Language beyond action

Ivan Toni [a,b], Floris P. de Lange [a], Matthijs L. Noordzij [a], Peter Hagoort [a,b,c,*]

[a] F.C. Donders Centre for Cognitive Neuroimaging, Radboud University Nijmegen, P.O. Box 9101, 6500 HB Nijmegen, Netherlands
[b] Nijmegen Institute for Cognition and Information, Radboud University Nijmegen, Netherlands
[c] Max Planck Institute for Psycholinguistics, Nijmegen, Netherlands

ABSTRACT

The discovery of mirror neurons in macaques and of a similar system in humans has provided a new and fertile neurobiological ground for rooting a variety of cognitive faculties. Automatic sensorimotor resonance has been invoked as the key elementary process accounting for disparate (dys)functions, like imitation, ideomotor apraxia, autism, and schizophrenia. In this paper, we provide a critical appraisal of three of these claims that deal with the relationship between language and the motor system. Does language comprehension require the motor system? Was there an evolutionary switch from manual gestures to speech as the primary mode of language? Is human communication explained by automatic sensorimotor resonances? A positive answer to these questions would open the tantalizing possibility of bringing language and human communication within the fold of the motor system.

We argue that the available empirical evidence does not appear to support these claims, and their theoretical scope fails to account for some crucial features of the phenomena they are supposed to explain. Without denying the enormous importance of the discovery of mirror neurons, we highlight the limits of their explanatory power for understanding language and communication.

© 2008 Elsevier Ltd. All rights reserved.

1. Introduction

It is beyond any doubt that there are intricate links and interactions between language and the motor system. If a speaker or signer wants to communicate his thoughts, it is through the execution of articulatory gestures (speech), or of hand and face movements in space (sign language). This can only be done via the recruitment of the motor system, and the involvement of premotor and motor cortices. The relatively recent discovery of mirror neurons in monkey brain (di Pellegrino et al., 1992a) and of the mirror neuron system as its human analogue (Rizzolatti and Craighero, 2004) has led to a boost of much stronger claims. If not explicitly, there is at least the implicit claim of a more direct or exclusive relation between language and action than between language and perception. More in particular, we will discuss in some detail three claims that have emerged in the literature about the nature of the relation between the motor system, mirror neurons, and human linguistic abilities. The first one is that not only speaking, but also speech perception, and more generally language comprehension, requires the involvement of motor systems (Galantucci et al., 2006). The second claim is that an action system (i.e., manual gestures) was the evolutionary precursor of speech as the primary mode of language (Arbib,

2005; Rizzolatti and Arbib, 1998). The final claim that we will address is that human communication can be effective without any cognitive mediation, by virtue of automatic sensorimotor resonances (Rizzolatti and Craighero, 2004). Our conclusion will be that due to the enormous appeal and importance of mirror neurons, we might have gone too far in drawing most of the language system into the domain of action understanding. Language clearly goes beyond action.

2. The involvement of the motor system in language processing and representation

The motor system has been claimed to play a role both at the level of word forms and word meanings. The former deals with addressing fundamental issues in speech perception, the latter refers to the semantics of action words. We will discuss the involvement of the motor system at both levels.

A fundamental problem in speech perception is that the realization of speech sounds is highly context dependent. Successive speech sounds are produced by vocal tract gestures that overlap temporally. This is referred to as coarticulation. The consequence of coarticulation is that there is no one-to-one correspondence between an acoustic event and the repertoire of phonemes or phonetic features in the language. This is known as the invariance problem. For instance, one of the acoustic cues for stop consonants, namely the burst of energy generated during the release of the

vocal tract constriction, is context dependent. The same holds for a second cue, which is the formant transition that occurs as the consonant gives way to the following vowel. If we measure the transition of second formant transition for the consonant /d/ into the vowels in the syllables /di/ and /du/, it turns out that this transition is markedly different in the two cases. The second formant (F2) contains information about place or articulation (e.g., specifying that this sound is a /d/ instead of a /b/). In general this F2 is lower for /b/ than for /d/, but in the syllable /bi/ it is in fact as high, or even higher, as in the syllable /du/. How do these very different second formant characteristics in the syllables /di/ and /du/ give rise to the invariant percept of the consonant /d/? This is the fundamental problem that theories of speech perception have to account for.

Different proposals have been made as to how invariance can be achieved (for a recent review, see, Diehl et al., 2004). A view that is shared by many in the field of speech perception is that speech is perceived by exploiting the same auditory mechanisms that analyse other classes of environmental sounds. On the way to activating word forms in the mental lexicon (lexical access), the acoustic information gives way to an intermediate input representation. Some models assume that this is a phonemic level of representation (Foss and Gernsbacher, 1983; Marslen-Wilson and Welsh, 1978; Pisoni and Luce, 1987). Others have suggested alternative 'units of perception' such as syllables or stress units (a stressed syllable plus sometimes one or more unstressed syllables) (see, Cutler and Clifton, 1999, for a review). Alternatively, the necessity of an intermediate representation has been denied. Instead, a continuous mapping of the acoustic input onto stored word form representations is assumed to take place. Despite the differences in the ideas about the transformation of the acoustic signal to a more abstract memory trace specifying the phonological word form, all these models assume that speech perception extracts the invariance in the acoustic input through some mapping onto an input-specific stored form representation (see also Levelt et al., 1999).

A radical alternative to auditory accounts of speech perception was formulated by Alvin Liberman and his colleagues at the Haskins Laboratories (Liberman, 1957; Liberman et al., 1952, 1954, 1956). According to Liberman, the listener does not solve the invariance problem in the auditory domain, but instead in the motor domain. The acoustic patterns might be different, but the articulatory gestures that are needed to produce them are the same. Liberman's fundamental idea was that the perceptual problem might be solved by recruiting the production system. A listener will understand speech by virtue of being a speaker. This idea has become known as the motor theory of speech perception. The objects of speech perception are thus the articulatory events. According to this theory, the human listener recovers the neuromotor commands to the articulators (e.g., tongue, lips, and vocal folds) (Liberman and Mattingly, 1985). The neuromotor commands provide the required invariance (for an insightful review of the different claim and the different versions of the motor theory of speech perception, see, Galantucci et al., 2006).

Despite its general appeal in the context of the discovery of a mirror neuron system in humans, the motor theory of speech perception faces a number of serious challenges. One challenge originates in the finding of speech recognition capabilities in species that lack a speech production system. For instance, it has been found that birds can extract the invariant properties of consonants in different vowel contexts. Kluender et al. (1987) trained Japanese quails to respond to the consonant /d/ in various vowel contexts, but to withold a response to /b/ and /g/ followed by the same vowels. In general, convincing evidence has been provided that animals are capable of showing perceptual skills that were considered to be uniquely human (e.g., categorical perception), or for which a speech production apparatus is claimed to be required by the pro-

ponents of the motor theory of speech perception (e.g., Kluender et al., 1987; Kuhl and Miller, 1975, 1978; Kuhl, 1981. Although this is not decisive evidence that humans do not recruit the motor system for perceiving speech, the animal data provide clear evidence that some of the invariants in the speech input can be extracted without a mapping on the vocal tract gestures. Other findings are also difficult to account for by the motor theory. One of these findings is that, if acoustic properties of speech are artificially transduced into vibrotactile patterns on the skin, listeners are able to extract information that helps in identifying phonemes (Sparks et al., 1978). It is not easy to see how these vibrotactile patterns can be meaningfully said to specify vocal gestures. For this and other reasons (see, Diehl et al., 2004, for an in-depth discussion) the motor theory of speech perception "[…] has few proponents within the field of speech perception, and many authors cite it primarily to offer critical commentary" (Galantucci et al., 2006). Most researchers in the domain of speech recognition follow what is sometimes referred to as the general approach (Diehl et al., 2004). In this approach, it is assumed that speech sounds are perceived with the help of the same mechanisms of audition and perceptual learning that have evolved in humans and their ancestors to process other classes of environmental sounds.

In short, among researchers of speech perception, the motor theory is not undisputed. Moreover, a clear idea about how the acoustic input gets transformed into a "gestural score", is lacking. The finding of mirror neurons, suggestive as it might be for a motor theory account of speech perception, in itself does not solve the problems that the motor theory of speech perception is facing. In addition, the finding that motor systems are activated during speech perception is necessary but not sufficient evidence that these systems are involved in speech recognition. Only when it can be shown that speech recognition is no longer possible if the motor areas that subserve the neuromotor commands of the articulatory gestures are lesioned, do we have sufficient reason to believe that the motor theory of speech perception is correct.

What we have seen so far is that the motor theory of speech perception introduced the claim that human listeners perceive speech sounds by mapping the input onto the domain of motor commands necessary to produce these very same sounds. This involves parity at the form level. Accounts that connect mirror neurons to language, or are inspired by the mirror neuron findings, also make the claim that there exists a *semantic* links between language and the motor system (Aziz-Zadeh et al., 2006; Rizzolatti and Arbib, 1998; Rizzolatti and Craighero, 2004). The idea is that the mirror neuron system is not only used for establishing parity between sender and receiver at the form level, but also at the semantic level [But see (Arbib, 2006) for an account that connects the *evolved* mirror system to phonology (word-as-action) not to semantics]. The empirical evidence is mainly derived from studies showing that action words result in activation in motor areas (Hauk et al., 2004). We do not deny that there is convincing evidence suggesting a role of motor areas in the conceptual representation of action words. However, this evidence is often presented in the context of a claim for a special relation between language and the motor system. This implication is not supported by the facts. There is no decisive evidence that motor systems play an exclusive role in semantics.

To make our argument more clear, we first need to discuss the issue of the representation of (concrete) concepts. A long tradition of neuropsychological evidence exists on the role of sensorimotor attributes in the semantic representation of concrete nouns (Hagoort, 1998). Important evidence comes from patients with category-specific deficits. Especially Elizabeth Warrington and her colleagues (for an overview, see, McKenna and Warrington, 1993) have reported a number of single cases in which semantic knowledge of either objects of nature (e.g., animals, fruits, vegetables) or

man-made artefacts (e.g., tools, furniture, kitchen utensils) was selectively impaired. A theoretically important issue raised by the reports of these category-specific deficits is to determine the precise nature of the semantic dimension that results in specific semantic impairments for living and non-living things. An influential view relates these impairments to the distinction between perceptual and functional attributes (Warrington and Shallice, 1984). Warrington and Shallice (1984) suggested that sensory attributes are very salient features for the identification of living things such as animals or fruits. In contrast, functional attributes are probably more important than perceptual characteristics for the identification of artefacts such as tools. Empirical evidence supports the claim that visual features are more salient in definitions of living things than of artefacts (Farah and McClelland, 1991). In a further refinement of this account, Warrington and McCarthy (1987) and McCarthy and Warrington (1990) have proposed that the contributions of sensory (perceptual) and motor (functional) channels are differentially weighted not only between but also within categories. For instance, within the category of artefacts, small manipulable tools are associated with a repertoire of skilled movements, and hence rely more heavily on motor channels than do large man-made objects such as aeroplanes. Aeroplanes are probably not too different from birds with respect to our reliance on sensory channels for their identification and categorization. In short, patient data provide strong evidence for a view that object concepts are grounded in perception and action, with presumably on top of this some additional organization by category for evolutionary wetwired categories such as animals, conspecifics, plants, and possibly tools (for recent reviews, see, Martin and Caramazza, 2003; Martin, 2007).

Imaging studies have provided further supportive evidence that there is a relation between conceptual knowledge and brain systems for perception and action. In a PET study, Martin et al. (1996) asked subjects to name pictures of animals and of tools. For pictures of both kinds bilateral activation was obtained in ventral regions of the temporal lobes. The naming of animals resulted in additional activation in the left medial occipital lobe, an area involved in visual processing. In contrast, the naming of tools led to additional activation in the left premotor area and an area in the left middle temporal gyrus. These areas are close to cortical tissue that is active when using objects and perceiving motion. The authors conclude that the brain circuitry underlying the conceptual representation of objects includes regions that are particularly well suited for the processing of their most salient meaning aspects (perceptual, functional). This evidence is largely compatible with an analysis of the lesion data of patients with disorders in the identification of living things vs. man-made artefacts. On the basis of a review of the lesion data of the known cases, Gainoti et al. (1995) concluded that the lesion distribution of these two patient types suggests a dominance of areas for visual object processing vs. areas that are especially important for somato-sensory and motor functions.

Hauk et al. (2004) took advantage of the somatotopic organization of the motor cortex to investigate the representation of action verbs. Subject read verbs describing actions performed with the feet, hands or face (e.g., 'kick', 'pick', 'lick'). Subsequently, they performed simple actions with foot, finger or tongue, which activated primary and premotor cortex in a somatotopic fashion, as expected. Interestingly, reading action verbs led to a similar somatotopic pattern of activation. Overlap between parts of (pre)motor cortex activated by action verbs and by action production was clearly observed for two of the three effectors (for a review of the imaging literature, see, Willems and Hagoort, 2007).

In conclusion, imaging studies provide evidence for the activation of premotor cortex in response to action words. Overall, the neural evidence seems to be compatible with the distinction between perceptual and functional attributes as an important metric for semantic categorization (for alternative accounts, see, Martin, 2007). Linguistic accounts of meaning are relevant in this regard as well. Jackendoff (1987, 1996, 2002) argues that in semantics one has to make a distinction between conceptual structure and spatial structure. Conceptual structure encodes abstract components such as category membership (a daffodil is a plant) and predicate-argument structure. Spatial structure encodes the physical features of objects referred to by concrete nouns, specified in the formats of the different sensorimotor systems. For the visual aspects, this includes an abstract visual-geometric description (a 3D model, in Marr's terminology), reflecting "the intuition that knowing the meaning of a word that denotes a physical object involves in part knowing what such an object looks like" (Jackendoff, 1987, p. 104). A similar idea is present in some instances of prototype theory in which word meaning representations contain an image of a stereotypical instance. But the proposed 3D model is a more abstract representation in which objects are spatially decomposed into parts and subparts in a viewer-independent orientation. In line with Allport's distributed model of conceptual information (Allport, 1985; Saffran and Sholl, 1999), Hagoort (1998) has suggested that this decomposition of word meanings into a restricted set of conceptual features, paired with an abstract visual description (a 3D model) has to be extended with matched pairs of conceptual structure and nonvisual sensory models (auditory, haptic), and with a model of action specified in a format that is tailored to the requirements of the motor system (Barsalou et al., 2003). But clearly, not all lexical items have a sensorimotor structure. Abstract concepts such as *freedom* and *ethics*, or logical concepts such as *if* and *not* have a conceptual structure only. It is hard to see how their meaning could be derived from motor system properties.

As we have seen above, an explicit account of meaning indicates that action words are not a unique class. Many concrete words have spatial properties as part of their semantics. For some of these words the visual characteristics are crucial for the identification of their meaning, for others it is the auditory features. In addition, there are words whose functional features are most distinctive. It is mostly these words that activate motor areas. But this does not provide any evidence that the mirror neuron system has a special or exclusive role for establishing parity of meaning between sender and receiver.

Finally, activation of motor areas during speech perception found with TMS (Fadiga et al., 2002; Watkins et al., 2003) or fMRI (Pulvermuller et al., 2006), is often cited as evidence for the motor theory of speech perception and, more generally, the mirror neuron system in humans. The reasoning is that if speech comprehension results in activation of areas involved in speech production, this is evidence that comprehension requires the involvement of the language production system. However, this conclusion does not follow. Even theorists that assume separate form representations for speech production and speech perception, do not deny that they are linked and often recruited in each others service. For instance, in speaking there is good evidence that we rely on the comprehension system to monitor our own speech (Levelt, 1989; Levelt et al., 1999). Likewise, the speech production system is presumably recruited during comprehension when the contextual constraints are high. A series of recent studies has shown that lexical items are predicted in comprehension even before any acoustic input has been present (Van Berkum et al., 2005). During this prediction, not only the semantic and syntactic properties of the upcoming item are activated, but even its phonological form seems to be retrieved (DeLong et al., 2005). The most parsimonious explanation is that the comprehension system recruits the speech production system when the contextual constraints are strong (Pickering and Garrod, 2007).

In summary, neither for speech recognition nor for the semantics of words, there is convincing evidence for a unique and exclusive relation between language and motor systems.

3. The evolutionary claim

It has been argued that the presence of mirror neurons in macaques' brains might untie the gordian knot of human language evolution, a notoriously intractable evolutionary problem given its apparent uniqueness. The claim is that "[...] the primate mirror system for grasping evolved into a key component of the mechanisms that render the human brain language-ready" (Arbib, 2005). This general claim is based on at least two inductive inferences. First, Rizzolatti and Arbib (1998) have claimed that macaque area F5 (Matelli et al., 1985), the premotor region where mirror neurons were first discovered (di Pellegrino et al., 1992b), is homologous with human Broca's area, namely Brodmann areas (BA) 44 and 45 (Brodmann, 1909). When put together with the claim that Broca's area controls language, these two premises generate the inference that macaque area F5 is the *anatomical* precursor of a language area (Rizzolatti and Arbib, 1998). Second, it is claimed that macaque area F5 is involved in controlling hand movements (Fogassi et al., 2001; Rizzolatti et al., 1988). When put together with the claim that language derives from hand gestures, it follows that macaque area F5 is the *functional* precursor of a language area. In the following section we will provide a review of the available evidences supporting the premises of these inferences, testing whether this evolutionary conjecture is consistent with current neurobiological findings.

3.1. Is macaque area F5 the anatomical precursor of human area 44?

Evolutionary homology between a feature occurring in two related species implies that the feature has been inherited from the latest common ancestor of the two species ("symplesiomorphy", Eldredge and Cracraft, 1980). Given the lack of evidence for the presence of mirror neurons in a premotor region in any common ancestor of humans and macaques, it appears at least premature to claim an evolutionary homology between macaque area F5c (the specific portion of area F5 where mirror neurons are localized in macaques, Fogassi et al., 2001) and human BA 44–45 (Brodmann, 1909). This homology was suggested by Rizzolatti and Arbib (1998), and since then it has been held as a proven matter by numerous reports dealing with the mirror neuron system. In fact, the available evidence appears to point against the possibility of an homology between macaque area F5c and human BA 44–45, given that macaques already have an area 44. In macaque it is this area 44, and not area F5c, that is functionally, cytoarchitectonically, and hodologically similar to human BA 44 (Petrides et al., 2005). In fact, macaque area F5c is part of macaque BA 6 (Geyer et al., 2000), and it appears parsimonious to assume that the same relationship holds for humans (Toni et al., 2001). More generally, it should be emphasized that claims of homology depends on the hierarchical level at which two structures are compared (wings in birds and bats are analogous as flying-enabling devices, but homologous as forelimbs in the tetrapods body plan). In the context of this commentary, the claim of homology between macaque area F5c and human BA 44–45 appears to be further undermined by the unlikely hierarchical correspondence between coarse cytoarchitectonic conglomerates (i.e., Brodmann's areas) and fine-grained anatomical elements (like F5c).

Recently, the discrepancy between the suggestion of (Rizzolatti and Arbib, 1998) and the findings of (Petrides et al., 2005) has been accommodated by introducing a new partition of macaque area F5 (Nelissen et al., 2005), a partition that closely overlaps with the location and the cytoarchitectonic characteristics of macaque area 44 (Petrides et al., 2005; Petrides and Pandya, 2002). This new partition, called area F5a (Nelissen et al., 2005), is rostral to both F5c (where precentral mirror neurons were found, Fogassi et al., 2001) and F5p (also known as F5ab or F5-bank, where visuomotor grasping-related neurons were found Fogassi et al., 2001; Rizzolatti and Luppino, 2001). Accordingly, macaque area F5a (as defined by Nelissen et al., 2005) appears to be part of macaque area 44 (as defined by Petrides et al., 2005). F5a shows increases in blood flow during observation of hand movements, either biological or mechanical, and during observation of *pantomimed* hand movements (Nelissen et al., 2005). In the original reports describing the properties of mirror neurons (di Pellegrino et al., 1992a; Gallese et al., 1996; Rizzolatti et al., 1996), particular emphasis was put on the absence of responses when an experimenter was pantomiming a grasping action (see, Fig. 4B of Gallese et al., 1996, and the following quote from di Pellegrino et al., 1992: "Control testing for the specificity of the hand-object interrelations showed that movements of the experimenter's hand alone or the combined movements of hand and object spatially separated one from the other were not effective in triggering neurons of all classes"). Given that macaques are not known for their pantomiming skills, this observation was important to support the conclusion that mirror neurons present matched perceptual and motor properties. By the same token, given that area F5a is metabolically active during observation of pantomimed hand movements (Nelissen et al., 2005), it remains unclear how the neurons in macaque area F5a/44 could support the same perceptuo-motor mirroring ascribed to mirror neurons. In other words, how can neurons in macaque area F5a/44 generate an "[...] automatically induced motor representation of the observed action" (Rizzolatti and Craighero, 2004) if those very neurons cannot motorically support a behaviour the animal cannot produce? Rizzolatti and colleagues argue that "[...] the basic essence of grasping appears to be coded here" (i.e., F5a) (Nelissen et al., 2005), but how can these platonic properties support a sensorimotor simulation based on the animal's motor repertoire when the neurons in macaque F5a respond to movements the animal is not capable of performing? A more prosaic interpretation of the properties of macaque area F5a/44 would consider the known involvement of this region in extracting arbitrary statistical regularities in input/output patterns (Forkstam et al., 2006; Grol et al., 2006; Wallis and Miller, 2003). It could be argued that this suggestion is considerably more complex than what the mirror neuron system theorists have proposed: "The proposed mechanism is rather simple. Each time an individual sees an action done by another individual, neurons that represent that action are activated in the observer's premotor cortex. This automatically induced, motor representation of the observed action corresponds to that which is spontaneously generated during active action and whose outcome is known to the acting individual. Thus, the mirror neuron system transforms visual information into knowledge" (Rizzolatti and Craighero, 2004). However, although this latter hypothesis might be simple to describe verbally, it does not specify how to map the visual information obtained from an observed action into the motor system of the observer (Csibra, 2004) [This was done by (Oztop and Arbib, 2002) for the case of grasping].

Is there an alternative account of the role of F5a/44? Here is what we believe to be a more viable account. In our view the sensorimotor mapping performed in F5a/44 is not centered on the ambiguous surface structure of a specific behaviour (Jacob and Jeannerod, 2005), but rather on more abstract (i.e., generalizable and learnable) properties of sensorimotor events. In other words, F5a/44 might "transform visual information into knowledge" not by virtue of a sensorimotor transformation tied to a particular sensory or effector system (a "motor representation of the observed

action"), that would entail a precise spatio-temporal relationship between stimuli and responses (as those required by grasping an object, for instance). Rather, F5a/44 might play a role in the guidance of behaviour according to learned rules, a role that fits with the crucial contribution of the inferior frontal region to the learning of arbitrary sensorimotor mappings (Nixon et al., 2004). This suggestion allows for mapping a variety of sensory features into multiple effector systems. In this sense, it appears more parsimonious than postulating ad hoc mirror neurons for any given sensorimotor scenario, with the associated computational explosion if a mirror neurons mechanism were to "transform visual information into knowledge". Irrespectively of these theoretical considerations, recent empirical findings support the notion that supposedly "mirror" responses depend on learned and arbitrary contingencies between stimuli and responses (Catmur et al., 2007; Kohler et al., 2002), as hypothesized by Oztop and Arbib (2002). Incidentally, this interpretation is also in line with the processing of hierarchically structured behaviours into superordinate representations by means of unification operations that have been suggested for human area 44 (Hagoort, 2005; Koechlin and Jubault, 2006). This processing might well be performed by the same neurons contributing to the hemodynamic signal described in Nelissen et al. (2005), without the need to attribute mirror properties to them.

In summary, the claim of an homology between macaque area F5 and human area 44, given the available evidence, appears unwarranted, if not unlikely. Furthermore, given the proliferation of F5 subareas with different anatomical and functional properties, it has become unclear which particular partition of macaque area F5 is supposed to be homologous with human area 44. These considerations cast doubts on the plausibility of an evolutionary link between precentral mirror neurons in macaques and human area 44, in particular a neural link based on the control of surface behaviours like grasping.

There is also an implicit claim at the basis of the contention that F5 is the *anatomical* precursor of a language area (Rizzolatti and Arbib, 1998), namely the apparently uncontroversial claim that Broca's area controls language. Yet, it is known that lesions restricted to Broca's area are neither necessary nor sufficient for the manifestation of deficits in language production or comprehension. These lesions simply do not result in a lasting aphasia (Dronkers et al., 2004; Mohr et al., 1978). Furthermore, it has become evident that, within the left inferior prefrontal cortex, Broca's area refers to a conglomerate of related but cytoarchitectonically distinct areas with a responsivity to distinct information types within the domains of language comprehension and production (Hagoort, 2005). Taken together, these observations indicate that Broca's area operates in the context of a larger cerebral network supporting human linguistic abilities, and that, even in an intact brain, its functional contributions are far from homogeneous. It is not immediately clear how to reconcile these observations with the claim of a direct line of descent between F5 mirror neurons and one (or more?) of the language faculties implemented in Broca's area.

3.2. Is macaque F5 the functional precursor of Broca's area?

As mentioned above, the evolutionary link between mirror neurons and language has also been related to the claim that F5 is the *functional* precursor of a language area (Arbib, 2005). This claim is based on two arguments. One is that macaque area F5 controls hand movements (Fogassi et al., 2001; Rizzolatti et al., 1988). This has been demonstrated by showing that local inactivation of F5p/F5ab markedly impairs the pre-shaping of the hand during a grasping movement (Fogassi et al., 2001; Rizzolatti and Luppino, 2001). However, the same experiment has also shown that mirror neurons (in F5c) are not necessary for grasping: "[…] after inactivation of F5 convexity, the hand shaping during reaching-to-grasp move-

ments was normal" (Fogassi et al., 2001). The intact grasping behaviour of the animal after inactivation of grasping mirror neurons raises the issue of their relevance for the actual production of an action. This is particularly important when considering that the human neuroimaging studies used to define the mirror neuron system have been based on passive observation of visual displays (reviewed in Rizzolatti and Craighero, 2004). Two recent studies have addressed this issue in humans (Heiser et al., 2003; Pobric and Hamilton, 2006), applying repetitive transcranial magnetic stimulation (rTMS) over the pars opercularis of the inferior frontal gyrus (this region shows a degree of overlap with Brodmann area 44, Amunts et al., 1999). These studies reported impairments on a weight-judgement task (Pobric and Hamilton, 2006), and in an imitation task (Heiser et al., 2003), but not in a matched motor execution task (Heiser et al., 2003). Therefore, these studies would suggest that human BA 44 is not relevant for the actual production of an action, but rather for assessing the kinematics of observed movements. These observations confirm the notion that macaque area F5c and human BA 44 are not necessary for controlling hand movements. It remains unclear how the clear dissociation between visuomotor and perceptual processes revealed by these interference studies (Fogassi et al., 2001; Heiser et al., 2003; Pobric and Hamilton, 2006) can support the notion that action understanding is supported by a sensorimotor matching between observation and execution occurring within the mirror neurons (Rizzolatti et al., 1999). Finally, these findings raise the collateral but important issue of the nature of the impairment evoked by rTMS of the left inferior frontal gyrus. The studies indicates that the subjects were impaired in assessing the kinematics of observed movements, but the results of these studies cannot be taken as evidence that the human mirror neuron system "understands" an action by recognizing its standard motor outcome (Gallese et al., 2004).

In summary, while it is clear that macaque area F5p is necessary for performing grasping movements, the available evidence indicates that visuomotor control and perceptual properties of area F5 are organized along functionally segregated circuits. This evidence weakens the hypothesis of a link between mirror neuron phenomena and visuomotor control.

There is a second argument at the basis of the contention that F5 is the *functional* precursor of a language area (Arbib, 2005). The claim is that language derives from hand gestures (Arbib, 2005; Corballis, 2003; Gentilucci and Corballis, 2006). Theoretically, this account partially overlaps with the issue discussed in the first section of this paper, and it appears to be marked by the same confusion between structural and symbolic aspects of language. This confusion inevitably percolates through the empirical findings. For instance, there is substantial evidence on neurophysiological links between manual and oral movements: performing or observing grasping movements towards objects of different size modulates vocal production (Gentilucci et al., 2001; Gentilucci, 2003), and silent vocalizations modulate kinematic parameters of grasping movements (Gentilucci et al., 2000; Glover et al., 2004). While these findings might be extremely relevant for understanding neurophysiological and cognitive characteristics of speech, it is unclear how they can tell us anything about language, let alone the evolution of language.

Finally, it is a long known but often forgotten difference between man and monkey that only the first has established a neuroanatomical pathway that allows voluntary control over the vocal apparatus (larynx, tongue, lips) (Ploog, 1979, 1988). In man, there is a direct connection between the primary motor cortex and the laryngeal motorneurons, whereas such a connection is not found in monkey. In addition, homo sapiens has dedicated an enlarged region of the spinal cord to be involved in voluntary control over breathing which is required for speech production (MacLarnon and Hewitt, 1999). If speech had gesture as its evolutionary

precursor, it remains to be explained why all these evolutionary expensive changes took place just for shifting the modality of communication from hands to mouth. Why would not a full-fledged sign language have developed instead of invoking a uniquely human speech system that has no more communicative power than a sign language (Emmorey, 2005)? The evolutionary cost–benefit function does not seem to favour the shift from gesture to speech.

Taken together, the theoretical considerations and the empirical findings presented in this section do not appear to provide unconditional support for the evolutionary scenario suggested by Rizzolatti and Arbib (1998). First, the evidence points to the anatomical and functional heterogeneity of both macaque area F5 and human Broca's area, casting serious doubts on the value and validity of claiming homology between these regions (see also Arbib and Bota, 2003). Second, while there is evidence of links between mirror-related phenomena and the processing of sensorimotor regularities, it is not immediately obvious how these could be relevant for linguistic phenomena. Finally, it is not clear why speech should have been the end product of this evolutionary scenario, instead of a fully-fledged sign language.

4. Is human communication based on motor resonance?

As discussed in the first section, the motor theory of speech perception has been recently revived and elaborated in the light of the discovery of mirror neurons (Arbib, 2005; Rizzolatti and Arbib, 1998). This elaboration is not trivial, since the claim is that automatic sensorimotor couplings (motor resonances, Rizzolatti et al., 1999) are the basis for solving a semantic and pragmatic parity problem – "[...] actions done by other individuals become messages that are understood by an observer without any cognitive mediation" (Rizzolatti and Craighero, 2004) [Note that other authors, like Arbib (2005) do not endorse the notion of motor resonance]. The hypothesis is that this problem is solved by the mirror neuron system by means of a simulation mechanism that uses the same inputs as those involved in motor control (Gallese et al., 2004; Rizzolatti and Craighero, 2004). In other words, a receiver can establish common ground with a sender by processing the communicative actions of the latter through the same predictive model used for the motor commands acting on one's own body (Wolpert et al., 2003). The appeal of this explanatory framework lies in its logical simplicity and in the possibility to link it to a specific neurobiological mechanism like the mirror neurons. For instance, the sensorimotor matching mechanism supported by the mirror neuron system might be adequate for reflexively sharing perceptuo-motor states between two agents, without considering how an observer could interpret a given action (Jacob and Jeannerod, 2005; Tomasello and Carpenter, 2007). These one-to-one mappings between physical and semantic properties of communicative behaviours have been documented in various species, and they involve broadcasted messages (Cheney and Seyfarth, 1990; Dyer, 2002; Mather, 2004). In these one-to-one mappings, the broadcasted message is unambiguously associated with a particular communicative effect. The mental representation of this type of actions predicts their physical outcome (Umilta et al., 2001), and it follows a mind-to-world direction of causation (Jacob and Jeannerod, 2003). Could this cognitive and neurobiological mechanism account for human communication? We believe it does not. There is more to human communication than emotional contagion, alarm calls, and social manipulation (Tomasello, 2006; Tomasello and Carpenter, 2007). Human communicative actions are tailored to a particular individual, taking into accounts the specific knowledge of the receiver (Clark, 1985). Framing human communication in terms of shared codes (Rizzolatti and Craighero, 2004) misses the point of its inferential nature (Levinson, 2000; Sperber and Wilson,

2001). The nature of these inferences cannot be exhausted within a one-to-one "mirror" matching between physical and semantic properties of an action, since the mapping between a communicative action and its intention is many-to-many rather than one-to-one (Levinson, 1983, 2000). In other words, the form of a communicative action underdetermines its content, and the contextual information required to resolve this ambiguity is to be found in the common ground between sender and receiver (Clark, 1996), i.e., the sender knowledge of the receiver knowledge (and possibly higher-order levels of this meta-knowledge). Accordingly, the goal of communicative actions is not to evoke a perceptual or a sensorimotor state, but to change the mental state of the receiver. Elaborating on the notation of Searle (1983) and Jacob and Jeannerod (2003), communicative actions have a $mind_{sender}$ to $mind_{receiver}$ direction of causation. It remains unclear how an automatic sensorimotor resonance between perception and action can account for these crucial features of human communication.

It might be argued that these are armchair arguments, and that the available empirical evidence points to the mirror neuron system as the biological substrate linking action generation and intention recognition (Fogassi et al., 2005; Iacoboni et al., 2005). These studies appear relevant for assessing how contextual information can be linked to kinematic features of a movement, such that it becomes possible to compute a motor program given an observable goal (Csibra, 2004; Jacob, 2008). Note that this inference is substantially different from the suggestion that mirror neurons extract a goal from an observable movement (Fogassi et al., 2005; Iacoboni et al., 2005). In fact, the latter inference appears computationally intractable (there are just too many goals that could be achieved with a given movement, given a purely sensorimotor inference), whereas the former might be computed (on the basis of the subjects past experience and motor repertoire). More importantly, the studies of Iacoboni et al. (2005) and Fogassi et al. (2005) do not deal with the generation of a communicative intention, i.e., how we organize our behaviour for conveying communicative intentions to a receiver. This is an important issue. If one takes seriously the inferential nature of human communication, then it follows that the sender of a message needs to generate a prediction about how a communicative intention will be inferred from an observable behaviour (de Ruiter et al., 2007). This generative process needs to start from the communicative intention of the sender, and this communicative intention cannot be an externally-driven mirroring of another individual's brain. In other words, when dealing with human communication, simply producing an action that an observer will automatically interpret by virtue of motor contagion is not an option (Jacob and Jeannerod, 2005). How can an automatic sensorimotor resonance distinguish between the reflexive and the meaningful cough (Levinson, 2006)? The mirror neuron system might account for the automatic interpretation of the sensorimotor outcome of an observed action (Fogassi et al., 2005; Iacoboni et al., 2005), but this says nothing about the crucial issue of how to arrange these functional aspects of an action such that they become tailored to alter the mental state of a particular receiver in a particular way. Moreover, sensorimotor resonances rely on previous shared experience, and they cannot explain the fast and successful interpretation of signal-to-meaning mappings that a receiver has never encountered before (Galantucci, 2005; Goldin-Meadow, 2003).

In a recent fMRI study we have directly addressed this issue (Noordzij et al., submitted for publication), asking pairs of subjects to participate in a real, on-line form of interaction for which no established communicative conventions were available. We found that planning novel communicative acts and understanding the communicative intention of these acts relied on the same cerebral tissue, namely the posterior part of the superior temporal sulcus (pSTS) of the right hemisphere. This activity was modulated by

the communicative difficulty of each trial, while remaining indifferent to sensory and motor events. Furthermore, the pSTS was embedded within the "mentalizing" network (Frith and Frith, 2006), a cerebral network clearly outside the motor system. These empirical observations, together with the theoretical considerations presented above, support the notion that human communication relies on cognitive processes operating on conceptual knowledge, rather than (or in addition to) sensorimotor resonances. This evidence points to the inadequacy of mirror neurons as the unique neural basis for human symbolic communication (see also, Csibra, 2004; Hurford, 2004; Jacob, 2008). [For a recent update on possible links between mirror neurons and human communication, see Arbib, 2008 (this issue)].

5. Conclusion

In this paper, we have presented a critical review of some claims that have gained widespread popularity following the discovery of mirror neurons in macaques and of similar effects in humans (Rizzolatti et al., 2006). We have provided empirical and theoretical arguments against three particular claims that use sensorimotor resonance as a mechanism for bringing language comprehension, language evolution, and human communication within the realm of the motor system. The available empirical evidence does not appear to support these claims, and their theoretical scope fails to account for some crucial features of the phenomena they are supposed to explain.

We are not arguing against important links and interactions between language and motor systems in the brain. For the semantics of certain word classes the action system might be invoked. Clearly at the level of phonetic realizations the motor system plays a crucial role. This could even be in language comprehension when under condition of high predictability the predicted word forms might actually be produced in the form of internal speech. However, all this does not imply that the highly complicated communication system of natural language can be fully reduced to sensorimotor properties and the contribution of sensorimotor areas. We, therefore, fully agree with the conclusion by Jacob and Jeannerod (2005) that "the motor properties of the mirror system are well designed for representing an agent's motor intention involved in an object-oriented action, not for representing an agent's social intention, let alone his communicative intention" (p. 24). Language goes beyond action. Without denying the enormous importance of the discovery of mirror neurons, their explanatory power for understanding human communication is limited. More generally, one might wonder about the scientific mileage provided by a theory that focuses on close links between surface behaviours (like grasping) and individual neurons (like mirror neurons) to account for cognitive events (like action understanding). Surface behaviour has been successfully used as a tool for studying mental processes, as clearly illustrated by psycholinguistic and neuropsychological research. However, the faculties ascribed to mirror neurons appear to go a step further, reducing mental processes to surface behaviour, with neurons able to encode "the basic essence of grasping" (Nelissen et al., 2005). We believe that surface behaviours are unlikely to provide a comprehensive metric for understanding mental processes, let alone highly non-linear, non-stationary, and multi-dimensional neuronal computations (Koch and Laurent, 1999).

Acknowledgements

References

Allport, D.A., 1985. Distributed memory, modular subsystems and dysphasia. In: Newman, S.K., Epstein, R. (Eds.), Current Perspectives in Dysphasia. Churchill Livingstone, Edinburgh, pp. 32–60.

Amunts, K., Schleicher, A., Burgel, U., Mohlberg, H., Uylings, H.B., Zilles, K., 1999. Broca's region revisited: cytoarchitecture and intersubject variability. J. Comp. Neurol. 412, 319–341.

Arbib, M.A., 2005. From monkey-like action recognition to human language: an evolutionary framework for neurolinguistics. Behav. Brain Sci. 28, 105–124.

Aziz-Zadeh, L., Wilson, S.M., Rizzolatti, G., Iacoboni, M., 2006. Congruent embodied representations for visually presented actions and linguistic phrases describing actions. Curr. Biol. 16, 1818–1823.

Barsalou, L.W., Kyle Simmons, W., Barbey, A.K., Wilson, C.D., 2003. Grounding conceptual knowledge in modality-specific systems. Trends Cogn. Sci. 7, 84–91.

Brodmann, K., 1909. Vergleichende Lokalisationslehre der Großhirnrinde.

Catmur, C., Walsh, V., Heyes, C., 2007. Sensorimotor learning configures the human mirror system. Curr. Biol. 17, 1527–1531.

Cheney, D.L., Seyfarth, R.M., 1990. How Monkeys See the World: Inside the Mind of Another Species. University of Chicago Press, Chicago.

Clark, H.H., 1985. Language use and language users. In: Lindsey, G., Aronson, E. (Eds.), The Handbook of Social Psychology. Harper & Row, New York.

Clark, H.H., 1996. Using language. Cambridge University Press, Cambridge England.

Corballis, M.C., 2003. From mouth to hand: gesture, speech, and the evolution of right-handedness. Behav. Brain Sci. 26, 199–208.

Csibra, G., 2004. Mirror neurons and action observation. Is simulation involved? Publication of the European Science Foundation. Available from: <http://www.interdisciplines.org/mirror/papers/4>.

Cutler, A., Clifton, C.E., 1999. Comprehending spoken language: a blueprint of the listener. In: Brown, C.M., Hagoort, P. (Eds.), The Neurocognition of Language. Oxford University Press, Oxford, pp. 123–166.

DeLong, K.A., Urbach, Th.P., Kutas, M., 2005. Probabilistic word pre-activation during language comprehension inferred from electrical brain activity. Nat. Neurosci. 8, 1117–1121.

de Ruiter, J.P., Noordzij, M.L., Newman-Norlund, S., Hagoort, P., Toni, I., 2007. On the origin of intentions. In: Haggard, P., Rossetti, Y., Kawato, M. (Eds.), Attention and Performance XXII: Sensorimotor foundation of higher cognition. Oxford University Press, Oxford, pp. 593–610.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., Rizzolatti, G., 1992a. Understanding motor events: a neurophysiological study. Exp. Brain Res. 91, 176–180.

di Pellegrino, G., Fadiga, L., Fogassi, L., Gallese, V., Rizzolatti, G., 1992b. Understanding motor events: a neurophysiological study 2. Exp. Brain Res. 91, 176–180.

Diehl, R.L., Lotto, A.J., Holt, L.L., 2004. Speech perception. Annu. Rev. Psychol. 55, 149–179.

Dronkers, N.F., Wilkins, D.P., Van Valin, R.D., Redfern, B.B., Jaeger, J.J., 2004. Lesion analysis of the brain areas involved in language comprehension. Cognition 92, 145–177.

Dyer, F.C., 2002. The biology of the dance language. Ann. Rev. Entomol. 47, 917–949.

Eldredge, N., Cracraft, J., 1980. Phylogenetic Patterns and the Evolutionary Process: Method and Theory in Comparative Biology. Columbia University Press, New York.

Emmorey, K., 2005. Sign languages are problematic for a gestural origins theory of language evolution. Behav. Brain Sci. 28, 130–131.

Fadiga, L., Craighero, L., Buccino, G., Rizzolatti, G., 2002. Speech listening specifically modulates the excitability of tongue muscles: a TMS study. Eur. J. Neurosci. 15, 399–402.

Farah, M.J., McClelland, J.L., 1991. A computational model of semantic memory impairment: modality specificity and emergent category specificity. J. Exp. Psychol. Gen. 120, 339–357.

Fogassi, L., Gallese, V., Buccino, G., Craighero, L., Fadiga, L., Rizzolatti, G., 2001. Cortical mechanism for the visual guidance of hand grasping movements in the monkey: a reversible inactivation study. Brain 124, 571–586.

Fogassi, L., Ferrari, P.F., Gesierich, B., Rozzi, S., Chersi, F., Rizzolatti, G., 2005. Parietal lobe: from action organization to intention understanding. Science 308, 662–667.

Forkstam, C., Hagoort, P., Fernandez, G., Ingvar, M., Petersson, K.M., 2006. Neural correlates of artificial syntactic structure classification. Neuroimage 32, 956–967.

Foss, D.J., Gernsbacher, M.A., 1983. Cracking the dual code: towards a unitary model of phoneme identification. J. Verbal Learn. Verbal Behav. 22, 609–632.

Frith, C.D., Frith, U., 2006a. How we predict what other people are going to do. Brain Res. 1079, 36–46.

Frith, C.D., Frith, U., 2006b. The neural basis of mentalizing. Neuron 50, 531–534.

Gainoti, G., Azzoni, A., Lanzillotta, M., Marra, C., Razzano, C., 1995. Some preliminary findings concerning a new scale for the assessment of depression and related symptoms in stroke patients. Ital. J. Neurol. Sci. 16, 439–451.

Galantucci, B., 2005. An experimental study of the emergence of human communication systems. Cogn. Sci. 29, 737–767.

Galantucci, B., Fowler, C.A., Turvey, M.T., 2006. The motor theory of speech perception reviewed. Psychon. Bull. Rev. 13, 361–377.

Gallese, V., Fadiga, L., Fogassi, L., Rizzolatti, G., 1996. Action recognition in the premotor cortex. Brain 119 (Pt 2), 593–609.

Gallese, V., Keysers, C., Rizzolatti, G., 2004. A unifying view of the basis of social cognition. Trends Cogn. Sci. 8, 396–403.

Gentilucci, M., 2003. Grasp observation influences speech production. Eur. J. Neurosci. 17, 179–184.

Gentilucci, M., Corballis, M.C., 2006. From manual gesture to speech: a gradual transition. Neurosci. Biobehav. Rev. 30, 949–960.

Gentilucci, M., Benuzzi, F., Bertolani, L., Daprati, E., Gangitano, M., 2000. Language and motor control. Exp. Brain Res. 133, 468–490.

Gentilucci, M., Benuzzi, F., Gangitano, M., Grimaldi, S., 2001. Grasp with hand and mouth: a kinematic study on healthy subjects. J. Neurophysiol. 86, 1685–1699.

Geyer, S., Matelli, M., Luppino, G., Zilles, K., 2000. Functional neuroanatomy of the primate isocortical motor system. Anat. Embryol. (Berl.) 202, 443–474.

Glover, S., Rosenbaum, D.A., Graham, J., Dixon, P., 2004. Grasping the meaning of words. Exp. Brain Res. 154, 103–108.

Goldin-Meadow, S., 2003. The Resilience of Language: What Gesture Creation in Deaf Children can tell us about How All Children Learn Language. Psychology Press, New York, NY.

Grol, M.J., de Lange, F.P., Verstraten, F.A.J., Passingham, R.E., Toni, I., 2006. Cerebral changes during performance of overlearned arbitrary visuomotor associations. J. Neurosci. 26, 117–125.

Hagoort, P., 1998. The shadows of lexical meaning in patients with semantic impairments. In: Stemmer, B., Whitaker, H.A. (Eds.), Handbook of Neurolinguistics. Academic Press, New York, pp. 235–248.

Hagoort, P., 2005. On broca, brain, and binding: a new framework. Trends Cogn. Sci. 9, 416–423.

Hauk, O., Johnsrude, I., Pulvermuller, F., 2004. Somatotopic representation of action words in human motor and premotor cortex. Neuron 41, 301–307.

Heiser, M., Iacoboni, M., Maeda, F., Marcus, J., Mazziotta, J.C., 2003. The essential role of Broca's area in imitation. Eur. J. Neurosci. 17, 1123–1128.

Hurford, J.R., 2004. Language beyond our grasp: what mirror neurons can, and cannot, do for the evolution of language. In: Oller, D.K., Griebel, U. (Eds.), Evolution of Communication Systems. MIT Press, Cambridge, MA, pp. 297–314.

Iacoboni, M., Molnar-Szakacs, I., Gallese, V., Buccino, G., Mazziotta, J.C., Rizzolatti, G., 2005. Grasping the intentions of others with one's own mirror neuron system. PLoS. Biol. 3, e79.

Jackendoff, R., 1987. On beyond zebra: the relation of linguistic and visual information. Cognition 26, 89–114.

Jackendoff, R., 1996. Conceptual semantic and cognitive semantics. Cogn. Linguist. 7, 93–129.

Jackendoff, R., 2002. Foundations of Language: Brain, Meaning, Grammar, Evolution. Oxford University Press, Oxford, UK.

Jacob, P., 2008. What do mirror neurons contribute to human social cognition? Mind Lang. 23, 190–223.

Jacob, P., Jeannerod, M., 2003. Ways of Seeing: The Scope and Limits of Visual Cognition. Oxford University Press, Oxford.

Jacob, P., Jeannerod, M., 2005. The motor theory of social cognition: a critique. Trends Cogn. Sci. 9, 21–25.

Kluender, K.R., Diehl, R.L., Killeen, P.R., 1987. Japanese quail can learn phonetic categories. Science 237, 1195–1197.

Koch, C., Laurent, G., 1999. Complexity and the nervous system. Science 284, 96–98.

Koechlin, E., Jubault, T., 2006. Broca's area and the hierarchical organization of human behavior. Neuron 50, 963–974.

Kohler, E., Keysers, C., Umilta, M.A., Fogassi, L., Gallese, V., Rizzolatti, G., 2002. Hearing sounds, understanding actions: action representation in mirror neurons. Science 297, 846–848.

Kuhl, P.K., 1981. Discrimination of speech by nonhuman animals: basic auditory sensitivities conducive to the perception of speech-sound categories. J. Acoust. Soc. Am. 95, 340–349.

Kuhl, P.K., Miller, J.D., 1975. Speech perception by the chinchilla: voiced-voiceless distinction in alveolar plosive consonants. Science 190, 69–72.

Kuhl, P.K., Miller, J.D., 1978. Speech perception by the chinchilla: identification functions for synthetic VOT stimuli. J. Acoust. Soc. Am. 63, 905–917.

Levelt, W.J.M., 1989. Speaking: From Intention to Articulation. MIT Press, Cambridge, MA.

Levelt, W.J.M., Roelofs, A., Meyer, A.S., 1999. A theory of lexical access in speech production. Behav. Brain Sci. 22, 1–38.

Levinson, S.C., 1983. Pragmatics. Cambridge University Press, Cambridge, Cambridgeshire.

Levinson, S.C., 2000. Presumptive Meanings: The Theory of Generalized Conversational Implicature. MIT Press, Cambridge, Mass.

Levinson, S.C., 2006. On the human interactional engine. In: Enfield, N.J., Levinson, S.C. (Eds.), Roots of Human Sociality. Berg Publisher, New York, pp. 39–69.

Liberman, A.M., 1957. Some results of research on speech perception. J. Acoust. Soc. Am. 29, 117–123.

Liberman, A.M., Mattingly, I.G., 1985. The motor theory of speech perception revised. Cognition 21, 1–36.

Liberman, A.M., Delattre, P.C., Cooper, F.S., 1952. The role of selected stimulus variables in the perception of the unvoiced stop consonants. Am. J. Psychol. 65, 497–516.

Liberman, A.M., Delattre, P.C., Cooper, F.S., Gerstman, L.J., 1954. The role of consonant–vowel transitions in the perception of the stop and nasal consonants. Psychol. Monogr.: Gen. Appl. 68, 1–13.

Liberman, A.M., Delattre, P.C., Gerstman, L.C., Cooper, F.S., 1956. Tempo of frequency change as a cue for distinguishing classes of speech sounds. J. Exp. Psychol. 52, 127–137.

MacLarnon, A.M., Hewitt, G.P., 1999. The evolution of human speech: the role of enhanced breathing control. Am. J. Phys. Anthropol., 109.

