Mental imagery of speech: linking motor and perceptual systems through internal simulation and estimation

Xing Tian* and David Poeppel

Poeppel Lab, Department of Psychology, New York University, New York, NY, USA

The neural basis of mental imagery has been investigated by localizing the underlying neural networks, mostly in motor and perceptual systems, separately. However, how modality-specific representations are top-down induced and how the action and perception systems interact in the context of mental imagery is not well understood. Imagined speech production ("articulation imagery"), which induces the kinesthetic feeling of articulator movement and its auditory consequences, provides a new angle because of the concurrent involvement of motor and perceptual systems. On the basis of previous findings in mental imagery of speech, we argue for the following regarding the induction mechanisms of mental imagery and the interaction between motor and perceptual systems: (1) Two distinct top-down mechanisms, *memory retrieval* and *motor simulation*, exist to induce estimation in perceptual systems. (2) Motor simulation is sufficient to internally induce the representation of perceptual changes that would be caused by actual movement (perceptual associations); however, this simulation process only has modulatory effects on the perception of external stimuli, which critically depends on context and task demands. Considering the proposed simulation-estimation processes as common mechanisms for interaction between motor and perceptual systems, we outline how mental imagery (of speech) relates to perception and production, and how these hypothesized mechanisms might underpin certain neural disorders.

Keywords: internal forward model, efference copy, corollary discharge, sensory-motor integration, mirror neurons, auditory hallucination, stuttering, phantom limb

INTRODUCTION

Mental imagery can be characterized as a quasi-perceptual experience, induced in the absence of external stimulation. