Perspective

Schizophrenia and the mirror system: an essay[☆]

Michael A. Arbib[*], T. Nathan Mundhenk

Computer Science, Neuroscience, and the USC Brain Project, University of Southern California, Los Angeles, CA 90089-2520, USA

Abstract

We analyze how data on the mirror system for grasping in macaque and human ground the mirror system hypothesis for the evolution of the language-ready human brain, and then focus on this putative relation between hand movements and speech to contribute to the understanding of how it may be that a schizophrenic patient generates an action (whether manual or verbal) but does not attribute the generation of that action to himself. We make a crucial discussion between self-monitoring and attribution of agency. We suggest that vebal hallucinations occur when an utterance progresses through verbal creation pathways and returns as a vocalization observed, only to be dismissed as external since no record of its being created has been kept. Schizophrenic patients on this theory then confabulate the agent.
© 2004 Published by Elsevier Ltd.

Keywords: FARS model; Grasping; Mirror system; Schizophrenia; Agency

1. Introduction

How do we as humans know the agency of actions? In particular, how does one discriminate one's actions from that of another person? If I am a normal adult, when I move my hand, I know I moved it and also know that someone else did not move it. The same goes for speech and thought. Yet schizophrenics may commit actions such as verbal utterances and hand movements that they erroneously attribute to other agents, and they may erroneously attribute the actions of others to themselves.

The symptoms of schizophrenia include delusions, hallucinations, disorganized speech, disorganized or catatonic behavior, and what are typically referred to as negative symptoms (affective flattening, alogia, or avolition). Hallucinations can take on the form of a constant verbal running commentary or even include two or more voices conversing with each other (DSM-IV, 1994). In schizophrenia, the auditory verbal hallucinations (AVH) and movements of the hand that are attributed to other agents (delusions of influence) are of particular interest to us since (a) AVH seem to be generated in a manner very similar to regular speech production (Stephane, Barton, & Boutros, 2001) and are sometimes observable in the form of sublingual vocalizations, and (b) we have developed an account of the brain mechanisms of language that relates them strongly to the mirror system for grasping, i.e., the system that is active both when the agent is grasping and when the agent observes similar actions performed by others (Arbib, 2002, in press-a; Arbib & Bota, 2003; Rizzolatti & Arbib, 1998; see Jeannerod, Arbib, Rizzolatti, & Sakata, 1995, for background). We focus on this putative relation between hand movements and speech as we seek to contribute to the understanding of how it may be that a schizophrenic patient generates an action but does not attribute the generation of that action to himself.

In addition to having an understanding of what one is doing, it is also important to know what other people are doing. For this, we need both a notion of *action*, what is being done, and *agency*, who is doing it. Indeed, humans and many other animals have a way of placing themselves in the actions of others (Arbib, Billard, Iacoboni, & Oztop, 2000; Frith, 2002; Rizzolatti, Fadiga, Gallese, & Fogassi, 1996). In this action-mirror paradigm, I imagine myself moving my hands in the

---

[*] Corresponding author. Tel.: +1 213 740 9220; fax: +1 213 740 5687.
E-mail addresses: arbib@pollux.usc.edu (M.A. Arbib), mundhenk@email.usc.edu (T.N. Mundhenk).

way another person does, or I imagine saying something like another person. Additionally, I do this while the other person is executing their actions. However, to function effectively, I must nonetheless correctly "bind" the various actions to the appropriate agents. We propose that that the binding for actions that I commit, or actions that are directed to me, may involve processes partially separate from those involved in binding of actions to other agents. An example might be the observation that delusions in schizophrenia seem to directed at the patient, or from the patient to another actor. If all agents, including the self, were created equal, we would expect that schizophrenics would experience as many third person delusions (actor to actor) as first person delusions (actor to self/self to actor)

2. The mirror system for grasping

The framework for our essay is provided by the Mirror System Hypothesis, which links the control and recognition of manual actions to the brain mechanisms, which support the production and perception of language. We introduce this Hypothesis in the next section. But, first we introduce the mirror system for grasping in monkeys, and outline a model of its development.

The system of the monkey brain for visuomotor control of hand movements has its premotor outpost in an area called F5, which contains a set of neurons, called *mirror neurons*. Each such mirror neuron is active not only when the monkey executes a specific grasp but also when the monkey observes a human or other monkey execute a more-or-less similar grasp

(Rizzolatti et al., 1996). Thus, F5 in monkey contains a *mirror system for grasping* which employs a common neural code for *executed* and *observed* manual actions. In addition, F5 contains neurons active only for execution of grasps, but not for observation of the grasps of others. These are called *canonical neurons*. Fagg and Arbib (1998) developed the FARS model for the control of the canonical F5 neurons. This was a computational model, which was used to generate a variety of interesting simulations. It will help our understanding of the mirror system and its putative role in schizophrenia if we first analyze the FARS model and then "lift" it to the mirror system.

As we see in Fig. 1, areas cIPS provides visual input to parietal area AIP concerning the position and orientation of the object's surfaces. The job of AIP is then to extract the affordances the object offers for grasping (i.e., the visually grounded encoding of "motor opportunities" for grasping the object, rather than its classification). The basic pathway AIP → F5 canonical → F1 (primary motor cortex, also known as M1) of the FARS model then transforms the (neural code for) affordance to the appropriate motor schema (F5) and thence to the appropriate detailed descending motor control signals (F1).

Going beyond the empirical data then available, Fagg and Arbib (1998) stressed that in general, even when attention is focused on a single object, there may be several ways to grasp that object. The original FARS model thus hypothesized:

(a) that object recognition (mediated by inferotemporal cortex IT) can bias the computation of working memory and task constraints and the effect of instruction stimuli in various areas of prefrontal cortex (PFC), and
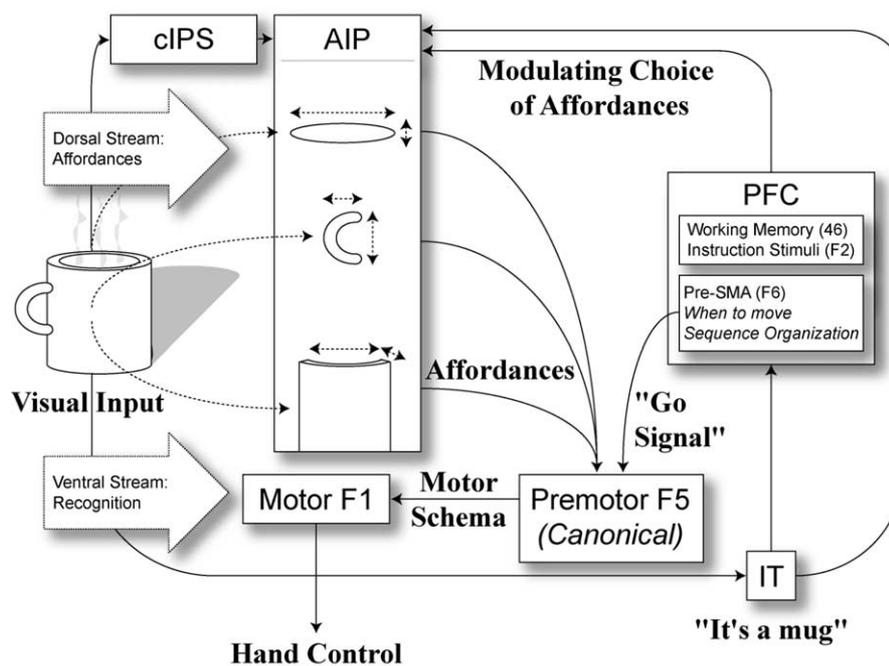

Fig. 1. FARS modificato: prefrontal systems modulate the choice of affordances via their influence on AIP. Other prefrontal influences determine whether potential actions encoded in premotor F5 will be released for execution (adapted from the original FARS figure of Fagg & Arbib, 1998).

(b) that strong connections between PFC and F5 provide the data for F5 to choose one affordance from the possibilities offered by AIP.

