



Gestures, vocalizations, and memory in language origins

Francisco Aboitiz *

Departamento de Psiquiatría, Facultad de Medicina y Centro Interdisciplinario de Neurociencia, Pontificia Universidad Católica de Chile, Santiago, Chile

Edited by:

Angela Dorkas Friederici, Max Planck Institute for Human Cognitive and Brain Sciences, Germany

Reviewed by:

Steven Chance, Oxford University, UK
Jonathan K. Burns, University of KwaZulu-Natal, South Africa

*Correspondence:

Francisco Aboitiz, Departamento de Psiquiatría, Facultad de Medicina y Centro Interdisciplinario de Neurociencia, Pontificia Universidad Católica de Chile, Avenue Marcoleta #391, Santiago, Chile.
e-mail: faboitiz@puc.cl

This article discusses the possible homologies between the human language networks and comparable auditory projection systems in the macaque brain, in an attempt to reconcile two existing views on language evolution: one that emphasizes hand control and gestures, and the other that emphasizes auditory–vocal mechanisms. The capacity for language is based on relatively well defined neural substrates whose rudiments have been traced in the non-human primate brain. At its core, this circuit constitutes an auditory–vocal sensorimotor circuit with two main components, a “ventral pathway” connecting anterior auditory regions with anterior ventrolateral prefrontal areas, and a “dorsal pathway” connecting auditory areas with parietal areas and with posterior ventrolateral prefrontal areas via the arcuate fasciculus and the superior longitudinal fasciculus. In humans, the dorsal circuit is especially important for phonological processing and phonological working memory, capacities that are critical for language acquisition and for complex syntax processing. In the macaque, the homolog of the dorsal circuit overlaps with an inferior parietal–premotor network for hand and gesture selection that is under voluntary control, while vocalizations are largely fixed and involuntary. The recruitment of the dorsal component for vocalization behavior in the human lineage, together with a direct cortical control of the subcortical vocalizing system, are proposed to represent a fundamental innovation in human evolution, generating an inflection point that permitted the explosion of vocal language and human communication. In this context, vocal communication and gesturing have a common history in primate communication.

Keywords: arcuate fasciculus, broca’s area, inferior parietal lobe, mirror neurons, phonological loop, superior longitudinal fasciculus, working memory

INTRODUCTION

In the last 15 years, there has been an increasing interest in understanding the evolutionary aspects of language and human communication. Several comparative analyses have been aimed at identifying a phylogenetic continuity between the brain networks involved in language processing in humans, and neural circuits present in the non-human primate. At least two lines of research have become particularly influential in this regard. One of them has focused on the search for auditory–premotor circuits in the macaque monkey, by assuming homology with the human’s language network based on cytoarchitectonic and connectivity criteria (Aboitiz and García, 1997; Petrides and Pandya, 2009). These findings are broadly consistent with those obtained through a comparative approach, which studies vocal learning in non-human species, particularly in songbirds, as both emphasize the development of auditory–vocal circuits as a crucial step in the acquisition of human language (Bolhuis et al., 2010; Berwick et al., 2011).

Another research program emerged somewhat unexpectedly from the study of grasping visuomotor neurons in the parietal and premotor cortex of the monkey, where the so-called “mirror neurons” were found to be activated both when executing an action and when observing this action (Di Pellegrino et al., 1992; Gallese et al., 1996; Rizzolatti et al., 1996). Based on these findings, Rizzolatti and Arbib (1998) developed the hypothesis that the grasping

mirror neuron system represented a scaffold from which language circuits emerged in the human. Mirror neurons are found in area F5 of the ventral premotor cortex, which has been proposed by some authors to be the homolog of Broca’s area in the human (Rizzolatti and Craighero, 2004).

These two approaches have largely been considered alternative possibilities, and there has been little cross-talk between authors supporting each view; in addition some misunderstanding of each other’s work has increased the difficulty of reaching some agreement or common view. While the gestural and mirror neuron perspective makes strong emphasis on the background conditions for the emergence of human language, it does not provide specific insights into how speech arose to become the predominant communication mode in our species. In this article I will discuss some of the evidence supporting both views, in order to propose an integrated perspective in which the evolution of human communication has been based on multimodal signals including facial, hand, and body gestures, together with vocalizations.

AN EARLY HYPOTHESIS OF MONKEY HOMOLOGIES ANCESTRY OF THE LANGUAGE CIRCUITS

Several years ago, we presented a hypothesis for the evolutionary origin of the language networks, based on the hodological evidence available for the monkey at that time (Aboitiz and García, 1997; see also Aboitiz, 1995). Homologs to human areas 44 and 45

(corresponding to Broca's region) in the ventrolateral prefrontal cortex (VLPFC) had been already identified in the chimpanzee by Brodmann, and there has been no further debate about their correspondence (Sherwood et al., 2003; Schenker et al., 2008; Keller et al., 2009). However, in the macaque the situation was more difficult. At that time, only area 45 could be identified in this species, inside the inferior arcuate sulcus, between subareas 6v and 8Ar (Preuss and Goldman-Rakic, 1991a). This region was viewed as a specialization of the premotor area 6v (area 6 ventralis), representing orofacial movements (Deacon, 1992; Preuss, 1995). On the other hand, area Tpt in the superior temporal lobe (which has been related to Wernicke's region by some authors) had been identified even in prosimians. This area was described as a multimodal zone receiving auditory and somatosensory projections from the temporal and parietal lobes, respectively (Galaburda and Sanides, 1980; Pandya and Yeterian, 1985; Preuss and Goldman-Rakic, 1991b).

Nonetheless, evidence for an arcuate fasciculus connecting Broca's and Wernicke's areas, was difficult to find in the monkey. Area Tpt was found to send projections to areas 8 and 46 of the prefrontal cortex, but not to area 45 or to the ventral arcuate sulcus (Petrides and Pandya, 1988). On the basis of these findings, area Tpt was proposed to play a role in head-turning movements aimed at localizing sound sources (Pandya and Yeterian, 1985). The only superior temporal projections to the ventral arcuate sulcus originated from the secondary auditory area ProA (Petrides and Pandya, 1988). However, another study at that time described some temporal projections from the superior temporal gyrus and the superior temporal sulcus (STS), to the inferior post-arcuate and the pre-arcuate region (Deacon, 1992). On the other hand, area 45 was found to receive projections from the anterior inferior parietal area 7b, which also sends afferents to area 46 in the dorsolateral prefrontal cortex (Petrides and Pandya, 1984; Preuss and Goldman-Rakic, 1991c; Seltzer and Pandya, 1994). Inside the intraparietal sulcus, area 7ip had been described as projecting to the dorsal and ventral aspects of the anterior arcuate sulcus (area 8), and to the posterior principal sulcus (area 46; Petrides and Pandya, 1984; Cavada and Goldman-Rakic, 1989; Preuss and Goldman-Rakic, 1991a,b,c). Pandya and Yeterian (1985) and Seltzer and Pandya (1978) described connections between the middle superior temporal lobe and somatosensory parietal regions via the middle longitudinal fasciculus (MLF), whereas Cavada and Goldman-Rakic (1989) reported projections from area 7ip to the posterior STS, and from areas 7b and 7a (the latter is posterior to 7b) to the superior bank of the STS. Finally, there is evidence for a projection from visual area TE in the inferior temporal lobe to the inferior arcuate sulcus, including area 45 (Bullier et al., 1996).

TRIPARTITE INPUT TO BROCA'S REGION

Based on these descriptions, we proposed – to our knowledge for the first time – a model for the organization of the language circuits and their possible homologies in the monkey, which emphasized a more complex network than did previous models. A significant component of this model was the inclusion of an inferior parietal projection to Broca's area and a connection between Wernicke's region and the inferior parietal lobe (Aboitiz and García, 1997). This schema implied three main inputs to Broca's area: (i) a direct

route running through the arcuate fasciculus; (ii) an indirect route from the posterior superior temporal lobe to the inferior parietal lobe, and from there to Broca's region; and (iii) projections from the anterior temporal lobe to Broca's area (although we proposed that these were mainly visual; Aboitiz and García, 1997). In human evolution, the direct projections from Wernicke's area to Broca's area via the arcuate fasciculus would have gained greater importance than in the monkey, in which no clear evidence for an arcuate fasciculus existed. Furthermore, at the time several imaging studies had shown a strong inferior parietal involvement in verbal working memory, especially in phonological storage tasks (for example, Paulesu et al., 1993; Awh et al., 1996; Salmon et al., 1996; see also Smith and Jonides, 1998), which was consistent with an inferior parietal input to Broca's region. These findings were in line with our hypothesis that working memory, particularly phonological working memory, was important for language acquisition in children (Baddeley et al., 1988) and also in early humans. Nonetheless, following Fuster (1995), we also argued strongly that rather than there being specific memory-dedicated regions, short-term memory should be considered a property of the whole network involved in sensorimotor integration, which interacted intensively with other associated networks (Aboitiz and García, 1997; see also Aboitiz et al., 2006a,b, 2010). In summary, we proposed that an expansion of working memory capacity was critically associated with the differentiation of the above mentioned language circuits, in the context of learning and processing complex phonological sequences that were acquired by imitation of conspecifics (Aboitiz and García, 1997).

NEW EVIDENCE

AUDITORY PROJECTIONS IN NON-HUMAN PRIMATES

After our original publication, there has been a wealth of new evidence on the existence of temporal–parietal–prefrontal connections, both in the monkey and in the human. In the macaque, auditory projections separate into a dorsal and a ventral stream, running to the parietal lobe and to the anterior temporal lobe, respectively. This arrangement has been viewed as being analogous to the bipartite arrangement of the visual system, in which the dorsal pathway is involved in spatiotemporal signal processing and is related to eye movement control, whereas the ventral pathway relates to stimulus recognition and emotional processing (Kaas and Hackett, 1999).

In the macaque, the dorsal auditory pathway originates in posterior auditory areas in the superior temporal lobe and is directed mainly to dorsal prefrontal areas (areas 8 and 46, related to eye movement control). It is noteworthy that this pathway does not fit in an obvious manner into the language network, as it terminates preferentially in dorsal rather than ventral prefrontal regions (in the human, a dorsal prefrontal projection of the dorsal pathway has been also described; see Frey et al., 2008). On the other hand, the ventral stream originates in different areas of the anterior and middle temporal gyrus, and conveys visual and auditory inputs directed mainly to areas 12 and 45 of the VLPFC (Kaas and Hackett, 1999; Romanski et al., 1999a,b; Belin and Zatorre, 2000; Rauschecker and Tian, 2000; Romanski, 2007). Consistent with this evidence, other reports described an auditory domain in the macaque inferior frontal areas 12 and 45, in

which vocalization-specific neurons were interspersed with facial-sensitive neurons, allowing for the integration of vocal auditory stimuli with the corresponding facial gestures (Romanski and Goldman-Rakic, 2002; Romanski et al., 2005; Romanski, 2007). Interestingly, this region was found to receive afferents from the anterior lateral belt auditory area (Rauschecker and Tian, 2000; Tian et al., 2001), which is preferentially activated by calls from conspecifics (Petkov et al., 2008).

INFERIOR PARIETAL PROJECTIONS TO THE VLPFC OF THE MACAQUE

Petrides and Pandya (1999, 2002) subdivided the monkey area 45 into areas 45A and 45B, and identified a dysgranular area 44 in the depth of the inferior arcuate sulcus (Petrides et al., 2005). Stimulation of neurons in area 44 triggered orofacial movements and sometimes hand movements, but not ocular movements; oculomotor responses occurred only when area 8Av was stimulated, far from the 44–8Av border (Petrides et al., 2005). Furthermore, stimulation sites in the most ventral aspect of area 8Av and in the 45–8Av border did not elicit any motor response.

Before discussing in more detail the different reports and interpretations on connectivity of the inferior parietal lobe, it must be noted that the cytoarchitectonic parcellation of this region has not been consistent across studies. Whereas earlier studies adopted Brodmann's early description of area 7, subdividing it into areas 7b and 7a, and area 7ip inside the intraparietal sulcus (Petrides and Pandya, 1984; Preuss and Goldman-Rakic, 1991a,b,c), more recent studies have used another parcellation scheme, defining area PF anteriorly (Brodmann's area 40, anterior supramarginal gyrus in the human), area PFG in the middle (area 39, posterior supramarginal gyrus in the human), and area PG posteriorly (area 39, angular gyrus in the human), with area anterior intraparietalis (AIP) inside the intraparietal sulcus (see Matelli et al., 1986; Frey et al., 2008; Petrides and Pandya, 2009; Gerbella et al., 2011).

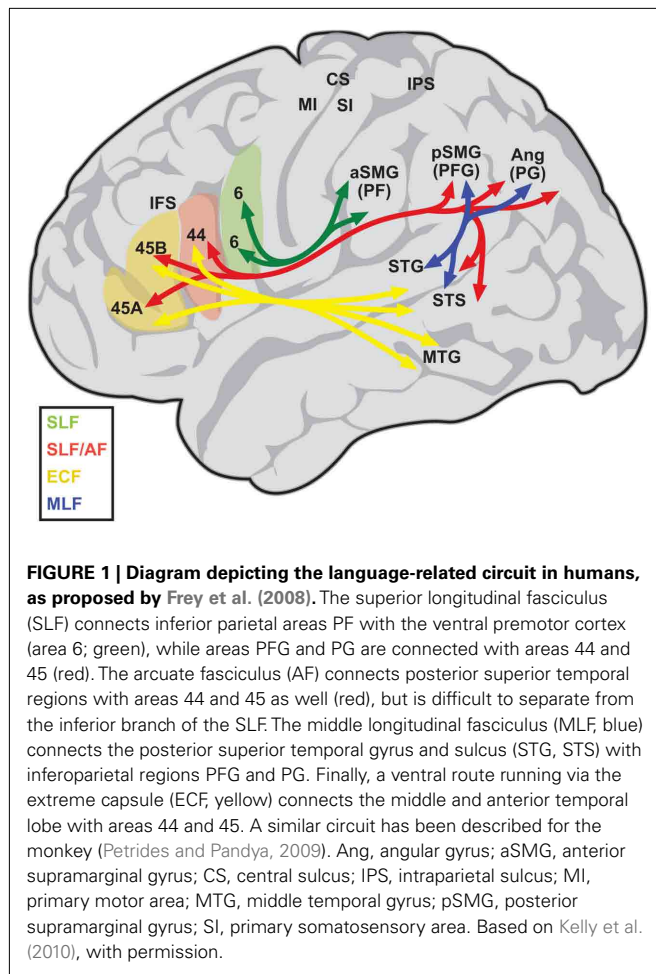
Petrides and Pandya (1999, 2002) described area 45 as being connected with the posterior inferior parietal lobe, while area 44 was viewed as receiving projections from the intraparietal and anterior inferior parietal lobe. Subsequently, Petrides and Pandya (2009) visualized a pattern of multiple afferents into areas 45A, 45B, and 44 from the inferior parietal and temporal regions. These projections consist of two main pathways: (i) axons running along the SLF into both areas 45 and 44, and which originate in the inferior parietal lobe (areas PFG and PG). Area PFG made a particularly strong projection into area 44. In addition, they described some axons from the ventral most inferior parietal lobe and the caudal STS, which formed an arcuate fasciculus, although this projection is not as prominent as it is in humans. Furthermore, there was a systematic relation between inferior parietal regions and the prefrontal regions to which they connected, with more rostral parts (area PF) connecting with the ventral premotor cortex (area 6 ventralis, controlling facial musculature), while intermediate regions of the inferior parietal lobe (area PFG) connected to area 44 and to a lesser extent to area 45. The second pathway (ii) consisted of multimodal axons running via the extreme capsule and uncinate fasciculus, originating in diverse auditory and visual cortical areas of the anterior and middle temporal lobe, and ending mainly in areas 45 and 47/12, but also to some extent in area 44. These authors argue that, in both the monkey and

in the human, the ventral projection to the VLPFC has a role in the mechanisms of memory retrieval, while the dorsal route (arcuate and superior longitudinal fasciculi) is involved in the control of vocal articulation only in humans (see also Saur et al., 2008).

TRACTOGRAPHIC STUDIES IN THE HUMAN BRAIN

Likewise, the advent of tractographic techniques in the living human yielded results consistent with the tripartite projection from the auditory regions into Broca's area that we originally described, with some modifications (Catani and ffytche, 2005; Parker et al., 2005; Friederici et al., 2006; Anwander et al., 2007; Frey et al., 2008; Glasser and Rilling, 2008; Friederici, 2009). Glasser and Rilling (2008) reported a two-component arcuate fasciculus in the left hemisphere, one connecting the superior temporal gyrus with areas 6 and 44, which according to them subserves phonological information; and the other connecting the middle temporal gyrus with areas 9, 44, and 45, and proposed to be involved in lexical-semantic aspects. In the right hemisphere they visualized a less prominent fasciculus, connecting the middle temporal gyrus with areas 6 and 44, which was proposed to convey prosodic information. They also reported a very small tract connecting the superior temporal lobe with areas 6 and 44 in the right hemisphere. Likewise, Parker et al. (2005) reported a strong asymmetry in the arcuate fasciculus, favoring the left hemisphere; a similar asymmetry was present in infants 1–4 months of age (Dubois et al., 2009). In a subsequent article, Rilling et al. (2008) visualized a progressive development of the arcuate fasciculus from human to macaque to the chimpanzee, while the ventral pathway, via the anterior temporal lobe, has remained more conservative during evolution (Rilling et al., 2008).

Frey et al. (2008) also described an arcuate fasciculus running from the posterior superior temporal gyrus to area 44 (in some cases to area 45; **Figure 1**). Note that the human arcuate fasciculus also projects to dorsal prefrontal areas 8 and 6, as in the monkey. However, the main focus of this report was on the inferior parietal and anterior temporal lobe projections to Broca's area. They found (i) a projection from the inferior parietal lobe (supramarginal gyrus) into area 44 via the SLF (in 10 of 12 subjects). The ventral posterior intraparietal region is claimed to receive auditory afferents from the superior temporal lobe via the middle and inferior longitudinal fasciculi, which might close a circuit from the posterior auditory cortex to area 44. In addition, they reported (ii) a ventral pathway connecting the anterior temporal areas with areas 47 and 45 via the extreme capsule and the uncinate fasciculus, as occurs in the monkey. This pathway has been described in other reports, and participates in the recognition of auditory stimuli including speech, identifying the speaker, mapping sound with meaning, verbal retrieval, echoic memory, and in simple grammatical processing (Buchsbbaum et al., 2005a,b; Hickok and Poeppel, 2007; Saur et al., 2008). According to Hickok and Poeppel (2007) the ventral pathway, unlike the dorsal stream, is represented bilaterally, being functional in both the left and right hemispheres. Recent evidence has shown that, instead of being separate pathways, the dorsal and the ventral streams operate synergistically during language processing (Lopez-Barroso et al., 2011; Rolheiser et al., 2011).



Despite the attractiveness of these studies, it is not entirely clear to what extent the tractographic evidence reveals a monosynaptic arcuate fasciculus, first because this technique lacks the resolution required to strongly confirm this possibility, and second, because this tract is difficult to separate from the adjacent SLF, which carries fibers from the parietal lobe. It must be also mentioned that Bernal and Altman (2010) were unable to find a strong projection from the SLF or the arcuate fasciculus into Broca's region, finding instead a strong termination of these tracts in the ventral premotor and motor cortices. However, in this study the parietal site of origin of the SLF was insufficiently characterized, which undermines somewhat the authors' main conclusions.

THE INFERIOR PARIETAL LOBE AND VERBAL WORKING MEMORY

There is now an important discussion about the role of different inferior parietal areas in verbal working memory, as the concept of a memory-dedicated, anatomically isolated component has been challenged by many studies (Hickok and Poeppel, 2007; Buchsbaum and D'Esposito, 2008; Hickok, 2009; see also Aboitiz et al., 2006a, 2010). More importantly, the only regions that have consistently shown sustained activation during verbal working memory tasks are the STS and the mid-superior temporal gyrus, especially an area located in the posterior planum temporale (area Spt; its

relation to area Tpt is not yet clear; Buchsbaum et al., 2005a,b; Hickok and Poeppel, 2007; Hickok, 2009). In particular, area Spt is proposed to serve as an interface between sensory and motor representations during the maintenance of phonological items on line (Buchsbaum and D'Esposito, 2008; Buchsbaum et al., 2011). In this interpretation, the "phonological loop" is considered to include a sensory phonological processing system partly represented (bilaterally) by the STS, a sensory–motor integration system in the left Spt area, and a left frontal articulatory system (Hickok, 2009). These authors interpret the role of the inferior parietal lobe as serving some higher-order functions that support verbal working memory. One possibility is that these regions participate in motor planning mechanisms that help stabilize perceptual memory traces (see below).

THE PHONOLOGICAL LOOP: A KEY INNOVATION PHONOLOGICAL CIRCUITS AND WORKING MEMORY

In subsequent reports, we emphasized the role of phonological working memory, supported by the development of a phonological sensory motor circuit (the phonological loop) in early humans, as a crucial element in early language and human evolution (Aboitiz et al., 2005, 2010). The ability to rehearse and keep newly learned phonological sequences in short-term memory became an inflection point that changed human sociality forever, being a fundamental factor in the evolution of complex language and culture. This "inner speech" capacity also allowed the elaboration of new and more complex messages by manipulating the phonemes being learned. Furthermore, we argued that this circuit was largely, although not exclusively, based on the strengthening of the dorsal pathway connecting Wernicke's and Broca's areas (including the direct projection via the arcuate fasciculus and the indirect one via the inferior parietal lobe), while the ventral pathway running via the extreme or external capsule had been more conservative in evolution (although not static; see below), being the dominant pathway for vocalization processing in the monkey (Aboitiz et al., 2006a, 2010).

SYNTAX

We also claimed that the origin of a complex syntax based on long-distance dependencies between linguistic elements required a robust phonological memory system in order to maintain the different items active while other elements were being processed online (Aboitiz et al., 2006a, 2010). Imaging studies indicate a participation of Broca's area in working memory processes associated with syntactical processing (Fiebach et al., 2002, 2005), and Friederici (2004) has argued that syntactic working memory involves the superior anterior portion of area 44, while syntactic processing relates to the inferior portion of left area 44. Furthermore, the dorsal pathway for language is involved in the processing of grammatical structures organized in a hierarchical manner, whereas a "middle pathway," similar to the ventral pathway but ending in area 45A, participates in the analysis of simple grammatical structures (Friederici et al., 2006; Anwander et al., 2007; see also Wilson et al., 2011).

In the adult, syntactical processing is probably automatic to a large extent – especially simple grammatical forms – and may depend on cortico-striatal circuits involved in procedural memory

(Ullman, 2004). Along the same line, patients with lesions in the left temporo-parietal cortex that have specific short-term memory deficits for numbers and words, do not display any major impairments in their spontaneous speech, supporting the participation of subcortical components in automatic language processing (Shallice and Warrington, 1970; Saffran and Marin, 1975). However, the initial acquisition of rules, the processing of complex syntactic forms (Friederici, 2005), and the online maintenance of semantic information during linguistic processing, require short-term memory mechanisms that bridge these procedural components with episodic memory networks (Reuland, 2010).

HANDEDNESS, GESTURES, AND MIRROR NEURONS BRAIN ASYMMETRIES FOR LANGUAGE AND HAND CONTROL

A critical issue in the context of language evolution is the conspicuous left hemispheric specialization for linguistic functions in most people, which is consistent with the evidence of gross-anatomical asymmetries in language-related regions (Ide et al., 1999; Josse and Tzourio-Mazoyer, 2004). Interestingly, asymmetry for language is correlated with some lateralized capacities like handedness, but not with other asymmetric capacities such as spatial attention (Badzakova-Trajkov et al., 2010). Apes tend to be right-handed, and there is evidence suggesting that Neanderthals were predominantly right-handed as well (Lonsdorf and Hopkins, 2005; Steele and Uomini, 2009). Notably, in chimpanzees, there is a correlation between throwing capacity, communicative ability and, the white-to-gray matter ratio in the homolog of Broca's area (Hopkins et al., 2012).

Thus, handedness, throwing ability, and rhythmic hammering have been related to language origins, which is in line with the mirror neuron hypothesis (see below; Calvin, 1983; Corballis, 2003). Along this line, several authors have made emphasis on gestural and manual communication as a first step in the acquisition of language (Hewes, 1973; Corballis, 1992; Armstrong et al., 1995; Kendon, 2004). More specifically, Corballis (1992, 2002) originally proposed that generativity, a key syntactic operation, was initially present in a system of manual gestures, but switched to a predominantly vocal system in modern humans. Corballis included evidences from different fields of comparative cognition and the mirror neuron literature in his hypotheses, which strictly imply a stage of predominantly manual communication before vocal language took over (Corballis, 2002, 2003, 2010).

A more general, but not alternative, interpretation of hemispheric dominance for language is that complex sequential motor patterns may be more efficiently programmed in one hemisphere than in two. This fits with comparative evidence of lateralization for song production in songbirds (Bolhuis et al., 2010). Other authors have proposed that differences in interhemispheric communication via the corpus callosum may have played a role in the origin or maintenance of human brain lateralization (Ringo et al., 1994; Aboitiz et al., 2003; Häberling et al., 2011).

THE DISCOVERY OF MIRROR NEURONS

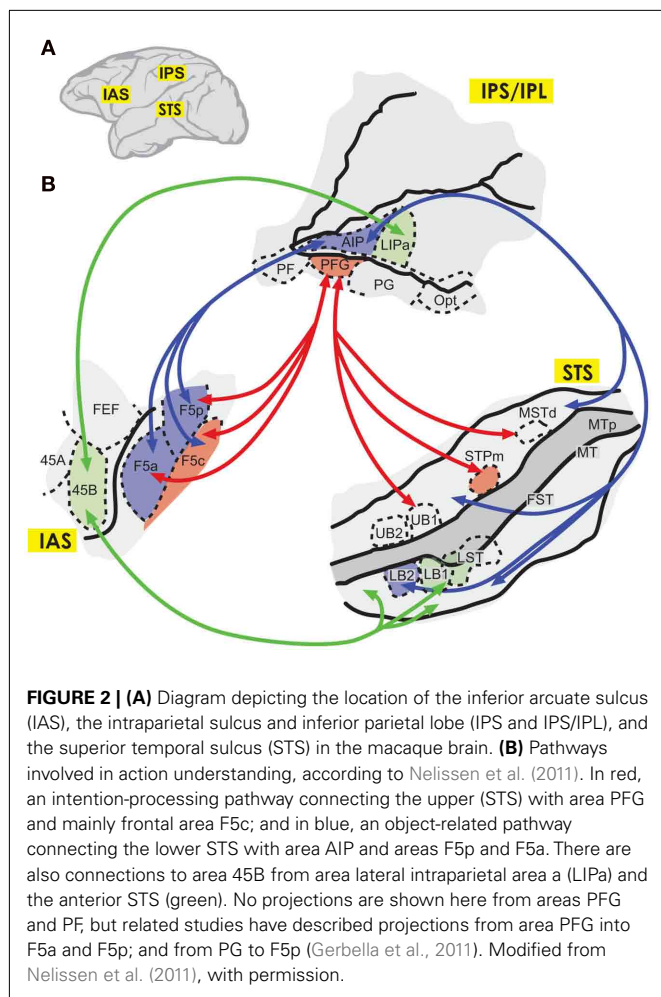
While studying the neurophysiology of visuomotor neurons involved in hand-grasping control in the monkey, Di Pellegrino et al. (1992) observed a group of motor neurons, termed "mirror neurons" that also became active when the animal observed

meaningful hand movements made by the experimenters or by another animal (see also Rizzolatti and Luppino, 2001; Rizzolatti and Craighero, 2004). Most mirror neurons were initially observed in the premotor area F5 (Brodmann's area 6v), located in the precentral gyrus and adjacent to the inferior arcuate sulcus (Belmalih et al., 2009). Area F5 has been subdivided into areas F5p, F5c, and F5a. Area F5a, which is adjacent to area 44 in the inferior arcuate sulcus (see below) has been proposed to be an integration site for parietal sensory-motor signals with signals from prefrontal and premotor areas (Gerbella et al., 2011). Of note, mirroring properties were also observed in face-selective neurons of the lateral aspect of F5, possibly allowing the animal to recognize gestures produced by conspecifics. Many of these responded to feeding behaviors, but some also fired when the animal observed a communicative gesture like a lip smacking (Ferrari et al., 2003; Rizzolatti and Craighero, 2004). Furthermore, some mirror neurons were found to fire not only in response to an observed action, but to action-related sounds, even in the absence of the visual presentation of the action (Keysers et al., 2003).

Mirror neurons have also been described in the rostral inferior parietal area, firing both to the observation of actions and to the execution of these or similar actions (Fogassi et al., 1998; Gallese et al., 2002). In the STS (which is connected with the inferior parietal region), there are sensory neurons selective for body actions rather than to grasping, although some of them also fire with the observation of goal-directed hand movements (Perrett et al., 1990). As seen by fMRI in the monkey, observation of grasping actions produces activations in inferior frontal areas F5, 45B, 45A, and 46; and on parietal areas PFG and AIP, plus the STS (Nelissen et al., 2005, 2011). These authors suggest that there are two pathways involved in the observation of actions, one running from the upper STS, relaying in area PFG, and projecting to the premotor area F5c which processes the agent's intentions (a context-dependent representation of the action); and the other, that originates in the lower STS, projects to area F5a/p via AIP and is more focused on the object (Figure 2; see also Luppino et al., 1999). They also describe connections of the STS and the lateral intraparietal area with area 45B. In the monkey, area PF was found to project strongly to F5a, F5c, and F5p; area PFG directed its axons to area F5a and F5p, while area PG was mainly directed to area F5p (Matelli et al., 1986; Gerbella et al., 2011). In other words, area F5a receives a robust input from areas PFG and AIP, a weaker input from PF, and practically no input from area PG (Gerbella et al., 2011). In addition, AIP projects to premotor area F5 (representing mostly the hand and mouth), while the ventral intraparietal sulcus is connected with the more dorsal premotor area F4 (representing the arm, neck, and face; Luppino et al., 1999). Note that this pattern is different from that described in the macaque by Petrides and Pandya (2009), who emphasize inferior parietal projections into the more anterior VLPFC.

A MIRROR SYSTEM IN HUMANS?

For obvious reasons, mirror neurons have been difficult to report in humans (Rizzolatti and Craighero, 2004). Nonetheless, there is a wealth of stimulation, electroencephalographic, and imaging data that is consistent with the notion that a mirror neuron system, i.e., a network involved in action recognition, imitation, and empathic



behavior, is present in the human (Rizzolatti and Craighero, 2004; Iacoboni and D’Apreto, 2006). However, there is discussion about whether this activity reflects or not the activity of mirror neurons as described in the monkey, and whether the human mirror neuron system does actually participate in language processing (Molenberghs et al., 2009; de Zubicaray et al., 2010). Below I will address some of the main findings of this research program, which nonetheless bears relevance to the issue of language and gesture interaction.

Unlike the monkey, humans show mirror-system activity with the observation of meaningless, not object-directed movements, and with pantomimes, which may be attributed to communication skills (Fadiga et al., 1995; Buccino et al., 2001; Maeda et al., 2002; Grèzes et al., 2003). In humans, the localization of mirror system activity encompasses a wide bilateral cortical network, including parietotemporal visual regions, the rostral inferior parietal lobe, and the inferior precentral and frontal gyri (Iacoboni and D’Apreto, 2006). More recent proposals also emphasize the participation of a ventral pathway running via the anterior temporal lobe, as an additional component involved in planning, decision making (Arbib, 2010), and in the prediction of the intentions and the goals of actions (Kilner, 2011).

In humans, the mirror system has been interpreted as participating in action understanding, which is critical for inferring

another’s intentions in a social context (Rizzolatti and Craighero, 2004). There are many studies that have reported an activation in Broca’s region during real and imagined hand movements (Binkofski et al., 1999; Iacoboni et al., 1999; Gerardin et al., 2000). Furthermore, activation of area 44 with object-related mouth movements and imitation of vocal gestures has been reported in several studies (Di Pellegrino et al., 1992; Buccino et al., 2001). In addition, the pars triangularis, corresponding to area 45, displays mirror activity with the observation of behavioral goals rather than with the action itself (Johnson Frey et al., 2003). Finally, the mirror system has been shown to be involved in imitation tasks (Iacoboni and D’Apreto, 2006). During a finger imitation task in humans, Iacoboni et al. (1999) found a specific activation of the left pars opercularis (area 44), while in a task requiring the learning of a motor sequence, the activated areas included the pars opercularis, ventral premotor area, and the STS (Buccino et al., 2004; Vogt et al., 2007).

MIRROR NEURONS AND LANGUAGE CIRCUITS

MIRROR NEURONS AS A REQUIREMENT FOR LANGUAGE

On the basis of these and other findings, Rizzolatti and Arbib (1998) and Arbib (2005) proposed the bold hypothesis that the neural circuits involved in language processing evolved as an elaboration of the mirror neuron circuitry present in monkeys, which provided a scaffolding for the elaboration of a more complex, phonological network involved in communication and eventually, in speech. Furthermore, and as we originally claimed (Aboitiz and García, 1997), imitation is a key element in learned communication, and mirror neurons provide an adequate neural substrate for its implementation (even if monkeys are not good imitators). More specifically, Arbib (2005) proposed a sequence of events starting with an imitation system for grasping, which developed into a complex gestural communication system in which pantomime came to be used as a conventionalized reference system. Afterward, a “protosign” stage that used hand symbols would have occurred that eventually incorporated vocal sounds, or “protospeech.” Likewise, words resembling or suggesting ingestive behavior were proposed to be particularly important for the origin of a primitive semantics (Ferrari et al., 2003; Rizzolatti and Craighero, 2004). More recently, Arbib (2010) proposed that the ventral pathway for actions may have been particularly relevant for the acquisition of a primitive semantics, as this may have evolved to support words-as-phonological-actions, with semantics provided by the linkage to neural systems supporting perceptual and motor schemas. This view is consistent with the current understanding of the ventral pathway as being involved in the transformation of sound into meaning (Buchsbbaum et al., 2005a).

WHERE IS THE MONKEY HOMOLOG OF BROCA’S AREA?

Initially, proposers of the mirror neuron hypothesis identified area F5 as the most likely homolog of the human Broca’s area (Rizzolatti and Craighero, 2004). More recently, Gerbella et al. (2007, 2010) confirmed Petrides and Pandya’s (2002, 2009) descriptions of the monkey VLPCF, but emphasized connectivity of area 45A and 45B with oculomotor regions. They found only weak connections between area 45A and the inferior parietal areas PFG and PG, and between area 45B and the lateral intraparietal area. Thus, area 45A

might be associated with eye movement control during communication, while area 45B would instead belong to the monkey pre-arcuate region, involved with other oculomotor processes. Instead, Petrides and Pandya (2009) assert that the presumed part of area 45 that has been linked to oculomotor function, cytoarchitectonically corresponds to the caudal oculomotor area 8. Gerbella et al. (2007, 2010) also confirmed the existence of area 44 in the monkey, and considered it as an anterior subdivision of area F5 (or area 6v), adjacent to area F5a (Belmalih et al., 2009; **Figure 3**). According to these authors, in their original description Petrides and Pandya (2002) described area 44 with a more posterior extension, overlapping with the anterior premotor area (F5), but they subsequently restricted the limits of this area to the actual fundus of the inferior arcuate sulcus (Petrides et al., 2005). This points to the concept of area 44 as a specialization of the ventral premotor area.

DISCUSSION

Much of the disagreement between scholars invoking auditory–vocal vs. hand-based ancestral circuits for language can be separated into two main issues: one concerns the correspondence in primates of the language-related circuits in humans, which focuses on identifying the cortical area ancestral to Broca's region in the monkey; and the other refers to the possibility that a specific hand-gestural communication system preceded the advent of speech, and on the likelihood that a hand-based mirror neuron system represents a critical scaffolding for the subsequent evolution of language.

HOMOLOGY ISSUES

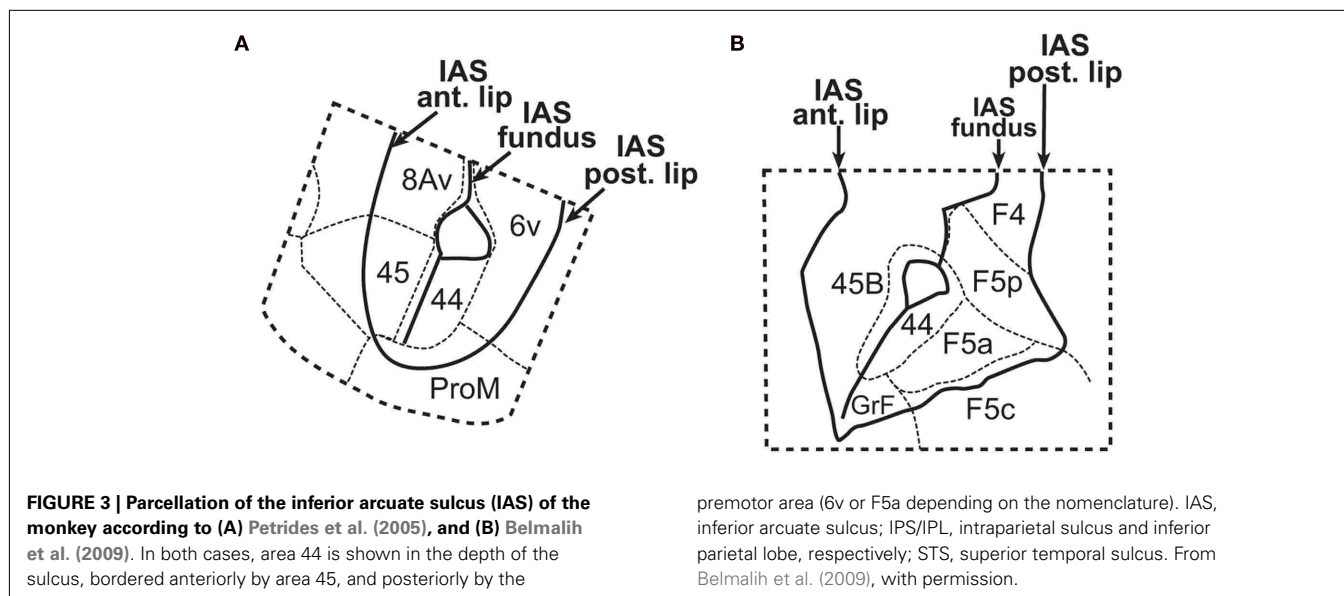
The unequivocal identification of areas 44 and 45 in the macaque, in the chimpanzee, and in the human, with practically identical topographies and cytoarchitectonic features suggests that those areas are most likely homologous to each other, deriving from the same germinal field in the embryonic telencephalon. On the other hand, classical grasping or mouth mirror neurons have been located in the premotor area F5 (area 6 ventralis), near the border

with area 44, in which there is yet no evidence of mirror neurons. However, it would be interesting to revisit the location of orofacial mirror neurons according to this cytoarchitectonic scheme (Ferrari et al., 2003), as in the monkey, stimulation of area 44 has been shown to elicit oral movements, either during communication or in feeding (Petrides et al., 2005).

Area 45 (and the adjacent area 12) fits the prefrontal auditory domain, receiving multimodal projections from the mid- and anterior-temporal lobe (Romanski, 2007; Petrides and Pandya, 2009; Gerbella et al., 2010). While Petrides and Pandya (2009) claim that this area participates in memory (semantic) retrieval processes, Belmalih et al. (2009) argue for a role in communication-directed eye movements, especially area 45B (see also Leichnetz, 2001). Although these discrepancies need to be resolved, it may safely be stated that the multimodal arrangement of area 45 corresponds to an auditory–motor interface that may be the evolutionary precursor of a speech-specialized region.

IS THE ARCUATE FASCICULUS THERE?

In the human left hemisphere, Glasser and Rilling (2008) described a tract connecting the superior temporal gyrus with areas 6 and 44 (involved in phonological processing), and a more robust one connecting more inferior temporal areas with areas 44, 45, and 9 (involved in semantic and lexical aspects). Frey et al. (2008) in the human, and Petrides and Pandya (2009) in the monkey, identified an arcuate fasciculus originating in the STS or in the inferior most parietal lobe, directed to the VLPFC and dorsal prefrontal areas. In the monkey, Yeterian et al. (2012) have recently made a claim for the existence of a direct projection between area Tpt in the superior temporal lobe and areas 44, 45, and dorsal prefrontal areas via the arcuate fasciculus. Furthermore, recent studies of verbal working memory point to the superior temporal gyrus (area Spt) as a key element involved in phonological sensorimotor integration (Buchsbaum and D'Esposito, 2008), which may perhaps contribute fibers to the arcuate fasciculus. Thus, the arcuate fasciculus is possibly an element involved in



auditory–vocal coordination and articulatory control, and might be involved in working memory processes by maintaining the functional connectivity between sensory and motor regions while holding items online. Nonetheless, tractographic studies to date are still insufficient to determine to what extent this is a mono-synaptic pathway, different from the SLF, connecting posterior temporal and VLPFC regions. In the monkey, chemical tracing studies suggest that if it is present, it is rather small (Petrides and Pandya, 2009).

THE INFERIOR PARIETAL CONNECTION

According to several studies in monkeys, areas 45 and 44 receive strong or moderate afferences from the inferior parietal lobe (Mesulam et al., 1977; Petrides and Pandya, 1984, 1999, 2002, 2009; Cavada and Goldman-Rakic, 1989; Preuss and Goldman-Rakic, 1991c; Leichnetz, 2001). However, in other studies only minor inferior parietal and intraparietal projections were described into areas 45A and 45B, respectively (Belmalih et al., 2009; Gerbella et al., 2010). More posteriorly, the premotor area 6v (or F5, where mirror neurons have been detected) receives strong projections from inferior parietal and intraparietal areas (Petrides and Pandya, 2009; Gerbella et al., 2011; Gharbawie et al., 2011). In humans, an inferoparietal projection to areas 44 and 45 has been described in several tractography studies (Catani and ffytche, 2005; Parker et al., 2005; Friederici et al., 2006; Anwander et al., 2007; Frey et al., 2008; Friederici, 2009; but see Bernal and Altman, 2010). Additional and more extensive connectivity studies are needed to determine the exact pattern of inferior parietal–prefrontal projections in the monkey and in the human.

An additional pathway involved in this circuit consists of a projection via the MLF to the inferior parietal lobe and intraparietal sulcus, which originates in the superior temporal lobe and STS. For some authors, this projection carries auditory information (Keyser et al., 2003; Frey et al., 2008; Petrides and Pandya, 2009), whereas others consider it as conveying body and arm positional information (Luppino et al., 1999; Nelissen et al., 2011). Both interpretations are not necessarily exclusive, as this projection likely transmits a multimodal input to the inferior parietal lobe.

A participation of inferior parietal regions in language circuits has been acknowledged by several researchers, although the precise role of these areas has yet to be resolved (Buchsbaum and D'Esposito, 2008; see above). Some authors have proposed a relation to phonological processing (Moser et al., 2009; Hartwigsen et al., 2010; Turkeltaub and Coslett, 2010), while others propose a role restricted to the sensorimotor control of writing (Brownsett and Wise, 2010). According to some authors, the inferior parietal lobe participates at an interface between speech audition and the articulatory code (Hickok, 2009; Moser et al., 2009). As mentioned above, one possibility is that these circuits maintain the stability of phonological sensorimotor circuits by codifying motor plans that contribute to maintain a behavioral goal during a working memory task. Along this line, the inferior parietal cortex of primates, and its projections into the frontal cortex, codify a diversity of orienting and object-directed behaviors, and have been proposed to participate in the selection of appropriate actions among competing circuits (Gharbawie et al., 2011; Kaas et al., 2011). At some

point in hominid evolution, these projections may have come to receive an increasingly stronger auditory input (especially phonological information from the STS), via the MLF, recruiting regions that were involved in face and especially mouth control to process vocalization information, and to perform action selection based on auditory input. It is very likely that this transition was concomitant with the elaboration of the direct cortical control over the hypoglossal motoneurons involved in vocalization (Jürgens and Alipour, 2002), thus closing a sensorimotor pathway from the acoustic system to the phonatory effectors. The recent finding of neurons controlling voluntary vocalizations in the ventral premotor cortex of the macaque is of great interest in this context, as it indicates that this rudimentary circuit was present at very early stages, possibly overlapping with other voluntary control systems (Coudé et al., 2011).

Finally in this section, the ventral pathway from the anterior temporal lobe to the anterior VLPFC (areas 45 and 47/12) has apparently suffered less structural changes in the lineage leading to humans (Rilling et al., 2008), which is consistent with our early hypothesis that in monkeys the ventral pathway is the dominant circuit involved in auditory–vocal integration (Aboitiz et al., 2006a). Nonetheless, in the human this pathway has been proposed to contribute to the processing of semantic, echoic information, and simple grammatical forms (Buchsbaum et al., 2005a; Anwander et al., 2007; Frey et al., 2008), indicating that it has suffered important modifications as well. It is also important to note that, despite being organized in several parallel streams, like the visual pathways, language processing operates in an integrated dynamics, in which all these streams converge on the common bottleneck of Broca's area, and very likely there is cross-pathway communication along the different functional routes (Rolheiser et al., 2011).

A MULTIMODAL COMMUNICATION SYSTEM

As discussed above, there is a confluence of facial gesture and vocalization information in the VLPFC of the monkey, mostly carried by the ventral visual and auditory pathways. Associations between the vocalization-sensitive region described in areas 45 and 47/12 (Romanski, 2007), the facial gesture-coding area 44 (Petrides et al., 2005), and the hand and body representations in the premotor area F5 are supported by neuroanatomy and make it plausible to visualize an integrated processing of hand and face gestures and vocalization patterns. Chimpanzees are able to match vocalizations with gesturing faces (Izumi and Kojima, 2004), and the chimpanzee homolog of Broca's area becomes active during both gestural and vocal communicative actions (Tagliatalata et al., 2008); activation is maximal when gestures are accompanied with vocalizations to call the other's attention (Tagliatalata et al., 2011). In humans, areas 44, 45, and 47 become activated during the integration of speech with gestures (Willems et al., 2007; Gentilucci and Dalla Volta, 2008), and there is evidence for activation of hand motor systems during speech (Gentilucci et al., 2001; Meister et al., 2003). Thus, communication is multimodal both in humans and monkeys, and makes use of overlapping circuits in both species (Aboitiz and García, 2009). This evidence supports the concept that early steps of language evolution also consisted of multimodal signals, instead of being predominantly hand-based or vocalization-based.

WERE GESTURES OR GRASPING REQUIRED FOR THE ADVENT OF SPEECH?

There is abundant evidence for vocalization plasticity in several mammalian species like elephants, bats, seals, and dolphins, not to speak of birds, especially songbirds (Bolhuis et al., 2010). More generally, we may argue that body gestural communication is a widespread characteristic of vertebrates, while vocal communication (innate or learned) has become an important communication pathway only in some lineages. Learned vocalizations are present in even fewer species, coexisting with hand or grasping abilities only in humans, whereas most other vocal learners lack this capacity. Interestingly, cerebral dominance for vocalizations has been reported in many species, both vocal learners and non-vocal learners (Corballis, 2003). Thus, at least in mammals there seems to be no phylogenetic association between grasping abilities and the capacity for vocal learning or imitation. Birds have grasping feet, but it is not known if this ability involves a mirror neuron system, or if its neural representation matches the neural substrate for vocalizations. More likely, imitation tends to be more conspicuous in animals that have developed vocal learning, suggesting that the latter is more closely associated with the acquisition of imitative capacities. Along this line, a vocalization mirror neuron system has been proposed to exist in songbirds, but this possibility and the relation of this putative circuit with a grasping mirror system have yet to be proved (Bonini and Ferrari, 2011).

The grasping mirror neuron network is an ancient characteristic of the primate brain, and therefore cannot by itself account for the origin of vocal language. Among other capacities, an emerging language may have needed shared intentionality, mirror neuron properties, and the capacity to understand actions (Premack, 2004; Tomasello et al., 2005; Corballis, 2010). However, the mirror neuron-gestural perspective does not provide any clue as to how or why speech emerged and became the dominant communication channel. More likely, the key event was the reinforcement of a primitive auditory–vocal sensorimotor circuitry, which, as it expanded, probably took advantage of circuits previously involved in other motor functions, recruiting them for vocal control mechanisms.

TOOL USE, GESTURES, AND A PRIMITIVE SEMANTICS

From the mirror neuron perspective, gestures have been proposed to be crucial for the acquisition of a primitive semantics (Arbib, 2005). In this process, grasping ability and voluntary hand control may have been important elements to facilitate shared attention, and possibly led to the appearance of pointing behavior, which is critical for making reference to the world (Call, 1980). From pointing, other meaningful hand gestures may have evolved, especially in the context of a primitive tool-making and tool-using technology in which the emulation of tool use may have conveyed a ritualized semantics.

There is an extensive literature on tool manufacturing and use in modern humans, early hominids, and non-human primates (Greenfield, 1991; Boesch, 1993; Call and Tomasello, 2007; Ambrose, 2010; Liebal and Call, 2012; Macellini et al., 2012). Observation of tool use produces activation of a sector in the inferior parietal lobe in humans but not in tool-trained monkeys (Peeters et al., 2009). However, the pattern of brain activation

during tool use depends on the tool being used. Comparing two different Paleolithic stone tool tasks, one early (Oldowan), and the other from a later period (Acheulean), Stout and Chaminade (2012) reported that both tasks activated the inferior parietal cortex and the ventral premotor cortex, but only the Acheulean task produced activation of the right inferior frontal gyrus (area 45). These authors and others further propose that tool use and manufacture are hierarchically organized and can be described in a nested syntax, comparable to the recursive syntax of language (Stout and Chaminade, 2012). However, and consistent with the present perspective, they indicate that parsing of behavioral sequences during tool manufacture or use may have provided a bridge between instrumental actions and vocal syntax without the need to invoke a separate communicative gestural stage.

Communicative gestures are derived from non-communicative actions like throwing, grasping or tool use, through a process called ontogenetic ritualization, which may become assimilated during phylogeny (Pika et al., 2005). Orangutans and gorillas have been shown to perform specific gestures that imply distinct meanings, being used intentionally and in a frame of contextual flexibility (Genty et al., 2009; Cartmill and Byrne, 2010). There is also evidence that apes usually incorporate objects in their gestures, and that this correlates with the species' use of tools in the wild (Call and Tomasello, 2007; Liebal and Call, 2012). Pantomimes are gestures resembling the actions they represent but are not effective in performing the action. Whereas in non-human primates these pantomimes are simple representations of actions lacking abstraction, in humans they involve an abstract content, accompany symbolic communication, and may support the signer's capacity for problem solving (Cartmill et al., 2012).

The fact that apes can be taught sign language but are unable to master learned vocalizations has been proposed as supporting a gestural origin for human language (Corballis, 2003). Nevertheless, there is a difference between ontogenetic plasticity and capacity for evolutionary change. A rapid selective trend toward increasing vocal plasticity and vocal control is perfectly possible, and is compatible with the evidence of vocal learning in other mammals and in songbirds (Bolhuis et al., 2010).

However, in this scenario there is little insight into how the transition from gestural references to vocal references could be made. In my view, a gestural pantomime may have been accompanied by the use of sounds imitating the referred object; this simultaneity of gesture and vocalization is likely to have been crucial for the establishment of meaning in vocal behavior (see Tagliatalata et al., 2011). Furthermore, increasing vocal plasticity may have facilitated vocal imitation of physical or animal sounds, rapidly taking over most symbolic contents. To what extent this primitive semantics was gesture-based or vocalization-based will probably never be known, but it is likely that there were several ways to convey meaning, and more importantly, individuals used whatever means they had available, be they gestures, signs, or other signals, to call attention to relevant events under different circumstances.

MIRROR NEURONS AND WORKING MEMORY

Recently, there has been an important debate as to whether motor functions are essential or not for speech processing, which impinges into the mirror neuron – vocal learning debate. A current

interpretation is that the motor system modulates, but does not obligate speech perception (Hickok et al., 2011a,b). However, this modulation may be what is needed to have a better learning capacity, as children with a stronger verbal working memory end up with a larger vocabulary some years later (Baddeley, 2003). In other words, although it may not be necessary for phonological processing, inner speech may protect a perceptual memory trace from interfering processes, helping its maintenance for a longer time (Baddeley, 2003; Marvel and Desmond, 2012).

Furthermore, mirror neurons may eventually prove to be involved in verbal working memory mechanisms. An important component of working memory capacity depends on the close integration between sensory and motor systems, in which audio-vocal mirror neurons may participate, as is perhaps the case in song-learning birds (Bolhuis et al., 2010; Bonini and Ferrari, 2011). The case of conduction aphasia, involving not only a disruption of the white matter as originally considered, but also lesions in the surrounding cortical areas, is characterized by a dysfunction in short-term memory and in imitative capacities (Trortais, 1974; Buchsbaum et al., 2011; Song et al., 2011), which stresses the relation between imitation, sensorimotor integration, and short-term memory. Again, a commonly involved cortical area in conduction aphasia is the posterior planum temporale, i.e., area Spt (Buchsbaum et al., 2011).

SPEECH, BIRDSONG, AND MIRROR NEURONS: DEEP HOMOLOGY?

Finally, some words on studies of vocal learning in songbirds may be worth mentioning here. This has become a rich scientific program in which very different processes, including adult neurogenesis, neural plasticity, gene expression patterns, and even syntactical learning have been addressed (Bolhuis et al., 2010; Abe and Watanabe, 2011; Berwick et al., 2011), confirming Darwin's original speculation of a parallel between speech and birdsong. Moreover, in songbirds, the vocal learning circuit has a similar (but not homologous) architecture as the language circuits, involving cortico-basal ganglia-thalamic circuits (Bolhuis et al., 2010).

In the present context, it may be relevant to mention the recent proposal of a "deep homology" (homology at the gene level) between vocal learning mechanisms in songbirds and humans, based on the participation of the gene FOXP2 in this process (especially in circuits involving the basal ganglia; Scharff and Petri, 2011). FOXP2 is a gene whose mutation causes an inherited verbal dyspraxia in humans, and was initially proposed to be a sort of master-language gene. However, the interpretations of the behavioral phenotype of the affected members are a matter of debate, some proposing that it relates to an inability to denote tense, gender, and other grammatical functions; others view this condition as a phonological articulatory disorder, and still others argue that it affects all levels of language processing (Varga-Khadem et al., 2005). Despite these disagreements, there is evidence that FOXP2 has been a target of selection in the human lineage; it differs from the chimp homolog in two point mutations (Enard et al., 2002; Zhang et al., 2002; Teramitsu et al., 2004; Krause et al., 2007) and is a common transcriptional target of genes displaying accelerated evolution in humans (Lambert et al., 2011). This gene also displays accelerated evolution in echolocating bats, another vocal learning group (Li et al., 2007).

Interestingly, in songbirds, FOXP2 expression is modulated during song learning (Haesler et al., 2004; Teramitsu et al., 2010), and its transcript is required for appropriate song learning (Haesler et al., 2007). Furthermore, diminishing FOXP2 expression produces a decrease in dendritic spine density in the basal ganglia song area of the zebra finch (Schulz et al., 2010). However, deficiency of this gene affects the intensity but not the structure of innate vocalizations in mouse pups (Gaub et al., 2010; Fischer and Hammerschmidt, 2011). Furthermore, mutations of FOXP2 produce generalized deficits in synaptic plasticity and motor learning in mice (Groszer et al., 2008). In light of this evidence, FOXP2, rather than a specific language master gene, is now considered to be involved in more general aspects of sensorimotor learning, and may be of particular relevance for the acquisition of complex, learned motor patterns which include birdsong and speech (Varga-Khadem et al., 2005). If this were the case, any FOXP2-dependent process of sensory-guided learning would represent deep homology with the language and the birdsong circuits.

A few years ago, Corballis (2004) suggested a possible link between FOXP2 and the mirror neuron system, based on evidence indicating underactivity in Broca's area in subjects bearing a mutation of this gene (Liégeois et al., 2003; see also Bosman et al., 2004). It is not yet known whether FOXP2 is specifically expressed in hand-grasping processes in non-human primates. If it were, this evidence would be consistent with the above interpretation, namely that FOXP2 underlies a variety of sensorimotor learning processes, including hand-grasping, speech, and birdsong.

CONCLUDING REMARKS

This review mostly uses information on neural connectivity to establish the phylogenetic continuity of neural circuits involved in speech processing. For reasons of space, other aspects like the comparative microanatomy, cross-species volumetric analyses, and the details of behavioral studies have been discussed only briefly.

Summarizing all the information presented, and considering the several discrepancies in some specific issues, I will take the opportunity to make some concluding remarks. First, the cytoarchitectonic homologs to human areas 44 and 45 are the homonymous areas in the monkey. In the latter, area 44 represents an orofacial specialization of the ventral premotor area 6v (F5), receiving inputs from area 45, which conveys facial and auditory information from the anterior temporal lobe. An arcuate fasciculus may be present in the monkey, but it is probably not a robust tract. Inferior parietal areas send projections to the ventral premotor areas and possibly to area 44 of the monkey. There are discrepancies as to the inferoparietal projection to area 45.

It is likely that the dorsal auditory-vocal pathway via the arcuate fasciculus/SLF did not arise out of nothing, and that a rudimentary auditory pathway to the VLPFC strengthened gradually from monkey to chimpanzee to human. In the chimp, these projections may only have a weak participation in vocalization, but in hominids, neighboring inferior parietal areas were recruited to participate in the planning of motor processes involving vocal articulation, using auditory projections carried by the MLF. The ventral pathway became adapted to transmit echoic and semantic information into the anterior Broca's area.

As is possibly the case in songbirds, it is likely that mirror neurons were included in the nascent phonological loop of early humans, an auditory–vocal sensorimotor pathway with sufficient plasticity and memory capacity to learn complex vocal utterances by imitation. Across species, imitative capacity appears to be associated more with vocal learning than with grasping ability. Nonetheless, it is possible that gestures and vocalizations were both initially used to generate shared attention, which may be a requisite for a primitive semantics. The simultaneity of gestures and vocalizations was likely an important element to transmit stronger messages, and as vocalizations became increasingly sophisticated, they became dominant over gestures.

Thus, human communication is, and has always been, multimodal and opportunistic, using whatever means are available to transmit the intended meaning. Indeed our species is characterized by the urge to communicate things (Tomasello et al., 2005). We have developed a specialized neural device, the phonological

loop that, together with other cognitive specializations, has propelled our communication capacities far beyond those of other animals. Whenever speech is incapable of transmitting information, we literally use the most handy channel at our disposal. That is why, besides sign language, we have developed writing, which is now being transformed into key-pressing, and may eventually become a fully digitalized system for which we may need minimal motor skills.

ACKNOWLEDGMENTS

This work was supported by the Millenium Center for the Neuroscience of Memory, Chile, NC10-001-F, which is developed with funds from the Innovation for Competitiveness from the Ministry for Economics, Fomentation, and Tourism, Chile. I am also grateful to Francisco Zamorano, who prepared the Figures presented in this article; and to John Ewer for his kind help in revising the english language.

