embedded (Figure 1C, left). Notably, this differential response is shown also by mirror neurons during observation of grasping embedded in different action sequences performed by another individual (Figure 1C, right). Together, these data suggest that neurons in parietal and premotor cortex are organized in motor chains, each coding a specific action goal (Chersi et al. 2011). The advantage of this organization is to facilitate the smoothness of action execution and an efficient action control, especially when different motor acts and effectors need to be activated in a specific sequence.

Summing up, the premotor-parietal motor system plus the prefrontal cortex can provide a substrate for sequential organization and hierarchical combination of motor elements. We have proposed that such an organization has been exploited in other domains including some aspects of the syntactic structure of language (see also Fogassi and Ferrari 2007).

Several investigations attempted to examine the involvement of the human inferior frontal gyrus (IFG – which is part of the mirror neuron system) in syntac-

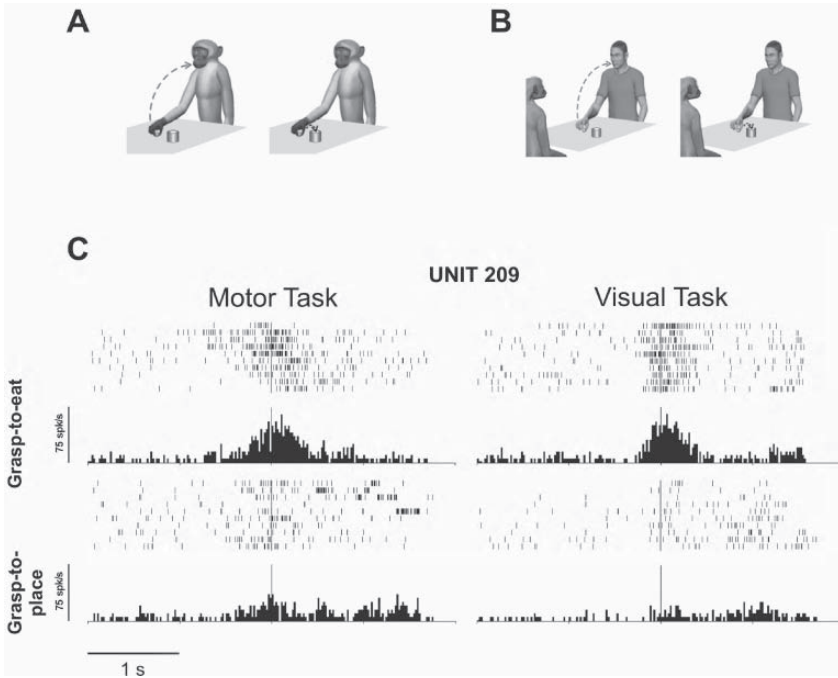


Fig. 1: (A) Motor task. The monkey, starting with its hand from a fixed position, reaches and grasps a piece of food (or an object), then it brings the food to the mouth and eats it (A, left, grasp-to-eat Condition) or places it (or the object) into a container (A, right, grasp-to-place condition). (B) Visual task. The experimenter performs the same two conditions as in (A) in front of the monkey, that simply observes them. (C) Example of the motor and the visual responses of an F5 mirror neuron during grasp-to-eat and grasp-to-place conditions. Rasters and histograms are aligned (vertical bar) with the moment when the monkey (left) or the experimenter (right) touched the food to be grasped. This neuron was differentially activated during grasping according to the final goal of the action in both the motor and the visual task.

tic construction and in the hierarchical structures of sentences. There is some evidence from Broca's patients (Fazio et al. 2009; Clerget et al. 2009; Pulvermüller and Fadiga 2010) that the type of deficits consist not only in the impairments in phono-articulation but also, depending on the extension of the lesion and on the involvement of the nearby areas, in the processing of the hierarchical structure of a sentence (Sapolsky et al. 2010; Wilson et al. 2010). In addition, fMRI studies in normal subject show that in the IFG region (BA 44, 45 and 47) there are subsectors that are involved in phonology, semantics and syntax (Bookheimer 2002; Hagoort 2005; Haller et al. 2005).

2 The voluntary control of vocalization evolved in parallel and in conjunction with that of gestures

The Mirror System Hypothesis (MSH) roots speech in communication-based manual activity. The mirror system would have expanded its roles in concert with other brain regions as the human brain evolved. MSH implies that a language or proto-language primarily based on manual gesture was used by hominids before spoken language. Throughout the book, many arguments are given to make this scenario highly plausible, but MSH appears to play down some important data concerning vocal communication. *First*, the presence of vocalization production neurons in the premotor cortex of macaque monkeys suggests that some early evolutionary pressures might have been critical for the acquisition of a better vocal voluntary control; *second*, the fact that vocal communication represents an important part of monkeys' and apes' communication suggests that a vocal processing system developed before or in conjunction with manual gesture.

Vocal production in nonhuman primates is considered to be controlled by the brainstem and by mesial cortical areas which, besides other functions, are also involved in emotional behavior (West and Larson 1995; Jürgens 2002). These characteristics are also underlined by Arbib in his book and represent a clear difference with the voluntary control of speech in humans, whose neural basis can be identified in the lateral motor and prefrontal cortex. This significant difference has led several scholars to deny the possible contribution of non-human primate vocalization to the emergence of speech. However, recent data allow us to suggest a possible role for monkey vocalization in language evolution. We recently conducted a study (Coudé et al. 2011) in which we recorded from ventral premotor cortex (PMv) of macaques trained to emit vocalizations (i.e. coo-calls). The results showed that the rostro-lateral part of PMv (mostly area F5; Figure 2a) contains neurons that fire during conditioned vocalization (Figure 2b). Interestingly, in the majority of these neurons the discharge started before sound onset (Figure 2c), thus suggesting their causal relation with vocal production. The involvement of these neurons in the motor control of vocalization has also been supported by electrical microstimulation of this sector, which in some cases and for the lateral-most penetration sites, elicited a combination of jaw and tongue movements together with larynx muscles contraction. Note that a control on larynx movements has been previously shown by a cortical surface electrical stimulation study of ventral premotor cortex (Hast et al. 1974).

Taken together, these results indicate that a partial voluntary vocal control was already taking place in the primate PMv cortex some 25 million years ago.