However, contra (b), anatomical evidence (reviewed by Rizzolatti & Luppino, 2001) was later found that demonstrated that connections from PFC to F5 in macaque are very limited whereas rich connections exist between prefrontal cortex and AIP. Furthermore AIP, unlike F5, receives direct input from IT (Webster, Bachevalier, & Ungerleider, 1994). Rizzolatti and Luppino (2003) thus suggested that FARS be modified so that information on object semantics and the goals of the individual directly influence AIP rather than F5. Thus, selection of an appropriate grip would occur in AIP by biasing those affordances that would lead to the grip appropriate to the individual's current intentions. In "FARS modificato" (Fig. 1), AIP still describes several affordances initially, but only one of these is selected to influence F5. This affordance then establishes in the F5 neurons a command which reaches threshold for the appropriate grip once it receive a "go signal" from F6 (pre-SMA) which (in concert with the basal ganglia) will determine whether external and/or internal contingencies allow the action execution. It is worth noting that this account associates *three* working memory systems with the canonical grasping system:

WM1: interactions between AIP and F5 keep track of current affordances in the environment;
WM2: area 46 or other PFC regions hold the location and related parameters of unattended or absent objects within the currently relevant environment (see Rolls & Arbib, 2003, for some of the relevant issues in scene perception);

WM3: the basal ganglia works with F6 to keep track of the place of the current action within some overall coordinated control program.
Later, we will suggest the importance of yet another working memory
WM4: a working memory, which holds information about aspects of the recently executed trajectory. This working memory decays rapidly over time.

Just as we have embedded the F5 canonical neurons in a larger system involving both the parietal area AIP and the inferotemporal area IT, so do we now stress that *the F5 mirror neurons are part of a larger mirror system that includes (at least) parts of the superior temporal gyrus (STS) and area PF of the parietal lobe*. We now discuss a model of this larger system, the MNS model (Fig. 2; Oztop & Arbib, 2002). (The reader may consult that paper both for modeling details and a review and references for the relevant neurophysiology.)

One path in Fig. 2 corresponds to the basic pathway AIP → F5 canonical → M1 of the FARS model (but MNS does not include the material on prefrontal influences). Another pathway (MIP/LIP/VIP → F4) completes the "canonical" portion of the MNS model, with intraparietal areas MIP/LIP/VIP providing object location information which enables F4 to instruct F1 to execute a reaching movement which positions the hand appropriately for grasping. The rest of Fig. 2 presents the core elements for the understanding of the mirror system. Mirror neurons do not fire when the monkey sees the hand movement or the object in isolation – it is the sight of the hand moving appropriately to grasp or otherwise manipulate a seen (or recently seen) object that is
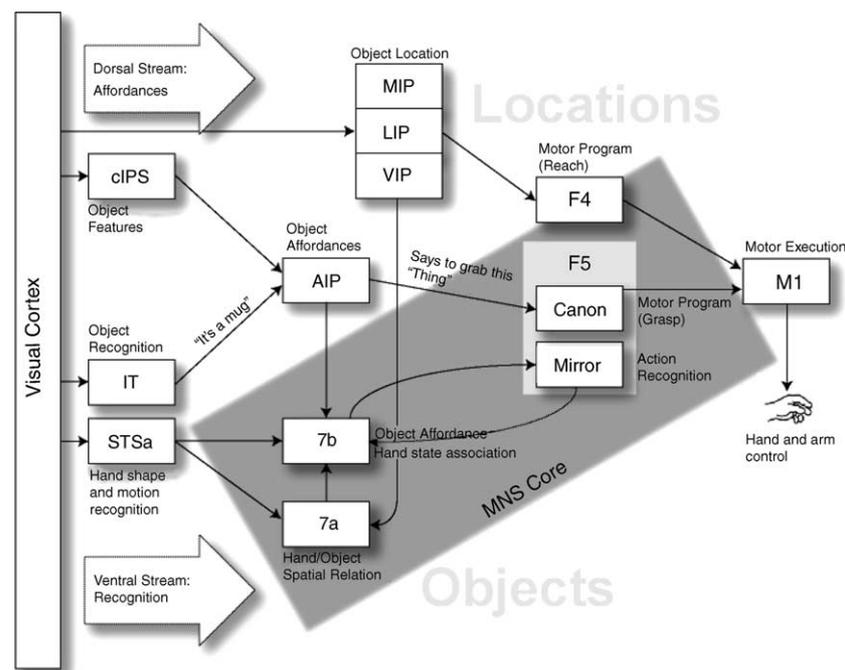

Fig. 2. A schematic view of the mirror neuron system. The mirror neuron system (MNS) model (Oztop and Arbib, 2002) focuses on the circuitry highlighted by the gray diagonal rectangle (note that this model omits the functionality of PFC included in the FARS model of Fig. 1).

required for the mirror neurons attuned to the given action to fire. This requires schemas for the recognition of both the shape of the hand and analysis of its motion (ascribed in the figure to STSa), and for analysis of the relation of these hand parameters to the location and affordance of the object (7a and 7b in Fig. 2; we identify 7b with PF in what follows).

In the MNS model, the *hand state* was defined as a vector whose components represented the movement of the wrist relative to the location of the object and of the hand shape relative to the affordances of the object. Oztop and Arbib (2002) showed that an artificial neural network corresponding to PF and F5$_{mirror}$ could be trained to recognize the grasp type from the *hand state trajectory*, with correct classification often being achieved well before the hand reached the object. The modeling assumed that the neural equivalent of a grasp being in the monkey's repertoire is that there is a pattern of activity in the F5 canonical neurons that commands that grasp. During training, the output of the F5 canonical neurons, acting as a code for the grasp being executed by the monkey at that time, was used as the training signal for the F5 mirror neurons to enable them to learn which hand-object trajectories corresponded to the canonically encoded grasps. Moreover, the input to the F5 mirror neurons encodes the trajectory of the relation of parts of the hand to the object rather than the visual appearance of the hand in the visual field. As a result of this training, the appropriate mirror neurons come to fire in response to the appropriate trajectories even when the trajectory is not accompanied by F5 canonical firing. What makes the modeling worthwhile is that the trained network responded not only to hand state trajectories from the training set, but also exhibited interesting responses to novel hand-object relationships. Despite the use of a non-physiological neural network, simulations with the model revealed a range of putative properties of mirror neurons that suggest new neurophysiological experiments.

This training prepares the F5 mirror neurons to respond to hand-object relational trajectories even when the hand is of the "other" rather than the "self" because the hand state is based on the movement of a hand relative to the object, and thus only *indirectly* on the retinal input of seeing hand and object – the latter can differ greatly between observation of self and other. However, the model only accepts input related to one hand and one object at a time, and so says nothing about the "binding" of the action to the agent of that action.

Recently, Miall (2003), building on the work of Iacoboni (in press) (Carr et al., 2003) has related mirror neurons to the notion of internal models. In fact, this relationship was first developed by Arbib and Rizzolatti (1997). Fig. 3 (their Fig. 4) presents the Arbib and Rizzolatti framework for analysis of the role of F5 in grasping. This combines mechanisms for (1) grasping a seen object (the right hand path from "view of object" to "grasp of object"); and (2) imitating observed gestures in such a way as to create expectations which not only play a role in "social learning" but also enable the visual feedback loop to eventually serve for (delayed) error correction during, e.g., reaching towards a target (the loop on the left of the figure).
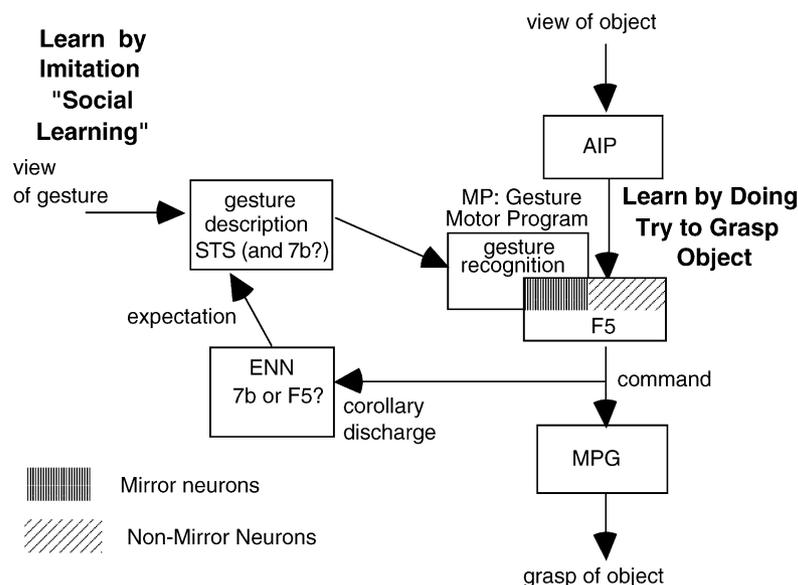

Fig. 3. An integrated conceptual framework for analysis of the role of F5 in grasping. The right hand, vertical, path is the execution system from "view of object" via AIP and F5 to "grasp of object" – it provides mechanisms for grasping a seen object. The loop on the left of the figure provides mechanisms for imitating observed gestures in such a way as to create expectations which enable the visual feedback loop to serve both for "social learning" (i.e., learning an action through imitation of the actions of others) and also for (delayed) error correction during, e.g., reaching towards a target. It combines the observation matching system from "view of gesture" via gesture description (STS) and gesture recognition (mirror neurons in F5 and possibly 7b) to a representation of the "command" for such a gesture, and the expectation system from an F5 command via the expectation neural network ENN to MP, the motor program for generating a given gesture. The latter path may mediate a comparison between "expected gesture" and "observed gesture" in the case of the monkey's self-generated movement (from Arbib & Rizzolatti, 1997; Fig. 4).
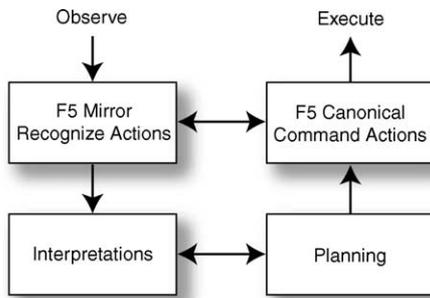
Fig. 4. The perceptuomotor coding for both observation and execution contained in the mirror system for manual actions in the monkey is linked to "conceptual systems" for interpretation and planning of such actions. The interpretation and planning systems themselves do not have the mirror property save through their linkage to the actual mirror system.

The Expectation Neural Network (ENN) is the "Direct Model" of Command → Response. When the animal gives a command (in the sense of brain regions issuing the neural signals that command a movement), ENN generates the expected neural code for the visual signal generated by the resulting gesture. We explicitly label the input to ENN, a copy of the motor command, as a corollary discharge. By contrast, the Motor Program MP provides an "Inverse Model" of Command → Response, by going from a desired response to a command, which can generate it. (This formulation was inspired in part by the work of Jordan & Rumelhart, 1992.) Arbib and Rizzolatti (1997) speculated that the inverse model which converts the view of a gesture to a corresponding command would most likely be located along the path leading from STS to F5 (possibly via 7b). The reciprocal path from F5 to superior temporal sulcus would provide the direct model, ENN. It is equally probable, however, that both ENN and MP are located in F5 and the interplay between stages occurs entirely within F5. If the latter interpretation is accepted, the role of STS areas would be that of giving a highly elaborated dynamic description of gestures – with the observation/execution system entirely located in the frontal lobe.

The integrated model of Fig. 3 thus relates the "grasp an object" system to the "view a gesture" system. The expectation network is driven by F5 irrespective of whether the motor command is "object-driven" (via AIP) or "gesture-driven". It thus creates expectations both for what a hand movement will look like when "object-driven" (an instrumental action directed towards a goal) or "gesture-driven" (a "social action" aimed at making a self-generated movement approximate an observed movement). The right hand path of Fig. 3 exemplifies "learning by doing" – the ILGM model described briefly below shows how this could be achieved by a process of reinforcement learning, in which the success/failure of the grasp acts as positive/negative reinforcement. The left hand path of Fig. 3 exemplifies another mode of learning (the two may be sequential or contemporary) which creates expectations about gestures as well as exemplifying "social learning" based on imitation of gestures made by others. We understand

that the former is present in monkeys, chimps, and humans, whereas the latter is fully developed only in humans.

Although MNS was constructed as a model of the development of mirror neurons in the monkey, we believe that it serves equally well as a model of the development of mirror neurons in the human infant. In any case, the model makes the crucial assumption that the grasps, which the mirror system comes to recognize are already in the (monkey or human) infant's repertoire. But this raises the question of how grasps entered the repertoire. To simplify somewhat, the answer has two parts:

(i) Children explore their environment and as their initially inept arm and hand movements successfully contact objects, they learn to reliably reproduce the successful grasps, with the repertoire being tuned through further experience.
(ii) With more or less help from caregivers, infants come to recognize certain novel actions in terms of similarities with and differences from movements already in their repertoires, and on this basis learn to produce some version of these novel actions for themselves.

Our Infant Learning to Grasp Model (ILGM; Oztop, Bradley, & Arbib, 2004) strongly supports the hypothesis of part (i) that grasps are acquired through experience as the infant learns how to conform the biomechanics of its hand to the shapes of the objects it encounters. It uses as reinforcement a signal of the stability of the grasp when the infant's hand contacts an object and attempts to enclose it, initially by means of the grasp reflex. Over time, the grasp reflex is suppressed and grasps are instead extracted from the repertoire built up by reinforcement learning. Further modeling is required to fully address the issues of imitation raised by in Fig. 3. The idea is that if MNS were augmented to have a population of mirror neurons which could acquire codes for observed actions not yet in the repertoire of self-actions, then the mirror neurons would provide training for the canonical neurons, reversing the information flow seen in the MNS model.

What is the adaptive value of mirror neuron activity to the monkey itself? Most writers have noted the adaptive advantage that such a system could have for social interaction, allowing one monkey to "understand" the actions of another, and thus position itself to compete or cooperate with the other monkey more effectively. However, monkey neurophysiology to date only shows that a macaque (the data are silent about other creatures) can "recognize" certain manual and oro-facial actions made by others in the very special sense that the neural pattern elicited in the F5 mirror neurons[1] by observing those actions is similar to that generated when he performs a similar action himself.

---

[1] And recall that the F5 mirror neurons are part of a larger mirror system that includes (at least) parts of the superior temporal gyrus and area PF of the parietal lobe.

However, this form of "recognition" is more akin to classification by a computer program than to the rich subjective dimensions that often accompany a human's recognition of an act or situation. We are *not* denying that the monkey's recognition of action may be quite rich (though we argue that human language and other cognitive abilities make human awareness very different from the monkey's; Arbib, 2001). We *do* deny that the mere activity of F5 mirror neurons alone suffices to provide such richness, or to constitute "understanding" the action. Consider a pattern recognition device that can be trained to classify pixel patterns from its camera into those, which resemble a line drawing of a circle and those, which do not (with the degree of resemblance cut off at some arbitrary threshold). It does not *understand* circles. However, to the extent that this recognition could be linked to circuitry for drawing a circle, or for forming associations like "the outline of the sun" or "an orthogonal cut through a cone" as yielding an appropriate stimulus, to that extent can one say that the *system* of which the pattern recognizer is part does exhibit some modicum of understanding. Understanding is thus not a binary concept but rather a matter of degree; some things may be encoded appropriately yet not understood at all, others may be understood in great richness because their neural encoding is linked to many other behaviors and perceptions.

3. The Mirror System Hypothesis for language mechanisms

We have seen that premotor area F5 in macaque contains a *mirror system for grasping* which employs a common neural code for *executed* and *observed* manual actions. The homologous region of the human brain is Brodmann's area 44, which is part of Broca's area, traditionally thought of as a speech area but which has been shown by brain imaging studies to be active when humans both execute and observe grasps (Arbib & Bota, in press). These findings are the basis for one account of how the human brain changed from, but built upon, that of ancestral primates to make humans "language-ready":

The Mirror System Hypothesis. (Arbib & Rizzolatti, 1997; Rizzolatti & Arbib, 1997): The parity requirement for language in humans – that what counts for the speaker must count approximately the same for the hearer[2] – is met because Broca's area evolved atop the mirror system for grasping with its capacity to generate and recognize a set of actions.

The key point in elaborating the Hypothesis is that changes in the primate brain might have adapted the use of the hands to support pantomime (intended communication) as well as praxis, and then further evolutionary changes linking protosign and protospeech would yield a brain that could support language. On this view, the "openness" or "generativity"

which some see as the hallmark of language (i.e., its openness to new constructions, as distinct from having a fixed repertoire like that of monkey vocalizations) is present in manual behavior, which can thus supply part of the evolutionary substrate for its appearance in language.[3] Arbib (2002, in press-a) provides more details, including the logical point that a brain that can support language need not be one that evolved for the purpose of supporting language, any more than our brains evolved under the pressure to ensure success at Web surfing. Specifically, the first hominids to have language-ready brains may have had a limited protosign and protospeech, but no full language in the sense of a symbol system equipped with a rich syntax that supports a compositional semantics.

Fig. 5 suggests a framework, which relates the perception and production of language to perception and action more generally. We distinguish Cognitive Form, which concerns the recognition and representation of objects and actions and their relationships, from Phonological Form, which may relate to signed language as much as to spoken language. We see Cognitive Form as present in monkeys as well as (in more complex form) in humans; while, Phonological Form is present only in humans.

For *Production*, the notion is that at any time we have much that we could possibly talk about which might be represented as cognitive structures (Cognitive Form; schema assemblages) from which some aspects are selected for possible expression. Further selection and transformation yields semantic structures (hierarchical constituents expressing objects, actions and relationships), which constitute a Semantic Form enriched by linkage to schemas for perceiving and acting upon the world. Finally, the ideas in the Semantic Form must be expressed in words whose markings and ordering reflect the relationships within Semantic Form. These words must be conveyed as "phonological" structures – with phonological form embracing a wide range of ordered expressive gestures, which may include speech, sign, and oro-facial expressions (and even writing and typing).

For *Perception*, the received sentence must be interpreted semantically with the result updating the "hearer's" cognitive structures. For example, perception of a visual scene may reveal "Who is doing what and to whom/which" as part of a non-linguistic *action-object frame* in cognitive form. By contrast, the *verb-argument structure* is an overt linguistic rep-

---

[2] To include sign language as well as spoken language, "speaker" and "hearer" may actually be using hand and face gestures rather than vocal gestures for communication.

[3] Kohler et al. (2002) found that 15% of mirror neurons in the hand area of F5 can respond to the distinctive sound of an action (breaking peanuts, ripping paper, etc.) as well as viewing the action. Ferrari, Gallese, Rizzolatti, and Fogassi (2003) show that the oro-facial area of F5 (adjacent to the hand area) contains a small number of neurons tuned to communicative gestures (lip-smacking, etc.) but the observation and execution functions of these neurons are not strictly congruent – most of the neurons are active for execution of ingestive actions, e.g., one "observed" lip protrusion but "executed" syringe sucking. Fogassi and Ferrari (in press) and Arbib (in press-b) discuss these data in terms of the relative weight to be given to hand movements and oro-facial movements in the evolution of human brain mechanisms supporting language.
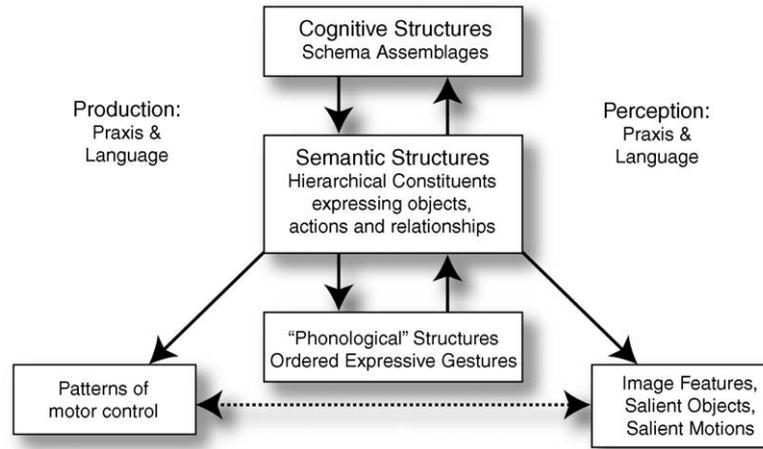
Fig. 5. A view of language which places production and perception of language within a broader framework of action and perception considered more generically. Language production and perception are viewed as the linkage of Cognitive Form (CF), Semantic Form (SF), and Phonological Form (PF), where the "phonology" may involve vocal or manual and gestures, or just one of these, with or without the accompaniment of facial gestures.

resentation in semantic form – in modern human languages, generally the action is named by a verb and the objects are named by nouns or noun phrases.

A *production grammar* for a language is then a specific mechanism (whether explicit or implicit) for converting verb-argument structures into strings of words (and hierarchical compounds of verb-argument structures into complex sentences) and vice versa for a *perception grammar.*

In summary, we argue that the monkey's F5 premotor area for the control of manual and oro-facial actions – and a fortiori the human homologues that include Broca's area and serves both praxic and language-related movements of face, hands and vocal articulators – are situated within the neural context that links the execution and observation of an action to the creature's planning of its own actions and interpretations of the actions of others. These linkages extract more or less coherent patterns from the creature's experience of the effects of its own actions as well as the consequences of actions by others (Fig. 4). Similarly, execution and observation of a communicative action must be linked to the creature's planning and interpretations of communication with others in relation to the ongoing behaviors, which provide the significance of the communicative gestures involved. As such we propose that while STS and other temporal regions may recognize an agent, they must be linked to mirror systems in F5 to bind the agent to the observed action. In terms of the Mirror System Hypothesis, this may ground the linkage of the Agent in the semantic representation of a sentence. Note, however, that language also supports alternative syntactic structures for a given semantic structure, as in the choice in English of active or passive voice, making the Agent subject or object of the sentence.

Rizzolatti and Arbib (1998) view the activity of canonical F5 neurons as part of the code for the cognitive form

$Grasp_A$(Object)

where $Grasp_A$ is a specific motor command directed toward an object with constrained physical properties. Note that neural activity for instrumental action must include many sublinguistic parameters to do with the specification of reach and grasp movements. We say F5 activity is *part* of the code because its encoding of the action $Grasp_A$(–) must be linked to activity elsewhere in the brain to bind the specific raisin to the role of Object. Moreover, neural activity must include many parameters to do with the specification of reach and grasp movements. The full neural representation of the Cognitive Form $Grasp_A$(Object) requires not only the regions AIP and F5 canonical shown in Fig. 1 but also inferotemporal cortex (IT), which holds the identity of the object. Other parts of the brain (e.g., pre-SMA [F6; see Fig. 1] and basal ganglia) then determine whether and when that command will be executed.