REFERENCES

- Abe, K., and Watanabe, D. (2011). Songbirds possess the spontaneous ability to discriminate syntactic rules. *Nat. Neurosci.* 14, 1067–1077.
- Aboitiz, F. (1995). Working memory networks and the origin of language areas in the human brain. *Med. Hypotheses* 44, 504–506.
- Aboitiz, F., Aboitiz, S., and García, R. (2010). The phonological loop: a key innovation in human evolution. *Curr. Anthropol.* 51(Suppl. S1), S55–S65.
- Aboitiz, F., and García, R. (1997). The evolutionary origin of the language areas in the human brain. A neuroanatomical perspective. *Brain Res. Rev.* 25, 381–396.
- Aboitiz, F., and García, R. (2009). Merging of phonological and gestural circuits in early language evolution. *Rev. Neurosci.* 20, 71–84.
- Aboitiz, F., García, R., Bosman, C., and Brunetti, E. (2006a). Cortical memory mechanisms and language origins. *Brain Lang.* 98, 40–56.
- Aboitiz, F., García, R., Brunetti, E., and Bosman, C. (2006b). “The origin of Broca’s area and its connections from an ancestral working/active memory network,” in *Broca’s Area*, eds K. Amunts and Y. Grodzinsky (Oxford: Oxford University Press), 3–16.
- Aboitiz, F., García, R., Brunetti, E., and Bosman, C. (2005). Imitation and memory in language origins. *Neural Netw.* 18, 1357.
- Aboitiz, F., López, J., and Montiel, J. (2003). Long distance communication in the human brain: timing constraints for inter-hemispheric synchrony and the origin of brain lateralization. *Biol. Res.* 36, 89–99.
- Ambrose, S. H. (2010). Coevolution of composite-tool technology, constructive memory, and language: implications for the evolution of modern human behavior. *Curr. Anthropol.* 51, S135–S147.
- Anwander, A., Tittgemeyer, M., Yves von Cramon, D., Friederici, A. D., and Knösche, T. R. (2007). Connectivity-based parcellation of Broca’s area. *Cereb. Cortex* 17, 816–825.
- Arbib, M. A. (2005). From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. *Behav. Brain Sci.* 28, 105–167.
- Arbib, M. A. (2010). Mirror system activity for action and language is embedded in the integration of dorsal and ventral pathways. *Brain Lang.* 112, 12–24.
- Armstrong, A. C., Stokoe, W. C., and Wilcox, S. E. (eds). (1995). *Gesture and the Nature of Language*. Cambridge: Cambridge University Press.
- Awh, E., Jonides, J., Smith, E. E., Schumacher, E. H., Koeppe, R. A., and Katz, S. (1996). Dissociation of storage and rehearsal in working memory: PET evidence. *Psychol. Sci.* 7, 25–31.
- Baddeley, A. (2003). Working memory: looking back and looking forward. *Nat. Rev. Neurosci.* 4, 829–839.
- Baddeley, A. D., Papagno, C., and Vallar, G. (1988). When long-term learning depends on short-term memory. *J. Mem. Lang.* 27, 586–595.
- Badzakova-Trajkov, G., Häberling, I. S., Roberts, R. P., and Corballis, M. C. (2010). Cerebral asymmetries: complementary and independent processes. *PLoS ONE* 5, e9682. doi:10.1371/journal.pone.0009682
- Belin, P., and Zatorre, R. J. (2000). ‘What,’ ‘where’ and ‘how’ in auditory cortex. *Nat. Neurosci.* 3, 965–966.
- Belmalih, A., Borra, E., Contini, M., Gerbella, M., Rozzi, S., and Luppino, G. (2009). Multimodal architectonic subdivision of the rostral part (area F5) of the macaque ventral premotor cortex. *J. Comp. Neurol.* 512, 183–217.
- Bernal, B., and Altman, N. (2010). The connectivity of the superior longitudinal fasciculus: a tractography DTI study. *Magn. Reson. Imaging* 28, 217–225.
- Berwick, R. C., Okanoya, K., Beckers, G. J. L., and Bolhuis, J. J. (2011). Songs to syntax: the linguistics of birdsong. *Trends Cogn. Sci.* 15, 113–122.
- Binkofski, F., Buccino, G., Posse, S., Seitz, R. J., Rizzolatti, G., and Freund, H. (1999). A fronto-parietal circuit for object manipulation in man: evidence from an fMRI-study. *Eur. J. Neurosci.* 11, 3276–3286.
- Boesch, C. (1993). “Transmission of tool-use in wild chimpanzees,” in *Tools, Language and Cognition in Human Evolution*, eds K. R. Gibson and T. Ingold (Cambridge: Cambridge University Press), 171–183.
- Bolhuis, J. J., Okanoya, K., and Scharff, C. (2010). Twitter evolution: converging mechanisms in birdsong and human speech. *Nat. Rev. Neurosci.* 11, 747–759.
- Bonini, L., and Ferrari, P. F. (2011). Evolution of mirror systems: a simple mechanism for complex cognitive functions. *Ann. N. Y. Acad. Sci.* 1225, 166–175.
- Bosman, C., García, R., and Aboitiz, F. (2004). FOXP2 and the language working memory system. *Trends Cogn. Sci.* 8, 251–252.
- Brownsett, S. L. E., and Wise, R. J. S. (2010). The contribution of the parietal lobes to speaking and writing. *Cereb. Cortex* 20, 517–523.
- Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., Seitz, R. J., Zilles, K., Rizzolatti, G., and Freund, H. J. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* 13, 400–404.
- Buccino, G., Lui, F., Canessa, N., Patteri, I., Lagravinese, G., Benuzzi, F., Porro, C. A., and Rizzolatti, G. (2004). Neural circuits involved in the recognition of actions performed by non-conspecifics: an fMRI study. *J. Cogn. Neurosci.* 16, 1–14.
- Buchsbaum, B. R., Baldo, J., Okada, K., Berman, K. F., Dronkers, N., D’Esposito, M., and Hickok, G. (2011). Conduction aphasia, sensory-motor integration, and phonological short-term memory – an aggregate analysis of lesion and fMRI data. *Brain Lang.* 119, 119–128.
- Buchsbaum, B. R., and D’Esposito, M. (2008). The search for the phonological store: from loop to convolution. *J. Cogn. Neurosci.* 20, 762–778.
- Buchsbaum, B. R., Olsen, R. K., Koch, P., and Berman, K. F. (2005a). Human dorsal and ventral auditory streams subserve rehearsal-based and echoic processes during verbal working memory. *Neuron* 48, 687–697.
- Buchsbaum, B. R., Olsen, R. K., Koch, P. F., Kohn, P., Kippenhan, J. S., and Berman, K. F. (2005b). Reading, hearing, and the planum temporale. *Neuroimage* 24, 444–454.

- Bullier, J., Schall, J. D., and Morel, A. (1996). Functional streams in occipito-frontal connections in the monkey. *Behav. Brain Res.* 76, 89–97.
- Call, J., and Tomasello, M. (eds). (2007). *The Gestural Communication of Apes and Monkeys*. Mahwah, NJ: Lawrence Erlbaum Associates.
- Call, J. D. (1980). Some prelinguistic aspects of language development. *J. Am. Psychoanal. Assoc.* 28, 259–289.
- Calvin, W. H. (1983). A stone's throw and its launch window: timing precision and its implications for language and hominid brains. *J. Theor. Biol.* 104, 121–135.
- Cartmill, E. A., Beilock, S., and Goldin-Meadow, S. (2012). A word in the hand: action, gesture, and mental representation in humans and non-human primates. *Philos. Trans. R. Soc. B* 367, 129–143.
- Cartmill, E. A., and Byrne, R. W. (2010). Semantics of primate gestures: intentional meanings of orangutan gestures. *Anim. Cogn.* 13, 793–804.
- Catani, M., and ffytche, D. M. (2005). The rises and falls of disconnection syndromes. *Brain* 128, 2224–2239.
- Cavada, C., and Goldman-Rakic, P. (1989). Posterior parietal cortex in rhesus monkey: II. Evidence for segregated cortico-cortical networks linking sensory and limbic areas with the frontal lobe. *J. Comp. Neurol.* 287, 422–445.
- Corballis, M. C. (1992). On the evolution of language and generativity. *Cognition* 44, 197–126.
- Corballis, M. C. (2002). *From Hand to Mouth: The Origins of Language*. Princeton: Princeton University Press.
- Corballis, M. C. (2003). From mouth to hand: gesture, speech, and the evolution of right-handedness. *Behav. Brain Sci.* 26, 199–208.
- Corballis, M. C. (2004). FOXP2 and the mirror system. *Trends Cogn. Sci.* 8, 95–96.
- Corballis, M. C. (2010). Mirror neurons and the evolution of language. *Brain Lang.* 112, 25–35.
- Coudé, G., Ferrari, P. F., Rodà, F., Maranesi, M., Borelli, E., Veroni, V., Monti, F., Rozzi, S., and Fogassi, L. (2011). Neurons controlling voluntary vocalization in the macaque ventral premotor cortex. *PLoS ONE* 6, e26822. doi:10.1371/journal.pone.0026822
- de Zubicaray, G., Postle, N., McMahon, K., Meredith, M., and Ashton, R. (2010). Mirror neurons, the representation of word meaning, and the foot of the third left frontal convolution. *Brain Lang.* 112, 77–84.
- Deacon, T. W. (1992). Cortical connections of the inferior arcuate sulcus cortex in the macaque brain. *Brain Res.* 573, 8–26.
- Di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., and Rizzolatti, G. (1992). Understanding motor events: a neurophysiological study. *Exp. Brain Res.* 91, 176–180.
- Dubois, J., Hertz-Pannier, L., Cachia, A., Mangin, J. F., Le Bihan, D., and Dehaene-Lambertz, G. (2009). Structural asymmetries in the infant language and sensorimotor networks. *Cereb. Cortex* 19, 414–423.
- Enard, W., Przeworski, M., Fisher, S. E., Lai, C. S., Wiebe, V., Kitano, T., Monaco, A. P., and Pääbo, S. (2002). Molecular evolution of FOXP2, a gene involved in speech and language. *Nature* 418, 869–872.
- Fadiga, L., Fogassi, L., Pavesi, G., and Rizzolatti, G. (1995). Motor facilitation during action observation: a magnetic stimulation study. *J. Neurophysiol.* 73, 2608–2611.
- Ferrari, P. F., Gallese, V., Rizzolatti, G., and Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *Eur. J. Neurosci.* 17, 1703–1714.
- Fiebach, C. J., Schlesewsky, M., and Friederici, A. D. (2002). Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German Wh-questions. *J. Mem. Lang.* 47, 250–272.
- Fiebach, C. J., Schlesewsky, M., Lohmann, G., von Cramon, D. Y., and Friederici, A. D. (2005). Revisiting the role of Broca's area in sentence processing: syntactic integration versus syntactic working memory. *Hum. Brain Mapp.* 24, 79–91.
- Fischer, J., and Hammerschmidt, K. (2011). Ultrasonic vocalizations in mouse models for speech and socio-cognitive disorders: insights into the evolution of vocal communication. *Genes Brain Behav.* 10, 17–27.
- Fogassi, L., Gallese, V., Fadiga, L., and Rizzolatti, G. (1998). Neurons responding to the sight of goal directed hand/arm actions in the parietal area PF (7b) of the macaque monkey. *Soc. Neurosci.* 24, abstr. 257.5.
- Frey, S., Campbell, J. S. W., Pike, G. B., and Petrides, M. (2008). Dissociating the human language pathways with high angular resolution diffusion fiber tractography. *J. Neurosci.* 28, 11435–11444.
- Friederici, A. D. (2004). "The neural basis of syntactic processes," in *The Cognitive Neurosciences III*, ed. M. S. Gazzaniga (Cambridge, MA: MIT Press), 789–801.
- Friederici, A. D. (2005). Revisiting the role of Broca's area in sentence processing: syntactic integration versus syntactic working memory. *Hum. Brain Mapp.* 24, 79–91.
- Friederici, A. D. (2009). Pathways to language: fiber tracts in the human brain. *Trends Cogn. Sci.* 13, 175–181.
- Friederici, A. D., Bahlmann, J., Heim, S., Schubotz, R. I., and Anwander, A. (2006). The brain differentiates human and non-human grammars: functional localization and structural connectivity. *Proc. Natl. Acad. Sci. U.S.A.* 103, 2458–2463.
- Fuster, J. M. (1995). *Memory in the Cerebral Cortex*. Cambridge, MA: MIT Press.
- Galaburda, A. M., and Sanides, F. (1980). Cytoarchitectonic organization of the human auditory cortex. *J. Comp. Neurol.* 190, 597–610.
- Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119, 593–609.
- Gallese, V., Fogassi, L., Fadiga, L., and Rizzolatti, G. (2002). "Action representation and the inferior parietal lobule," in *Attention and Performance XIX. Common Mechanisms in Perception and Action*, eds W. Prinz and B. Hommel (Oxford: Oxford University Press), 247–266.
- Gaub, S., Groszer, M., Fisher, S. E., and Ehret, G. (2010). The structure of innate vocalizations in Foxp2-deficient mouse pups. *Genes Brain Behav.* 9, 390–401.
- Genilucci, M., Benuzzi, F., Gangitano, M., and Grimaldi, S. (2001). Grasp with hand and mouth: a kinematic study on healthy subjects. *J. Neurophysiol.* 86, 1685–1699.
- Genilucci, M., and Dalla Volta, R. (2008). Spoken language and arm gestures are controlled by the same motor control system. *Q. J. Exp. Psychol.* 61, 944–957.
- Genty, E., Breuer, T., Hobaiter, C., and Byrne, R. W. (2009). Gestural communication of the gorilla (*Gorilla gorilla*): repertoire, intentionality and possible origins. *Anim. Cogn.* 12, 527–546.
- Gerardin, E., Sirigu, A., Lehericy, S., Poline, J. B., Gaymard, B., Marsault, C., Agid, Y., and Le Bihan, D. (2000). Partially overlapping neural networks for real and imagined hand movements. *Cereb. Cortex* 10, 1093–1104.
- Gerbella, M., Belmalih, A., Borra, E., Rozzi, S., and Luppino, G. (2007). Multimodal architectonic subdivision of the caudal ventrolateral prefrontal cortex of the macaque monkey. *Brain Struct. Funct.* 212, 269–301.
- Gerbella, M., Belmalih, A., Borra, E., Rozzi, S., and Luppino, G. (2010). Cortical connections of the macaque caudal ventrolateral prefrontal areas 45A and 45B. *Cereb. Cortex* 20, 141–168.
- Gerbella, M., Belmalih, A., Borra, E., Rozzi, S., and Luppino, G. (2011). Cortical connections of the anterior (F5a) subdivision of the macaque ventral premotor area F5. *Brain Struct. Funct.* 216, 43–65.
- Gharbawie, O. A., Stepniewska, I., Qi, H., and Kaas, J. H. (2011). Multiple parietal-frontal pathways mediate grasping in macaque monkeys. *J. Neurosci.* 31, 11660–11677.
- Glasser, M. F., and Rilling, J. K. (2008). DTI tractography of the human brain's language pathways. *Cereb. Cortex* 18, 2471–2482.
- Greenfield, P. M. (1991). Language, tools and brain: the ontogeny and phylogeny of hierarchically organized sequential behavior. *Behav. Brain Sci.* 14, 531–551.
- Grèzes, J., Armony, J. L., Rowe, J., and Passingham, R. E. (2003). Activations related to "mirror" and "canonical" neurons in the human brain: an fMRI study. *Neuroimage* 18, 928–937.
- Groszer, M., Keays, D. A., Deacon, R. M., de Bono, J. P., Prasad-Mulcare, S., Gaub, S., Baum, M. G., French, C. A., Nicod, J., Coventry, J. A., Enard, W., Fray, M., Brown, S. D., Nolan, P. M., Pääbo, S., Channon, K. M., Costa, R. M., Eilers, J., Ehret, G., Rawlins, J. N., and Fisher, S. E. (2008). Impaired synaptic plasticity and motor learning in mice with a point mutation implicated in human speech deficits. *Curr. Biol.* 18, 354–362.
- Häberling, I. S., Badzakova-Trajkov, G., and Corballis, M. C. (2011). Callosal tracts and patterns of hemispheric dominance: a combined fMRI and DTI study. *Neuroimage* 54, 779–786.
- Haesler, S., Rochefort, C., Georgi, B., Licznarski, P., Osten, P., and Scharff, C. (2007). Incomplete and inaccurate vocal imitation after knockdown of FoxP2 in songbird basal ganglia nucleus Area X. *PLoS Biol.* 5, e321. doi:10.1371/journal.pbio.0050321
- Haesler, S., Wada, K., Nshdejan, A., Morrissey, E. E., Lints, T., Jarvis, E. D., and Scharff, C. (2004). FoxP2 expression in avian vocal learners