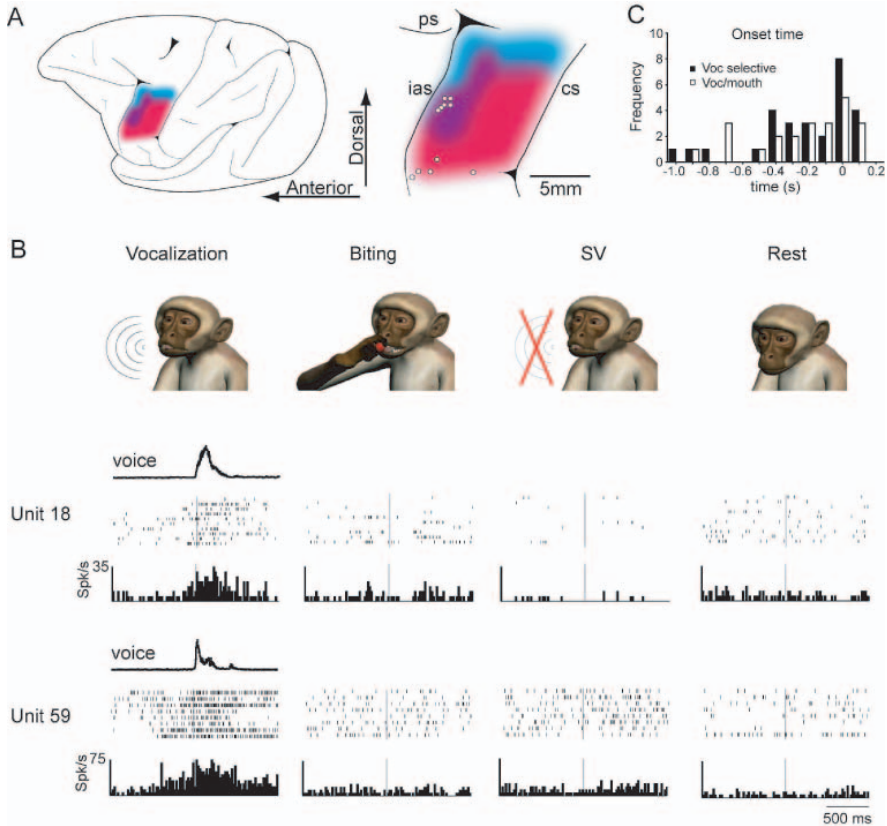


Fig. 2: Recorded region and vocalization-selective neurons. **A.** *Left*. Lateral view of the left hemisphere of one of the recorded monkeys. Colored sectors indicate hand (blue), mouth (red) and overlapping hand and mouth (purple) motor representations. *Right*. Enlarged view of the recorded area showing the position of electrode penetrations (white dots) where vocalization-selective neurons were found. cs = central sulcus, ias = inferior arcuate sulcus, ps = principal sulcus. **B.** Examples of two vocalization-selective neurons recorded during four different behaviors. For each unit, rasters and histograms illustrate the neuronal discharge aligned (vertical gray line) with behavioral events. They correspond to monkey sound emission onset during vocalization, contact with food during biting and maximum lips protrusion during silent vocalization (SV). During rest, the activity alignment corresponded to the midpoint of a period in which the monkey did not produce any movement. The voice trace is depicted for each unit above the raster of the vocalization-related neuronal discharge. **C.** Frequency of vocalization-selective (black bars) and vocalization/mouth related neurons (white bars) according to discharge time onset with respect to the beginning of sound emission. Modified from Coudé et al. (2011).

While acknowledging the involvement of PMv in vocal production, Arbib mentions that the behavior described in our study underscores the poor level of vocal control reached by the animals. The behavioral side of our study highlighted that vocalizing on command is achievable in macaques, but remains limited. In fact, about half of the trials consisted in failed attempts to vocalize in which the articulatory oro-facial gestures involved in the coos were made without sound emission. We referred this behaviour as to “silent vocalization”. However, the vocalization neurons we found (see Fig 2b) only discharged when the complete vocalization pattern was produced (i.e. respiration, phonation and articulation, see MacNeilage, 1998), while they were never active when the monkeys performed silent vocalization. This aspect emphasizes that a circuit at least partly dedicated to voluntary vocal control started to emerge in the common ancestors of old world monkeys and apes. These data are at odd with Arbib’s contention according to which the emergence of voice modulation and thus of an articulatory movement repertoire could have become associated with, or even prompted by, the preexisting manual action repertoire. In fact, from the standpoint of evolution, macaque data suggest a timescale for the emergence of vocal control such that some evolutionary pressure must have come into play well before the use of proto-signs (i.e. communication based on conventionalized manual gestures) developed.

If we look at our closest relatives, the apes, it has been reported that, even though they fail to produce words as a consequence of training (Hayes 1951; Kellogg and Kellogg 1967; Gardner and Gardner 1969), nonetheless they seem capable of voluntarily emitting utterances in order to gather attention or solicit interactions (Brosnan and De Waal 2001; Crockford et al. 2012; Hopkins et al. 2011; Leavens et al. 2010).

What appears to be a new acquisition in apes with respect to monkeys is the voluntary use of vocalizations in combination with brachio-manual gestures (Hostetter et al. 2001) to communicate. Although the repertoire of these vocalizations is limited, the use of arm/hand gestures is richer in terms of combinatorial possibilities and flexibility (Lyn et al. 2011; Gómez 2007; Leavens et al. 2005; Leavens et al. 2004). The behavioral coupling between vocalization and gestures suggests that at the neural level their motor control could be underpinned by common structures or by partially overlapping representations. This hypothesis seems to be confirmed by the few imaging studies available in chimpanzees. One of these studies demonstrated the activation of the homolog of human Broca’s area during the production of communicative vocal and hand gestures (Tagliabeta et al. 2008). A more recent PET study (Tagliabeta et al. 2011) showed that the same region is also involved in communicative oro-facial/vocal signaling. It is also interesting to note that this same region, together with precentral cortex, showed mirror properties during an execution/observation grasping task (Hecht

et al. 2011). These findings are hard to reconcile with an origin of language based only on the brachiomanual gestural communication system. Rather they point to a multimodal origin, with vocalization being already associated to it in the early phases of hominid evolution.

Acknowledgment: This work has been supported by the Italian Institute of Technology (RTM), the European Commission Grant Cogsystem (FP7-250013), the Italian PRIN n.2010MEFNF7_005 and the NIH Program P01 HD064653-01.