Similarly, Rizzolatti and Arbib (1998) view the firing of mirror F5 neurons as part of the code for the cognitive form $Grasp_A$(Agent, Object)

where $Grasp_A$ denotes the specific kind of grasp applied to the Object by the Agent. Again, this is an "action description". If attention is focused on the agent's hand, then the appropriate case structure would be $Grasp_A$(Hand, Object) as a special case of $Grasp_A$(Instrument, Object). Thus, the same act can be perceived in different ways: "Who" grasps versus "With what" the grasp is made. It is worth noting that the monkey's mirror neurons do not fire when the monkey initially observed the experimenter grasping a raisin with pliers rather than his hand but did come to fire after repeated observation. We thus see the ability to learn new constraints on a "slot" – in this case the observed generalization of the Instrument role from hands alone to include pliers. The full neural representation of the Cognitive Form $Grasp_A$(Agent, Object) requires not only the regions
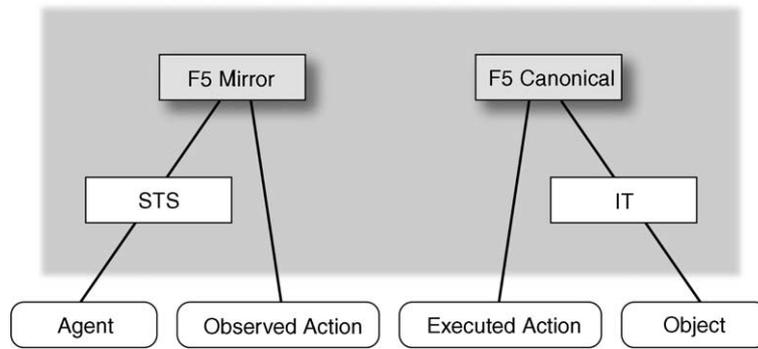
Fig. 6. A schematic of the canonical and mirror neurons stressing that executing an action requires that executing an action requires linking that action to the goal object (as recognized, e.g., by IT), while recognizing an action executed by another requires the further linkage of the action to the agent (who might be recognized by STS).

AIP, STS, 7a, 7b and F5 mirror shown in the MNS diagram (Fig. 1) and inferotemporal cortex to hold the identity of the object (as seen for canonical neurons) but also regions of, for example, the superior temporal sulcus not included in MNS which hold the identity of the agent. Fig. 6 summarizes the structures, which support $Grasp_A$(Object) and $Grasp_A$(Agent, Object) for the macaque brain - but we reiterate that there are no "Linguistic Forms" in the monkey's brain.

Fig. 7 (from Arbib & Bota, 2003) extends the canonical grasping system (FARS model of Fig. 1) and the mirror system for grasping (MNS model of Fig. 2) to give a sketch of how (according to one version of the Mirror System Hy-

pothesis) the language system evolved "atop" this. Note that this simple figure neither asserts nor denies that the extended mirror system for grasping and the language-supporting system are anatomically separable, nor does it address issues of lateralization.

4. Self-monitoring versus attribution of agency

Daprati et al. (1997) had 60 subjects (30 normal control subjects and 30 schizophrenic patients) perform a requested movement with the right hand, and monitor its execution by looking at an image of a hand movement – either a display of the subject's own movement, or a movement starting at
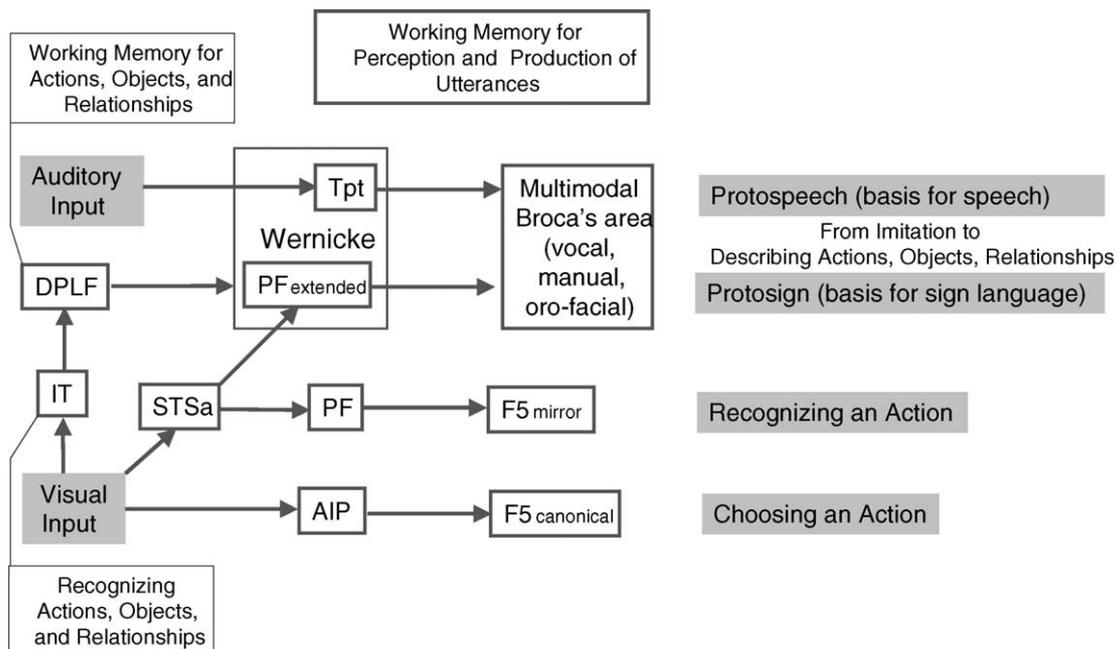
Fig. 7. Extending the FARS model (bottom row) to include the mirror system for grasping (next row) and the grounding of language by integrating Wernicke's area and Broca's area "atop" this (as homologues of PF and related parietal regions, and F5, respectively) to handle protosign and protospeech. Note that this simple figure neither asserts nor denies that the extended mirror system for grasping and the language-supporting system are anatomically separable, nor does it address issues of lateralization (from Arbib & Bota, in press).

the same time with the experimenter moving his/her gloved hand starting from the identical initial position. Once the movement was performed and the screen had blanked out, the subject was asked:

You have just seen the image of a moving hand. Was it your own hand? Answer YES if you saw your own hand performing the movement you have been executing. Answer NO in any other case, that is if you doubt that it was your own hand or your own movement.

One of four possible movements of the fingers was required in each trial:

1. extend thumb,
2. extend index,
3. extend index and middle finger,
4. open hand wide.

One of three possible images of the hand could be presented to the subjects in each trial:

- their own hand (Condition: *Subject*);
- the experimenter's hand performing the same type of movement (Condition: *Experimenter Same*); or
- the experimenter's hand performing a different movement (Condition: *Experimenter Different*).

Both normals and schizophrenics made virtually no errors when subjects saw their own hand, or a hand performing a different movement. The only errors occurred in the *Experimenter Same* condition, where the median error rate was 5% in the control group, 17% in the non-delusional group and 23% in the delusional group.

Note that these results show that *the experiment has little to do with attribution of agency*. In each case, it seems that the subject knows that he has made a movement and which type of movement it is – it is just a case of monitoring that movement accurately enough to tell whether a slight variant is indeed different.

To allow precise control of parameters which can affect judgment in the *Experimenter same* condition, but at the cost of increased "unreality" of the task, Franck et al. (2001) conducted similar experiments using a realistic virtual hand whose movement could reproduce that of the subject's hand but with systematic distortions. This time, subjects had to answer the question:

Did the movement you saw on the screen exactly correspond to that you have made with your hand?

They had to answer YES or NO. Note that this is no longer an agency issue – the hand is a computer hand and the subject has to judge whether the hand "corresponds to mine", not whether the hand is a video of the subject's own hand. Clearly, the difference will be easier to recognize if the two performances involve greater differences in timing or extent.

All subjects showed errors when shown a distorted view that varied little from their own hand movement. Non-delusional patients showed a sharp decrease in erroneous YES responses (down to 50% of maximum number of errors)

once the bias in displayed angles reaches between 15° and 20°, a value not very different from that of controls, whereas influenced patients did not reach the same score until the bias increased to 30–40°. As for the temporal bias, whereas control subjects show a clear decrease in YES responses for a relatively small bias (100–150 ms), both influenced and non-influenced patients do not show a decrease in the rate of YES responses until the bias reaches 300 ms. In discussing this result, Franck et al. (2001) assert that:

A decrease in sensitivity of this mechanism would explain the greater difficulties met by the schizophrenic patients. Even normal subjects misjudged the ownership of the experimenter's hand in about 30% of trials. This finding suggests *that the mechanism for recognizing actions and attributing them to their true origin* [our italics] operates with a relatively narrow safety margin: In conditions where the visual cues are degraded or ambiguous, it is barely sufficient for making correct judgments about the origin of action, although it remains compatible with correct determination of agency in everyday life.