- and non-learners. *J. Neurosci.* 24, 3164–3375.
- Hartwigsen, G., Baumgaertner, A., Price, C. J., Koehnke, M., Ulmer, S., and Siebner, H. R. (2010). Phonological decisions require both the left and right supramarginal gyri. *Proc. Natl. Acad. Sci. U.S.A.* 107, 16494–16499.
- Hewes, G. W. (1973). Primate communication and the gestural origin of language. *Curr. Anthropol.* 12, 5–24.
- Hickok, G. (2009). The functional neuroanatomy of language. *Phys. Life Rev.* 6, 121–143.
- Hickok, G., Houde, J., and Rong, F. (2011a). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* 69, 407–422.
- Hickok, G., Costanzo, M., Capasso, R., and Miceli, G. (2011b). The role of Broca's area in speech perception: evidence from aphasia revisited. *Brain Lang.* 119, 214–220.
- Hickok, G., and Poeppel, D. (2007). The cortical organization of speech processing. *Nat. Rev. Neurosci.* 8, 393–402.
- Hopkins, W. D., Russell, J. L., and Schaeffer, J. A. (2012). The neural and cognitive correlates of aimed throwing in chimpanzees: a magnetic resonance image and behavioral study on a unique form of social tool use. *Philos. Trans. R. Soc. B* 367, 37–47.
- Iacoboni, M., and D'Apretto, M. (2006). The mirror neuron system and the consequences of its dysfunction. *Nat. Rev. Neurosci.* 7, 942–951.
- Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., and Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science* 286, 2526–2528.
- Ide, A., Dolezal, C., Fernández, M., Labbé, E., Mandujano, R., Montes, S., Segura, P., Verschae, G., Yarmuch, P., and Aboitiz, F. (1999). Hemispheric differences in the variability of fissural patterns in parasyllvian and cingulate regions of human brains. *J. Comp. Neurol.* 410, 235–242.
- Izumi, A., and Kojima, S. (2004). Matching vocalizations to vocalizing faces in a chimpanzee (*Pan troglodytes*). *Anim. Cogn.* 7, 179–184.
- Johnson Frey, S. H., Maloof, F. R., Newman-Norlund, R., Farrer, C., Inati, S., and Grafton, S. T. (2003). Actions or hand-objects interactions? Human inferior frontal cortex and action observation. *Neuron* 39, 1053–1058.
- Josse, G., and Tzourio-Mazoyer, N. (2004). Hemispheric specialization for language. *Brain Res. Rev.* 44, 1–12.
- Jürgens, U., and Alipour, M. (2002). A comparative study on the corticohypoglossal connections in primates, using biotin dextran amine. *Neurosci. Lett.* 328, 245–248.
- Kaas, J. H., Gharbawie, O. A., and Stepniewska, I. (2011). The organization and evolution of dorsal stream multisensory motor pathways in primates. *Front. Neuroanat.* 5:34. doi:10.3389/fnana.2011.00034
- Kaas, J. H., and Hackett, T. A. (1999). 'What' and 'where' processing in auditory cortex. *Nat. Neurosci.* 2, 1045–1047.
- Keller, S. S., Roberts, N., and Hopkins, W. (2009). A comparative magnetic resonance imaging study of the anatomy, variability, and asymmetry of Broca's area in the human and chimpanzee brain. *J. Neurosci.* 29, 14607–14616.
- Kelly, K., Uddin, L. Q., Shehzad, Z., Margulies, D. S., Castellanos, F. X., Milham, M. P., and Petrides, M. (2010). Broca's region: linking human brain functional connectivity data and nonhuman primate tracing anatomy studies. *Eur. J. Neurosci.* 32, 383–398.
- Kendon, A. (2004). *Gesture: Visible Action*. Cambridge: Cambridge University Press.
- Keyzers, C., Kohler, E., Umiltà, M. A., Nanetti, L., Fogassi, L., and Gallese, V. (2003). Audiovisual mirror neurons and action recognition. *Exp. Brain Res.* 153, 628–636.
- Kilner, K. M. (2011). More than one pathway to action understanding. *Trends Cogn. Sci.* 15, 352–357.
- Krause, J., Lalueza-Fox, C., Orlando, L., Enard, W., Green, R. E., Burbano, H. A., Hublin, J. J., Hänni, C., Fortea, J., de la Rasilla, M., Bertranpetit, J., Rosas, A., and Pääbo, S. (2007). The derived FOXP2 variant of modern humans was shared with Neanderthals. *Curr. Biol.* 17, 1908–1912.
- Lambert, N., Lambot, M. A., Bilheu, A., Albert, V., Englert, Y., Libert, F., Noel, J. C., Sotiriou, C., Holloway, A. K., Pollard, K. S., Detours, V., and Vanderhaeghen, P. (2011). Genes expressed in specific areas of the human fetal cerebral cortex display distinct patterns of evolution. *PLoS ONE* 6, e17753. doi:10.1371/journal.pone.0017753
- Leichnetz, G. R. (2001). Connections of the medial posterior parietal cortex (area 7m) in the monkey. *Anat. Rec.* 263, 215–236.
- Li, G., Wang, J., Rossiter, S. J., Jones, G., and Zhang, S. (2007). Accelerated FOXP2 evolution in echolocating bats. *PLoS ONE* 2, e900. doi:10.1371/journal.pone.0000900
- Liebal, K., and Call, J. (2012). The origins of non-human primates' manual gestures. *Philos. Trans. R. Soc. Lond. B* 367, 118–128.
- Liégeois, F., Baldeweg, T., Connelly, A., Gadian, D. G., Mishkin, M., and Vargha-Khadem, F. (2003). Language fMRI abnormalities associated with FOXP2 gene mutation. *Nat. Neurosci.* 6, 1230–1237.
- Lonsdorf, E. V., and Hopkins, W. D. (2005). Wild chimpanzees show population-level handedness for tool use. *Proc. Natl. Acad. Sci. U.S.A.* 102, 12634–12638.
- Lopez-Barroso, D., de Diego-Balaguer, R., Cunillera, T., Camara, E., Münte, T. F., and Rodriguez-Fornells, A. (2011). Language learning under working memory constraints correlates with microstructural differences in the ventral language pathway. *Cereb. Cortex* 21, 2742–2750.
- Luppino, G., Murata, A., Govoni, P., and Matelli, M. (1999). Largely segregated parietofrontal connections linking rostral intraparietal cortex (areas AIP and VIP) and the ventral premotor cortex (areas F5 and F4). *Exp. Brain Res.* 128, 181–187.
- Macellini, S., Maranesi, M., Bonini, L., Simone, L., Rozzi, S., Ferrari, P. F., and Fogassi, L. (2012). Individual and social learning processes involved in the acquisition and generalization of tool use in macaques. *Philos. Trans. R. Soc. B* 367, 24–36.
- Maeda, F., Kleiner-Fisman, G., and Pascual-Leone, A. (2002). Motor facilitation while observing hand actions: specificity of the effect and role of observer's orientation. *J. Neurophysiol.* 87, 1329–1335.
- Marvel, C. L., and Desmond, J. E. (2012). From storage to manipulation: how the neural correlates of verbal working memory reflect varying demands on inner speech. *Brain Lang.* 120, 42–51.
- Matelli, M., Camarda, R., Glickstein, M., and Rizzolatti, G. (1986). Afferent and efferent projections of the inferior area 6 in the macaque monkey. *J. Comp. Neurol.* 251, 281–298.
- Meister, I. G., Boroojerdi, B., Foltys, H., Sparing, R., Huber, W., and Topper, R. (2003). Motor cortex hand area and speech: implications for the development of language. *Neuropsychologia* 41, 401–406.
- Mesulam, M. M., Van Hoesen, G. W., Pandya, D. N., and Geschwind, N. (1977). Limbic and sensory connections of the inferior parietal lobule (area PG) in the rhesus monkey: a study with a new method for horseradish peroxidase histochemistry. *Brain Res.* 136, 393–414.
- Molenberghs, P., Cunnington, R., and Mattingley, J. B. (2009). Is the mirror neuron system involved in imitation? A short review and meta-analysis. *Neurosci. Biobehav. Rev.* 33, 975–980.
- Moser, D., Baker, J. M., Sanchez, C. E., Rorden, C., and Fridriksson, J. (2009). Temporal order processing of syllables in the left parietal lobe. *J. Neurosci.* 29, 12568–12573.
- Nelissen, K., Borra, E., Gerbella, M., Rozzi, S., Luppino, G., Vanduffel, W., Rizzolatti, G., and Orban, G. A. (2011). Action observation circuits in the macaque monkey cortex. *J. Neurosci.* 31, 3743–3756.
- Nelissen, K., Luppino, G., Vanduffel, W., Rizzolatti, G., and Orban, G. A. (2005). Observing others: multiple action representation in the frontal lobe. *Science* 310, 332–336.
- Pandya, D. N., and Yeterian, E. H. (1985). "Architecture and connections of cortical association areas," in *Cerebral Cortex*, Vol. 4, eds A. Peters and E. G. Jones (New York: Plenum Press), 3–61.
- Parker, G. J. M., Luzzi, S., Alexander, D. C., Wheeler-Kingshott, C. A. M., Ciccarelli, O., and Ralph, M. A. L. (2005). Laterality of ventral and dorsal auditory language pathways in the human brain. *Neuroimage* 24, 656–666.
- Paulesu, E., Frith, C. D., and Frackowiak, R. S. J. (1993). The neural correlates of the verbal component of working memory. *Nature* 362, 34–25.
- Peeters, R., Simone, L., Nelissen, K., Fabbri-Destro, M., Vanduffel, W., Rizzolatti, G., and Orban, G. A. (2009). The representation of tool use in humans and monkeys: common and uniquely human features. *J. Neurosci.* 29, 11523–11539.
- Perrett, D. I., Mistlin, A. J., Harries, M. H., and Chitty, A. J. (1990). "Understanding the visual appearance and consequence of hand actions," in *Vision and Action: The Control of Grasping*, ed. M. A. Goodale (Norwood, NJ: Ablex), 163–342.
- Petkov, C. I., Kayser, C., Steudel, T., Whittingstall, K., Augath, M., and Logothetis, N. K. (2008). A voice region in the monkey brain. *Nat. Neurosci.* 11, 367–374.
- Petrides, M., Cadoret, G., and Mackey, S. (2005). Orofacial somatomotor responses in the macaque monkey homologue of Broca's area. *Nature* 435, 1235–1238.
- Petrides, M., and Pandya, D. N. (1984). Projections from the frontal cortex from the posterior parietal region in the rhesus monkey. *J. Comp. Neurol.* 228, 105–116.