References

- Bonini L., F. Serventi, L. Simone, S. Rozzi, P. F. Ferrari & L. Fogassi. 2011. Grasping neurons of monkey parietal and premotor cortices encode action goals at distinct levels of abstraction during complex action sequences. *The Journal of Neuroscience* 31(15). 5876–5886.
- Bookheimer S. 2002. Functional MRI of language: new approaches to understanding the cortical organization of semantic processing. *Annual Review of Neuroscience* 25. 151–188.
- Brosnan S. F. & F. B. M. de Waal. 2001. Regulation of vocalizations by chimpanzees finding food in the presence or absence of an audience. *Evolution of Communication* 4(2). 211–224.
- Chersi F., P. F. Ferrari & L. Fogassi. 2011. Neuronal chains for actions in the parietal lobe: A computational model. *PLoS one* 6(11). e27652.
- Clerget E., A. Winderickx, L. Fadiga & E. Olivier. 2009. Role of Broca's area in encoding sequential human actions: a virtual lesion study. *Neuroreport* 20(16). 1496–1499.
- Coudé G., P. F. Ferrari, F. Rodà, M. Maranesi, E. Borelli, V. Veroni, F. Monti, S. Rozzi & L. Fogassi. 2011. Neurons controlling voluntary vocalization in the macaque ventral premotor cortex. *PLoS one* 6(11). e26822.
- Crockford C., R. M. Wittig, R. Mundry & K. Zuberbühler. 2012. Wild chimpanzees inform ignorant group members of danger. *Current Biology* 22(2). 142–146.
- Fazio P., A. Cantagallo, L. Craighero, A. D'Ausilio, A. C. Roy, T. Pozzo, F. Calzolari, E. Granieri & L. Fadiga. 2009. Encoding of human action in Broca's area. *Brain* 132(Pt 7). 1980–1988.
- Ferrari P. F., S. Rozzi & L. Fogassi. 2005. Mirror neurons responding to observation of actions made with tools in monkey ventral premotor cortex. *Journal of Cognitive Neuroscience* 17(2). 212–226.
- Fogassi L. & P. F. Ferrari. 2007. Mirror neurons and the evolution of embodied language. *Current Directions in Psychological Science* 16(3). 136–141.
- Fogassi L. & P. F. Ferrari. 2012. Cortical Motor organization, mirror neurons, and embodied language: An evolutionary perspective. *Biolinguistics* 6(3–4). 308–337.
- Fogassi L., P. F. Ferrari, B. Gesierich, S. Rozzi, F. Chersi & G. Rizzolatti. 2005. Parietal lobe: from action organization to intention understanding. *Science* 308(5722). 662–667.
- Gardner R. A. & B. T. Gardner. 1969. Teaching sign language to a chimpanzee. *Science* 165(3894). 664–672.
- Gómez J. C. 2007. Pointing behaviors in apes and human infants: A balanced interpretation. *Child Development* 78(3). 729–734.