However, the assertion "that the mechanism for recognizing actions and attributing them to their true origin" seems to us to be a confusion between the delusional situation and situation here where the patient knows that it is he (or she, as the case may be) who acted. In this paradigm, the patient *knows* he acted. It is thus mistaken to lump all schizophrenic symptoms under the rubric of deficits in attribution of agency. To clarify this, we analyze what we consider to be two *different* factors that may affect the symptoms of schizophrenia: self-monitoring, and attribution of agency.[4]

Self-monitoring involves maintaining a working memory of one's recent actions as a basis for evaluating their consequences (more on this below).

Agency attribution is different from, e.g., *face recognition*, yet in general must rest on the success of such processes. We assume that a schizophrenic can correctly identify the picture of a face (no prosopagnosia). But the issue is how to link that recognition to an ongoing action. We do not, generally, attribute agency to movements of a disembodied hand. Rather, we seek to link the hand to an arm and thus to a body and head whose face we can recognize. Of course, we can also recognize a person by their voice or by a characteristic walk, for example. But the point is that the recognition of a person must be linked to the local analysis of the action for the action to be attributed to that person as agent. This needs some Working Memory (WM) for parts of a scene to compute their relationship. Presumably, again, a schizophrenic can recognize a moderately complex scene given enough time to scan the contents (but if this process is too slow, one representation may decay while another forms) – but the catch with

---

[4] Stephens and Graham (1994) make a different distinction – that between sense of agency and sense of subjectivity (≈what Gallagher, 2000 calls sense of ownership) – distinguishing the possibility of self-ascription of a psychological state from the subjective experience of the state.

action is that it may change too quickly to be 'recorded" if the "writing" to WM is impaired.

In the simplest case, we see an action and we see a result, and we perceive both who was the agent of the action and whether the result was indeed the consequence of that action. For the moment, inspired by the study of the mirror system for grasping, let us consider a hand holding a knife, which pierces someone. Normally (i.e., with highest probability), we would attribute agency to the person of whose body the hand is part. But how do we know who that person is? We might turn in horror from the hand holding the knife to the face of its owner and recognize the face as one we know, or remember a new face for possible later identification in a line-up. Or it might be the voice of the person that serves to identify them if the face is obscured. How would we deny agency in this case? (Agency is a separate issue from guilt – we are not asking whether the knifing was a "free decision", or in self-defense, or a response to coercion.) We might, for example, deny agency to the knife holder if we recognized (and this recognition might well rest on our working memory of the last few seconds) that the arm of the knife-holder had been jolted forward by someone else, or that the knife-holder had not moved but instead the victim had suddenly lunged toward him. All this is simply to make the point that the attribution of agency is in general a complex process integrating multiple neural systems for perception, working memory and inference – of which the role of the STS in face recognition (cf. Fig. 6) is one (sometimes relevant) part.

Extending the view of the role of the mirror system in understanding (in the section "The Mirror System for Grasping", and recalling footnote 4) from manual, vocal and oro-facial actions, a number of authors have suggested that a common set of premotor centers is involved in both action imagination and action creation (Frith, 2002; Gallese & Goldman, 1998; Jeannerod, 1994, in press), but less attention has been given to the mechanisms whereby the brain can distinguish the "simulation" involved in recognizing the action of another from the actual creation of an action by the self. This suggests that this premotor activity must be compared against the working memories that track the intention (conscious and unconscious) for the actions. We suggest that when one has a thought about saying something or commit to a movement, one stores the intention of the action in working memory, and this includes knowledge of the action and the relating agent, thus grounding expectations for the behavior of oneself and others. The binding of agent to action stored in working memory plays a crucial role in our behavior and our understanding of behavior.

When I execute an action I either have the intention to commit this action or recognize the action as part of a larger intention when I receive feedback on that action. For example, if I suddenly swerve while driving, I will not have intended that swerve in advance but will recognize that it was an appropriate (but not premeditated) response to, say, an unexpected obstacle and that it fits within my overall intention. It is relevant here that my brain can take account of feedback both at and below the conscious level of my intentions, refining what Arbib (1981) has called the coordinated control program that coordinates the perceptual and motor schemas involved in the action (a conceptual framework inspired by the empirical data of Jeannerod & Biguer, 1982). As an example of unconscious feedback, a microslip of an object I am holding causes me to tighten my grip and this in turn leads me to provide a tighter initial grip on such objects in future (Johansson & Westling, 1984). Similarly, when I speak I may be most conscious of feedback on the effect of what I say on the person I am speaking to – i.e., in relation to my communicative intention – yet am constantly making adjustments at many levels down to the detailed effects of articulation. In general, the issuing of any command for action within the brain is accompanied by an expectation of the outcome of that action, and current actions generally unfold within the context of recent actions and ongoing plans which situate potential future actions with respect to current goals. (Recall the discussion of the diverse forms of working memory we may associate with the elaboration of the FARS model of Fig. 1 and the discussion of corollary discharge in Fig. 3.) All of these – goals, plans, intentions, actions, and expectations – require "working memories", whether the data they contain are accessible to conscious introspection or not. These contribute on-line during the creation of action (Fletcher, Shallice, & Dolan 2000) as well as in reconciling expectations and outcomes and learning (whether consciously or unconsciously) from the result. My conscious intention, i.e., the overall idea of what I would like to do, requires elaboration of all the subconscious details to assemble a coherent action. This assembly involves many brain areas with especial importance resting on the interaction of parietal and frontal areas.

How does this tie back, then, to the work of Daprati et al. (1997) and Franck et al. (2001)? We would claim that they do not tap the problem of agency, but rather tap the function of self-monitoring. We see this function as separate from agency as we first explain, but nonetheless relevant to agency, as we shall explain thereafter.

4.1. Self-monitoring as distinct from agency

We have seen that, in implementing the MNS model, Oztop and Arbib (2002) train the system to recognize what "hand-state trajectories" – multi-dimensional trajectories relating the motion and opposition axes of the hand to the location and affordances of the object – correspond to a given grasp type (as coded by the firing of canonical F5 neurons). The aim is to recognize the grasp type earlier rather than later in the trajectory, to the extent this is possible, but in any case the role of a working memory linking movement of an agent or instrument to objects in the world is a crucial ingredient of the learning process. The same is true of motor skill learning more generally.[5]

---

[5] In experiments in which perturbations are systematically applied across multiple trials (as distinct from an isolated perturbation whose cause I may

4.2. Self-monitoring in relation to agency

Suppose I sweep my hand backward in an expressive gesture and a moment later hear a glass drop to the floor behind me. This near coincidence has me worrying that I might have caused someone to drop the glass. Turning round and seeing where the glass hit the floor and the position of the person who dropped it, I extract from working memory an estimate of how far my hand swept backwards to infer whether it was likely that I caused the accident. Such a scenario is indicative of the way in which working memory may be crucial to my understanding of my agency with respect to certain observed consequences. Similarly, in the case of the swerving car, I may compare a trajectory with an expected trajectory to decide (consciously or unconsciously) whether the departure was such that I should posit an external cause. But in either case, I know that I am the agent of my primary action, even if it departs from my expectations whether in terms of an unexpected consequence or an unexpected perturbation whose cause I may (or may not) then seek.

In this perspective, Daprati et al. (1997) and Franck et al. (2001) demonstrate that their patients have an impaired working memory for self-monitoring (but a memory against which comparisons are more tolerant of errors; not a working memory that is lost) but do not address the issue of attribution of agency.

5. Thoughts on agency

In cases of hemineglect following a parietal lesion, a patient will deny that the arm on his neglected side is his own. Indeed, this denial is so strong that it confounds logic. Susan Bookheimer (personal communication, 2002) reports working with a patient who was asked to follow the arm up from hand to shoulder. The patient conceded "Well, it attaches to me" yet still asserted "But I don't think it's mine!" and suggested that a doctor had brought in the arm and attached it to play a trick on him. What can separate the observation of attachment from the sense of ownership? We will not answer the question but instead take the point that certain elements of personal experience can be so strong that we may refuse to reject this "testimony of the senses" and instead confabulate, seeking a pattern of explanation – no matter how implausible – that can bring the various observations into seeming coherence. It is this general observation that we bring back to this essay on schizophrenia.