Marslen-Wilson, W.D., Welsh, A., 1978. Processing interactions and lexical access during word-recognition in continuous speech. Cogn. Psychol. 10, 29–63.

Martin, A., 2007. The representation of object concepts in the brain. Annu. Rev. Psychol. 58, 25–45.

Martin, A., Caramazza, A., 2003. Neurophysiological and neuroimaging perspectives on conceptual knowledge: an introduction. Cogn. Neuropsychol. 20, 95–212.

Martin, A., Wiggs, C.L., Ungerleider, L.G., Haxby, J.V., 1996. Neural correlates of category-specific knowledge. Nature 379, 649–652.

Matelli, M., Luppino, G., Rizzolatti, G., 1985. Patterns of cytochrome oxidase activity in the frontal agranular cortex of the macaque monkey. Behav. Brain Res. 18, 125–136.

Mather, J.A., 2004. Cephalopod skin displays: from concealment to communication. In: Oller, D.K., Griebel, U. (Eds.), Evolution of Communication Systems. MIT Press, Cambridge, pp. 193–214.

McCarthy, R., Warrington, E.K., 1990. The dissolution of semantics. Nature 343, 599.

McKenna, P., Warrington, E.K., 1993. The neuropsychology of semantic memory. In: Boller, F., Grafman, J. (Eds.), Handbook of Neuropsychology, vol. 8. Elsevier, Amsterdam, pp. 193–213.

Mohr, J.P., Pessin, M.S., Finkelstein, S., Funkenstein, H.H., Duncan, G.W., Davis, K.R., 1978. Broca aphasia – pathologic and clinical. Neurology 28, 311–324.

Nelissen, K., Luppino, G., Vanduffel, W., Rizzolatti, G., Orban, G.A., 2005. Observing others: multiple action representation in the frontal lobe. Science 310, 332–336.

Nixon, P., Lazarova, J., Hodinott-Hill, I., Gough, P., Passingham, R., 2004. The inferior frontal gyrus and phonological processing: an investigation using rTMS. J Cogn. Neurosci. 16, 289–300.

Noordzij, M.L., Newman-Norlund, S.E., De Ruiter, J.P., Hagoort, P., Levinson, S.C., Toni, I., submitted for publication. Cerebral mechanisms underlying human communication.

Petrides, M., 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., Cadoret, G., Mackey, S., 2005. Orofacial somatomotor responses in the macaque monkey homologue of Broca's area. Nature 435, 1235–1238.

Pickering, M.J., Garrod, S., 2007. Do people use language production to make predictions during comprehension? Trends Cogn. Sci. 11, 105–110.

Pisoni, D.B., Luce, P.A., 1987. Acoustic-phonetic representations in word recognition. In: Frauenfelder, U.H., Tyler, L.K. (Eds.), Spoken Word Recognition. MIT Press, Cambridge, pp. 21–52.

Ploog, D., 1979. Phonation, emotion, cognition, with reference to the brain mechanisms involved. Ciba Found Symp., 79–98.

Ploog, D., 1988. An outline of human neuroethology. Human Neurobiol. 6, 227–238.

Pobric, G., Hamilton, A.F., 2006. Action understanding requires the left inferior frontal cortex. Curr. Biol. 16, 524–529.

Pulvermuller, F., Huss, M., Kheriff, F., Moscoso del Prado Martin, F., Hauk, O., Shtyrov, Y., 2006. Motor cortex maps articulatory features of speech sounds. Proc. Natl. Acad. Sci. USA 103, 7865–7870.

Rizzolatti, G., Arbib, M.A., 1998. Language within our grasp. Trends Neurosci. 21, 188–194.

Rizzolatti, G., Craighero, L., 2004. The mirror-neuron system. Annu. Rev. Neurosci. 27, 169–192.

Rizzolatti, G., Fogassi, L., Gallese, V., 2006. Mirrors in the mind. Sci. Am. 295, 54–61.

Rizzolatti, G., Luppino, G., 2001. The cortical motor system. Neuron 31, 889–901.

Rizzolatti, G., Camarda, R., Fogassi, L., Gentilucci, M., Luppino, G., Matelli, M., 1988. Functional organization of inferior area 6 in the macaque monkey. II. Area F5 and the control of distal movements. Exp. Brain Res. 71, 491–507.

Rizzolatti, G., Fadiga, L., Gallese, V., Fogassi, L., 1996. Premotor cortex and the recognition of motor actions. Brain Res. Cogn. Brain Res. 3, 131–141.

Rizzolatti, G., Fadiga, L., Fogassi, L., Gallese, V., 1999. Resonance behaviors and mirror neurons. Arch. Ital. Biol. 137, 85–100.

Saffran, E., Sholl, A., 1999. Clues to the functional and neural architecture of word meaning. In: Brown, C.M., Hagoort, P. (Eds.), The Neurocognition of Language. Oxford University Press, Oxford, pp. 241–273.

Searle, J.R., 1983. Intentionality. An Essay in the Philosophy of Mind. Cambridge University Press, Cambridge.

Sparks, D.W., Kuhl, p., Edmonds, A.E., Gray, G.P., 1978. Investigating the MESA (Multipoint Electrotactile Speech Aid): the transmission of segmental features of speech. J. Acoust. Soc. Am. 63, 246–257.

Sperber, D., Wilson, D., 2001. Relevance: Communication and Cognition. Blackwell Publishers, Oxford.

Tomasello, M., 2006. Why don't apes point? In: Enfield, N.J., Levinson, S.C. (Eds.), Roots of Human Sociality: Culture, Cognition and Interaction. Berg, London, pp. 506–524.

Tomasello, M., Carpenter, M., 2007. Shared intentionality. Dev. Sci. 10, 121–125.

Toni, I., Rushworth, M.F., Passingham, R.E., 2001. Neural correlates of visuomotor associations: Spatial rules compared with arbitrary rules. Exp. Brain Res. 141, 359–369.

Umilta, M.A., Kohler, E., Gallese, V., Fogassi, L., Fadiga, L., Keysers, C., Rizzolatti, G., 2001. I know what you are doing. A neurophysiological study. Neuron 31, 155–165.

Van Berkum, J.J.A., Brown, C.M., Zwitserlood, P., Kooijman, V., Hagoort, P., 2005. Anticipating upcoming words in discourse: evidence from ERPs and reading times. J. Exp. Psychol.: Learn. Memory Cogn. 31, 443–467.

Wallis, J.D., Miller, E.K., 2003. From rule to response: neuronal processes in the premotor and prefrontal cortex. J. Neurophysiol. 90, 1790–1806.

Warrington, E.K., McCarthy, R.A., 1987. Categories of knowledge. Further fractionations and an attempted integration. Brain 110 (Pt 5), 1273–1296.

Warrington, E.K., Shallice, T., 1984. Category specific semantic impairments. Brain 107 (Pt 3), 829–854.

Watkins, K.E., Strafella, A.P., Paus, T., 2003. Seeing and hearing speech excites the motor system involved in speech production. Neuropsychologia 41, 989–994.

Willems, R.M., Hagoort, P., 2007. Neural evidence for the interplay between language, gesture, and action: a review. Brain Lang. 101, 278–289.

Wolpert, D.M., Doya, K., Kawato, M., 2003. A unifying computational framework for motor control and social interaction. Philos. Trans. Royal Soc. London B: Biol. Sci. 358, 593–602.