- Hagoort P. 2005. On Broca, brain, and binding: A new framework. *Trends in Cognitive Sciences* 9(9). 416–423.
- Haller S., E. W. Radue, M. Erb, W. Grodd & T. Kircher. 2005. Overt sentence production in event-related fMRI. *Neuropsychologia* 43(5). 807–814.
- Hast M. H., J. M. Fischer, A. B. Wetzel & V. E. Thompson. 1974. Cortical motor representation of the laryngeal muscles in *Macaca mulatta*. *Brain Research* 73(2). 229–240.
- Hayes C. 1951. *The ape in our house*. New York: Harper.
- Hecht E., E. L. Davis & L. A. Parr. 2011. Do chimpanzees “mirror” others’ actions? A functional neuroimaging study of action execution and observation. *Society For Neuroscience Abstracts* 932.05.
- Hopkins W. D., J. P. Tagliatalata & D. A. Leavens. 2011. Do chimpanzees have voluntary control of their facial expressions and vocalizations? In A. Vilain, J. L. Schwartz, C. Abry & J. Vauclair (eds.), *Primate communication and human language*, 71–88. Amsterdam: John Benjamins.
- Hostetter A. B., M. Cantero & W. D. Hopkins. 2001. Differential use of vocal and gestural communication by chimpanzees (*Pan troglodytes*) in response to the attentional status of a human (*Homo sapiens*). *Journal of Comparative Psychology* 115(4). 337–343.
- Jürgens U. 2002. Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews* 26(2). 235–258.
- Kellogg W. N. & L. A. Kellogg. 1967. *The ape and the child*. New York: Hafner Publishing Company.
- Leavens D. A., W. D. Hopkins & K. A. Bard. 2005. Understanding the point of chimpanzee pointing: Epigenesis and ecological validity. *Current Directions in Psychological Science* 14(4). 185–189.
- Leavens D. A., W. D. Hopkins & R. K. Thomas. 2004. Referential communication by chimpanzees (*Pan troglodytes*). *Journal of Comparative Psychology* 118(1). 48–57.
- Leavens D. A., J. L. Russell & W. D. Hopkins. 2010. Multimodal communication by captive chimpanzees (*Pan troglodytes*). *Animal Cognition* 13(1). 33–40.
- Liebal K., J. Call & M. Tomasello. 2004. Use of gesture sequences in chimpanzees. *American Journal of Primatology* 64(4). 377–396.
- Lyn H., P. M. Greenfield, S. Savage-Rumbaugh, K. Gillespie-Lynch & W. D. Hopkins. 2011. Nonhuman primates do declare! A comparison of declarative symbol and gesture use in two children, two bonobos, and a chimpanzee. *Language & Communication* 31(1). 63–74.
- MacNeilage P. F. 1998. The frame/content theory of evolution of speech production. *The Behavioral and Brain Sciences* 21(4). 499–511.
- Pinker S. & R. Jackendoff. 2005. The faculty of language: What’s special about it? *Cognition* 95(2). 201–236.
- Pulvermüller F. & L. Fadiga. 2010. Active perception: Sensorimotor circuits as a cortical basis for language. *Nature Reviews. Neuroscience* 11(5). 351–360.
- Rizzolatti G. & L. Craighero. 2004. The mirror-neuron system. *Annual Review of Neuroscience* 27. 169–192.
- Rochat M. J., F. Caruana, A. Jezzini, L. Escola, I. Intskirveli, F. Grammont, V. Gallese, G. Rizzolatti & M. A. Umiltà. 2010. Responses of mirror neurons in area F5 to hand and tool grasping observation. *Experimental Brain Research* 204(4). 605–616.
- Sapolsky D., A. Bakkour, A. Negreira, P. Nalipinski, S. Weintraub, M. M. Mesulam, D. Caplan & B. C. Dickerson. 2010. Cortical neuroanatomic correlates of symptom severity in primary progressive aphasia. *Neurology* 75(4). 358–366.

- Tagliabata J. P., J. L. Russell, J. A. Schaeffer & W. D. Hopkins. 2008. Communicative signaling activates “Broca’s” homolog in chimpanzees. *Current Biology* 18(5). 343–348.
- Tagliabata J. P., J. L. Russell, J. A. Schaeffer & W. D. Hopkins. 2011. Chimpanzee vocal signaling points to a multimodal origin of human language. *PLoS one* 6(4). e18852.
- Tanji J. 2001. Sequential organization of multiple movements: involvement of cortical motor areas. *Annual Review of Neuroscience* 24. 631–651.
- Tanji J. & E. Hoshi. 2008. Role of the lateral prefrontal cortex in executive behavioral control. *Physiological Reviews* 88(1). 37–57.
- Umiltà M. A., L. Escola, I. Intskirveli, F. Grammont, M. Rochat, F. Caruana, A. Jezzini, V. Gallese & G. Rizzolatti. 2008. When pliers become fingers in the monkey motor system. *PNAS* 105(6). 2209–2213.
- West R. A. & C. R. Larson. 1995. Neurons of the anterior mesial cortex related to faciovocal activity in the awake monkey. *Journal of Neurophysiology* 74(5). 1856–1869.
- Wilson S. M., M. L. Henry, M. Besbris, J. M. Ogar, N. F. Dronkers, W. Jarrold, B. L. Miller & M. L. Gorno-Tempini. 2010. Connected speech production in three variants of primary progressive aphasia. *Brain* 133(Pt 7). 2069–2088.