We ask how we are to view attribution of self-agency and the attribution of self-induced movements. In this essay, we simply elaborate on the diagram of Fig. 1.[6] In the FARS Model, visual input from an object is processed to extract affordances (AIP) and object identity (IT). In "feedforward mode", the AIP activity leads to the activation of a motor schema in F5, which in turn causes the execution of the given action. More generally, the input may afford a variety of actions, and then the prefrontal cortex (PFC) will invoke various constraints (including information from IT) to bias the choice of affordance (acting on F5 in the original FARS model, but via AIP in the "FARS Modificato" of Fig. 1). Going beyond Fig. 1, the implemented FARS model also supports the execution of sequences. To simplify somewhat, the idea is that the sequence is represented in pre-SMA, and that a sequence A–B–A–C of motor schemas would be controlled by a sequence x_1–x_2–x_3–x_4 of states of pre-SMA, with trigger conditions y_1, y_2, y_3 (whether external events or internal states) for the transitions. Each x_i of SMA activates its own motor schema, call it $m(x_i)$, but also primes the next x_{i+1}, and thus $m(x_{i+1})$. One of the jobs of the basal ganglia (BG) is to inhibit x_{i+1} from going above threshold for executing $m(x_{i+1})$ until y_i has occurred. Once it occurs, then x_{i+1} and thus $m(x_{i+1})$ must be disinhibited, x_i and thus $m(x_i)$ must be inhibited, and x_{i+2} and thus $m(x_{i+2})$ must be primed but held below the execution threshold until the trigger y_{i+1} is encountered. Abstracting from this, we may say that an action m is intended only if there is explicit prefrontal activity x to prime it, and other prefrontal activity y to release the inhibition that holds its premotor (and thus motor) activity below the threshold for execution. In the simplest case (the canonical FARS case, one might say), the mere perception of an affordance for grasping an object will not be enough to yield the execution of that grasp as an intentional act. This requires priming of this act among other acts, and the disinhibition that effectively acts as a go signal.

We would argue that in general the brain simply processes self-actions without need to attribute agency to the action – the brain simply does its job of integrating perception, plan and action. Putting it another way, I usually brush my teeth without in addition reflecting it is *I* who is doing the brushing.[7] However, we hypothesize that each action is accompanied by a more or less accurate motor working memory of the trajectory of the action. Thus, if the need arises to question the agency of the action, the brain may consult its working memories (the plural is significant) to determine whether there was the (x, y) of priming and disinhibition prior to the action and if so whether the working memory of expected outcome of the action sufficiently matches the observed trajectory of the outcome. On this basis, the normal brain can decide "I am the agent", "I was the agent but for some reason the action did not come out as intended", or "I am not the agent".

---

try to understand), I may learn to adjust forces along an entire trajectory (this adjustment probably rests more on the cerebellum than on the regions of cerebral cortex that occupy us here).

[6] Modeling now underway is not only filling in the details in terms of analyzing realistic excitatory and inhibitory interactions of the neurons in the areas shown in Figs. 1 and 2, but is further playing specific attention to the roles of dopamine and serotonin in working memory in prefrontal cortex, since changes in these neuromodulators have been implicated in various disorders of working memory as well as in schizophrenia itself.

[7] Note the nice grammatical point that it is "is" rather than "am" that is the correct verb here – the "I" is viewed as a third person when I reflect on my own agency.

We close, then, with a few remarks linking all this back to schizophrenia. Schizophrenic patients hallucinate voices that they attribute to external agents; they also have delusions that other people are creating their actions for them and they also have delusions of influencing others to act (DSM-IV, 1994). In addition, patients with schizophrenia have difficulty determining whether they spoke or thought an utterance (Brébion, Gorman, Amador, Malaspina, & Sharif, 2002; Franck et al., 2000). It has also been observed that schizophrenic patients can project their intentions onto other agents (Sarfati, Hardy-Bayle, Besche, & Widlocher, 1997). For instance, a schizophrenic patient has the experience of controlling another agent as they observe that other agent's actions. From this, it seems that schizophrenic patients have difficulty not only in understanding the nature of their actions, but also in terms of knowing who is controlling an action.

We admit that our above account appears quite complex at first, but our experience with the FARS and MNS models and with modeling the basal ganglia suggests that it a useful simplification rather than an unprincipled complexification. Given this complexity, there may be many places at which the attribution of agency breaks down. In our forthcoming work, we will first explore the notion that that the primary deficit is in the lack of adequate control of disinhibition. Thus an action may be committed without need for a disinhibitory signal x that represents the decision to execute the action. Lacking any memory of having intended the action, the patient concludes "I am not the agent" and then proceeds to confabulate, to provide an account for the agency of the observed action. Our account is consistent with – but offers a fresh perspective on – the hypothesis that the problems of schizophrenia involve working memory (Perlstein, Dixit, Carter, Noll, & Cohen, 2003; Posada, Franck, Georgieff, & Jeannerod, 2001). Frith, Blakemore, and Wolpert (2000) (Frith, 2002) argue that when a normal person executes an action, the observation of the action is fed back and compared to the action representation held in working memory, whereas in a schizophrenic patient the record of the action is lost. However, the previous section looked at cases where the patient is well aware of having himself made a movement – the issue is whether he judges a similar movement on a TV screen as (based on) his own or that of the experimenter, and the data suggest not so much that the record of the action is lost as that is "blurred". In short, the knowledge or memory that one has committed an action seems distinguishable from the accuracy with which that action is recalled – *we must distinguish attribution of agency from self-monitoring.* The Franck et al. data suggest that the schizophrenic certainly knows that many of his actions are self-initiated and then remembers his action but with "permissive parameters". In other words, the "memory of intending" is perhaps more important than "the details of what was intended."

Schizophrenic misattributions of agency are commonly linked to hand movements and language. While delusions of influence are not as common as auditory verbal hallucinations, in most cases they take the form that the schizophrenic hallucinates that another agent is causing his hand to move. This leads us to stress the relevance of the Mirror System Hypothesis for the study of schizophrenia. Extending the Hypothesis, we suggest that the working memories for language production are evolved from, yet still closely related to, those for hand movements. This would explain why the disease does not strike all working memories and all "releasers of intention" equally but most affects those for hand movements and language. We suggest that schizophrenia is a disorder of the combined system, but also stress that the disorder leads to an impairment of this working memory system that is statistical in effect, rather than simply excising the whole system. Thus, depending on "where the dice fall", the patient's misattribution of agency may be related more to hands or voices, or may affect both in large part. We thus suggest that auditory verbal hallucinations are accounted for by the observation that auditory pathways are active during hallucination (Stephane et al., 2001) and produce a verbal process of some internal voice, but since no record is kept of the voice being created, it is treated as external. That is, an utterance is created and progresses through verbal creation pathways, and returns as a vocalization *observed*, only to be dismissed as external since no record of it being created has been kept. Schizophrenic patients by this theory then confabulate the agent. The confabulated agent then takes on a strong identity persisting across hallucinatory episodes, even if the fictitious agent is nowhere to be found, or does not even exist.

References

Arbib, M.A. (1981). Perceptual structures and distributed motor control. In V.B. Brooks (Ed.), *Handbook of physiology, Section 2: The nervous system, Vol. II, Motor control, Part 1.* (pp. 1449–1480). American Physiological Society.

Arbib, M. A. (2001). Co-evolution of human consciousness and language. In P. C. Marijuan (Ed.), *Cajal and consciousness: Scientific approaches to consciousness on the centennial of Ramón y Cajal's Textura, Vol. 929* (pp. 195–220). Annals of the New York Academy of Sciences.

Arbib, M. A. (2002). The mirror system, imitation and the evolution of language. In C. Nehaniv & K. Dautenhahn (Eds.), *Imitation in animals and artifacts* (pp. 229–280). The MIT Press.

Arbib, M. A. (in press-a). From monkey-like action recognition to human language: An evolutionary framework for neurolinguistics. *Behavioral and Brain Sciences*.