- Petrides, M., and Pandya, D. N. (1988). Association fiber pathways to the frontal cortex from the superior temporal region in the rhesus monkey. *J. Comp. Neurol.* 273, 52–66.
- Petrides, M., and Pandya, D. N. (1999). Dorsolateral prefrontal cortex: comparative cytoarchitectonic analysis in the human and the macaque brain and corticocortical connection patterns. *Eur. J. Neurosci.* 11, 1011–1036.
- Petrides, M., and Pandya, D. N. (2002). Comparative cytoarchitectonic analysis of the human and the macaque ventrolateral prefrontal cortex and corticocortical connection patterns in the monkey. *Eur. J. Neurosci.* 16, 291–310.
- Petrides, M., and Pandya, D. N. (2009). Distinct parietal and temporal pathways to the homologues of Broca's area in the monkey. *PLoS Biol.* 7, e1000170. doi:10.1371/journal.pbio.1000170
- Pika, S., Liebal, K., Call, J., and Tomasello, M. (2005). Gestural communication of apes. *Gesture* 5, 41–56.
- Premack, D. (2004). Is language the key to human intelligence? *Science* 303, 318–320.
- Preuss, T. (1995). "The argument from humans to animals in cognitive neuroscience," in *The Cognitive Neurosciences*, ed. M. S. Gazzaniga (Boston: MIT Press), 1227–1241.
- Preuss, T., and Goldman-Rakic, P. S. (1991a). Myelo- and cytoarchitecture of the granular frontal cortex and surrounding regions in the strepsirrhine primate Galago and the anthropoid primate Macaca. *J. Comp. Neurol.* 310, 429–474.
- Preuss, T., and Goldman-Rakic, P. S. (1991b). Architectonics of the parietal and temporal association cortex in the strepsirrhine primate Galago compared to the anthropoid primate Macaca. *J. Comp. Neurol.* 310, 475–506.
- Preuss, T., and Goldman-Rakic, P. S. (1991c). Ipsilateral cortical connections of granular frontal cortex in the strepsirrhine primate Galago, with comparative comments on anthropoid primates. *J. Comp. Neurol.* 310, 507–549.
- Rauschecker, J. P., and Tian, B. (2000). Mechanisms and streams for processing of "what" and "where" in auditory cortex. *Proc. Natl. Acad. Sci. U.S.A.* 97, 11800–11806.
- Reuland, E. (2010). Imagination, planning, and working memory: the emergence of language. *Curr. Anthropol.* 51(Suppl. 1), S99–S110.
- Rilling, J. K., Glasser, M. F., Preuss, T. M., Ma, X., Zhao, T., Hu, X., and Behrens, T. E. J. (2008). The evolution of the arcuate fasciculus revealed with comparative DTI. *Nat. Neurosci.* 11, 426–428.
- Ringo, J. L., Doty, R. W., Demeter, S., and Simard, P. Y. (1994). Time is of the essence: a conjecture that hemispheric specialization arises from inter-hemispheric conduction delay. *Cereb. Cortex* 4, 331–343.
- Rizzolatti, G., and Arbib, M. A. (1998). Language within our grasp. *Trends Neurosci.* 21, 188–194.
- Rizzolatti, G., Fadiga, L., Fogassi, L., and Gallese, V. (1996). Premotor cortex and the recognition of motor actions. *Cogn. Brain Res.* 3, 131–141.
- Rizzolatti, G., and Luppino, G. (2001). The cortical motor system. *Neuron* 31, 889–901.
- Rizzolatti, G. W., and Craighero, L. (2004). The mirror-neuron system. *Ann. Rev. Neurosci.* 27, 169–192.
- Rolheiser, T., Stamatakis, E. A., and Tyler, L. K. (2011). Dynamic processing in the human language system: synergy between the arcuate fasciculus and extreme capsule. *J. Neurosci.* 31, 16949–16957.
- Romanski, L. M. (2007). Representation and integration of auditory and visual stimuli in the primate ventral lateral prefrontal cortex. *Cereb. Cortex* 17(Suppl. 1), i61–i69.
- Romanski, L. M., Averbeck, B. B., and Diltz, M. (2005). Neural representation of vocalizations in the primate ventrolateral prefrontal cortex. *J. Neurophysiol.* 93, 734–747.
- Romanski, L. M., Bates, J. F., and Goldman-Rakic, P. S. (1999a). Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. *J. Comp. Neurol.* 403, 141–157.
- Romanski, L. M., Tian, B., Mishkin, M., Goldman-Rakic, P. S., and Rauschecker, J. P. (1999b). Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nat. Neurosci.* 2, 1131–1136.
- Romanski, L. M., and Goldman-Rakic, P. S. (2002). An auditory domain in primate prefrontal cortex. *Nat. Neurosci.* 5, 15–16.
- Saffran, E. M., and Marin, O. S. (1975). Immediate memory for word lists and sentences in a patient with deficient auditory short-term memory. *Brain Lang.* 2, 420–433.
- Salmon, E., Van der Linden, M., Collette, F., Delfiore, G., Maquet, P., Degueldre, C., Luxen, A., and Franck, G. (1996). Regional brain activity during working memory tasks. *Brain* 119, 1617–1625.
- Saur, D., Kreher, B. W., Schnell, S., Kümmerer, D., Kellmeyer, P., Vry, M.-S., Umarova, R., Musso, M., Glauche, V., Abel, S., Huber, W., Rijntjes, M., Hennig, J., and Weiller, C. (2008). Ventral and dorsal pathways for language. *Proc. Natl. Acad. Sci. U.S.A.* 105, 18035–18040.
- Scharff, C., and Petri, J. (2011). Evo-devo, deep homology and FoxP2: implications for the evolution of speech and language. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 366, 2124–2140.
- Schenker, N. M., Buxhoeveden, D. P., Blackmon, W. L., Amunts, K., Zilles, K., and Semendeferi, K. (2008). A comparative quantitative analysis of cytoarchitecture and minicolumnar organization in Broca's area in humans and great apes. *J. Comp. Neurol.* 510, 117–128.
- Schulz, S. B., Haesler, S., Scharff, C., and Rochefort, C. (2010). Knockdown of FOXP2 alters spine density in Area X of the zebra finch. *Genes Brain Behav.* 9, 732–740.
- Seltzer, B., and Pandya, D. N. (1978). Afferent cortical connections and architectonics of the superior temporal sulcus and surrounding cortex in the rhesus monkey. *Brain Res.* 149, 1–24.
- Seltzer, B., and Pandya, D. N. (1994). Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: a retrograde tracer study. *J. Comp. Neurol.* 343, 445–463.
- Shallice, T., and Warrington, E. K. (1970). Independent functioning of verbal memory stores: a neuropsychological study. *Q. J. Exp. Psychol.* 22, 261–273.
- Sherwood, C. C., Broadfield, D. C., Holloway, R. L., Gannon, P. J., and Hof, P. R. (2003). Variability of Broca's area homologue in African great apes: implications for language evolution. *Anat. Rec. A Discov. Mol. Cell. Evol. Biol.* 271, 276–285.
- Smith, E. E., and Jonides, J. (1998). Neuroimaging analyses of human working memory. *Proc. Natl. Acad. Sci. U.S.A.* 95, 12061–12068.
- Song, X., Dornbos, D. III., Lai, Z., Zhang, Y., Li, T., Chen, H., and Yang, Z. (2011). Diffusion tensor imaging and diffusion tensor imaging-fibre tractograph depict the mechanisms of Broca-like and Wernicke-like conduction aphasia. *Neurol. Res.* 33, 529–535.
- Steele, J., and Uomini, N. (2009). Can the archaeology of manual specialization tell us anything about language evolution? A survey of the state of play. *Camb. Archaeol. J.* 19, 97–110.
- Stout, D., and Chaminade, T. (2012). Stone tools, language and the brain in human evolution. *Philos. Trans. R. Soc. Lond. B* 367, 75–87.
- Tagliatalata, J. P., Russell, J. L., Schaeffer, J. A., and Hopkins, W. D. (2008). Communicative signaling activates "Broca's" homolog in chimpanzees. *Curr. Biol.* 18, 343–348.
- Tagliatalata, J. P., Russell, J. L., Schaeffer, J. A., and Hopkins, W. D. (2011). Chimpanzee vocal signaling points to a multimodal origin of human language. *PLoS ONE* 6, e18852. doi: 10.1371/journal.pone.0018852
- Teramitsu, I., Kudo, L. C., London, S. E., Geschwind, D. H., and White, S. A. (2004). Parallel FOXP1 and FOXP2 expression in songbird and human brain predicts functional interaction. *J. Neurosci.* 24, 3152–3163.
- Teramitsu, I., Poopatanapong, A., Torrisi, S., and White, S. A. (2010). Striatal FOXP2 is actively regulated during songbird sensorimotor learning. *PLoS ONE* 5, e8548. doi:10.1371/journal.pone.0008548
- Tian, B., Reser, D., Durham, A., Kustov, A., and Rauschecker, J. P. (2001). Functional specialization in rhesus monkey auditory cortex. *Science* 292, 290–293.
- Tomasello, M., Carpenter, M., Call, J., Behne, T., and Moll, H. (2005). Understanding and sharing intentions: the origins of cultural cognition. *Behav. Brain Sci.* 28, 675–691; discussion 691–735.
- Trortais, A. (1974). Impairment of memory for sequences in conduction aphasia. *Neuropsychologia* 12, 355–366.
- Turkeltaub, P. E., and Coslett, H. B. (2010). Localization of sublexical speech perception components. *Brain Lang.* 114, 1–15.
- Ullman, M. T. (2004). Contributions of memory circuits to language: the declarative/procedural model. *Cognition* 92, 231–270.
- Varga-Khadem, F., Gadian, D. G., Copp, A., and Mishkin, M. (2005). FOXP2 and the neuroanatomy of speech and language. *Nat. Rev. Neurosci.* 6, 131–138.
- Vogt, S., Buccino, G., Wohlschläger, A. M., Canessa, N., Shah, N. J., Zilles, K., Eickhoff, S. B., Freund, H. J., Rizzolatti, G., and Fink, G. R. (2007). Prefrontal involvement in imitation learning of hand actions: effects of practice and expertise. *Neuroimage* 37, 1371–1383.

- Willems, R. M., Ozyürek, A., and Hagoort, P. (2007). When language meets action: the neural integration of gesture and speech. *Cereb. Cortex* 17, 2322–2333.
- Wilson, S. M., Galantucci, S., Tartaglia, M. C., Rising, K., Patterson, D. K., Henry, M. L., Ogar, J. M., DeLeon, J., Miller, B. L., and Gorno-Tempini, M. L. (2011). Syntactic processing depends on dorsal language tracts. *Neuron* 72, 397–403.
- Yeterian, E. H., Pandya, D. N., Tomaiuolo, F., and Petrides, M. (2012). The cortical connectivity of the prefrontal cortex in the monkey brain. *Cortex* 48, 58–81.
- Zhang, J., Webb, D. M., and Podlaha, O. (2002). Accelerated protein evolution and origins of human-specific features: FOXP2 as an example. *Genetics* 162, 1825–1835.
- Conflict of Interest Statement:** The author declares that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.
- Received: 28 October 2011; paper pending published: 06 December 2011; accepted: 11 January 2012; published online: 01 February 2012.*
- Citation: Aboitiz F (2012) Gestures, vocalizations, and memory in language origins. Front. Evol. Neurosci. 4:2. doi: 10.3389/fnevo.2012.00002*
- Copyright © 2012 Aboitiz. This is an open-access article distributed under the terms of the Creative Commons Attribution Non Commercial License, which permits non-commercial use, distribution, and reproduction in other forums, provided the original authors and source are credited.