Arbib, M. A. (in press-b). Interweaving protosign and protospeech: Further developments beyond the mirror. In *Interaction studies: social behaviour and communication in biological and artificial systems*.

Arbib, M. A., Billard, A., Iacoboni, M., & Oztop, E. (2000). Synthetic brain imaging: Grasping, mirror neurons and imitation. *Neural Networks*, *13*, 975–997.

Arbib, M. A., & Bota, M. (2003). Language evolution: Neural homologies and neuroinformatics. *Neural Networks*, *16*, 1237–1260.

Arbib, M., & Rizzolatti, G. (1997a). Neural expectations: A possible evolutionary path from manual skills to language. *Communication and Cognition*, *29*, 393–424 [reprinted in Ph. Van Loocke (Ed.), *The nature, representation and evolution of concepts*. London/New York: Routledge].

Arbib, M., & Rizzolatti, G. (1997b). Neural expectations: A possible evolutionary path from manual skills to language. *Communication and Cognition*, *29*, 393–424.

Brébion, G., Gorman, J. M., Amador, X., Malaspina, D., & Sharif, Z. (2002). Source monitoring in schizophrenia: Characterization and associations with positive and negative symptomatology. *Psychiatry Research*, *112*, 27–39.

Carr, L., Iacoboni, M., Dubeau, M. C., et al. (2003). *Proc Natl Acad Sci USA*, *100*, 5497–5502.

Daprati, E., Franck, N., Georgieff, N., Proust, J., Pacherie, E., Dalery, J., et al. (1997). Looking for the agent: An investigation into consciousness of action and self-consciousness in schizophrenic patients. *Cognition*, *65*, 71–86.

DSM-IV. (1994). *Diagnostic and statistical manual of mental disorders* (4th ed.). Washington, DC: American Psychiatric Association.

Fagg, A. H., & Arbib, M. A. (1998). Modeling parietal-premotor interactions in primate control of grasping. *Neural Networks*, *11*, 1277–1303.

Ferrari, P. F., Gallese, V., Rizzolatti, G., & Fogassi, L. (2003). Mirror neurons responding to the observation of ingestive and communicative mouth actions in the monkey ventral premotor cortex. *European Journal of Neuroscience*, *17*, 1703–1714.

Fletcher, P. C., Shallice, T., & Dolan, R. J. (2000). "Sculpting the Response Space" – an account of left prefrontal activation at encoding. *NeuroImage*, *12*, 404–417.

Fogassi, L., & Ferrari, P. F. (in press). Mirror neurons, gestures and language evolution. In *Interaction studies: Social behaviour and communication in biological and artificial systems*.

Franck, N., Rouby, P., Daprati, E., Dalery, J., Marie-Cardine, M., & Georgieff, N. (2000). Confusion between silent overt reading in schizophrenia. *Schizophrenia Research*, *41*, 357–364.

Franck, N., Farrer, C., Georgieff, N., Marie-Cardine, M., Dalery, J., d'Amato, T., et al. (2001). Defective recognition of one's own actions in patients with schizophrenia. *American Journal of Psychiatry*, *158*, 454–459.

Frith, C. D. (2002). Attention to action and awareness of others minds. *Consciousness and Cognition*, *11*, 481–487.

Frith, C. D., Blakemore, S., & Wolpert, D. M. (2000). Explaining the symptoms of schizophrenia: Abnormalities in the awareness of action. *Brain Research Reviews*, *31*, 357–363.

Gallagher, S. (2000). Self-reference and schizophrenia: A cognitive model of immunity to error through misidentification. In D. Zahavi (Ed.), *Exploring the self: Philosophical and psychopathological perspectives on self-experience* (pp. 203–239). Benjamins: Amsterdam.

Gallese, V., & Goldman, A. (1998). Mirror neurons and the simulation theory of mind-reading. *Trends in Cognitive Sciences*, *2*, 493–501.

Iacoboni M. (in press). Understanding others: Imitation, language, empathy. In S. Hurley, N. Chater (Eds.), *Perspectives on imitation: From cognitive neuroscience to social science: Vol. 1. Mechanisms of imitation and imitation in animals*. Cambridge, MA: MIT Press.

Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioural Brain Science*, *17*, 187–245.

Jeannerod, M. in press). How do we decipher others' minds? In J.-M. Fellous, M. A. Arbib (Eds.), *Who needs emotions? The brain meets the robot*. New York: Oxford University Press.

Jeannerod, M., & Biguer, B. (1982). Visuomotor mechanisms in reaching within extra-personal space. In D. J. Ingle, R. J. W. Mansfield, & M. A. Goodale (Eds.), *Analysis of visual behavior* (pp. 387–409). The MIT Press.

Jeannerod, M., Arbib, M. A., Rizzolatti, G., & Sakata, H. (1995). Grasping objects: the cortical mechanisms of visuomotor transformation. *Trends in Neurosciences*, *18*, 314–320.

Johansson, R. S., & Westling, G. (1984). Influences of cutaneous sensory input on the motor coordination during precision manipulation. In C. von Euler & O. Franzen (Eds.), *Somatosensory mechanisms* (pp. 249–260). MacMillan.

Jordan, M. I., & Rumelhart, D. E. (1992). Forward models: Supervised learning with a distal teacher. *Cognitive Science*, *16*, 307–354.

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

Miall, R. C. (2003). Connecting mirror neurons and forward models. *NeuroReport*, *14*, 2135–2137.

Oztop, E., & Arbib, M. A. (2002). Schema design and implementation of the grasp-related mirror neuron system. *Biological Cybernetics*, *87*, 116–140.

Oztop, E., Bradley, N. S., & Arbib, M. A. (2004). Infant grasp learning: A computational model. *Experimental Brain Research*, *158*, 480–503.

Perlstein, W. M., Dixit, N. K., Carter, C. S., Noll, D. C., & Cohen, J. D. (2003). Prefrontal cortex dysfunction mediates deficits in working memory and prepotent responding in schizophrenia. *Biological Psychiatry*, *53*, 24–38.

Posada, A., Franck, N., Georgieff, N., & Jeannerod, M. (2001). Anticipating incoming events: an impaired cognitive process in schizophrenia. *Cognition*, *81*, 209–225.

Rizzolatti, G., & Arbib, M. A. (1998). Language within our grasp. *Trends in Neuroscience*, *21*, 188–194.

Rizzolatti, G., & Luppino, G. (2001). The cortical motor system. *Neuron*, *31*, 889–901.

Rizzolatti, G., & Luppino, G. (2003). Grasping movements: Visuomotor transformations. In M. A. Arbib (Ed.), *The handbook of brain theory and neural networks* (2nd ed., pp. 501–504). Cambridge, MA: The MIT Press.

Rizzolatti, G., Fadiga, L., Gallese, V., & Fogassi, L. (1996). Premotor cortex and the recognition of motor actions. *Cognitive Brain Research*, *3*, 131–141.

Rolls, E. T., & Arbib, M. A. (2003). Visual scene perception, neurophysiology principles. In M. A. Arbib (Ed.), *The handbook of brain theory and neural networks* (2nd ed., pp. 1210–1215). Cambridge, MA: A Bradford Book/The MIT Press.

Sarfati, Y., Hardy-Bayle, M. C., Besche, C., & Widlocher, D. (1997). Attribution of intentions to others in people with schizophrenia: A non-verbal exploration with comic strips. *Schizophrenia Research*, *25*, 199–209.

Stephane, M., Barton, S., & Boutros, N. N. (2001). Auditory verbal hallucinations and dysfunctions of the neural substrates of speech. *Schizophr Research*, *50*, 61–78.

Stephens, G. L., & Graham, G. (1994). Self-consciousness, mental agency, and the clinical psychopathology of thought insertion. *Philosophy, Psychiatry, and Psychology*, *1*, 1–12.

Webster, M. J., Bachevalier, J., & Ungerleider, L. G. (1994). Connections of inferior temporal areas TEO and TE with parietal and frontal-cortex in macaque monkeys. *Cerebral Cortex*, *4*, 470–483.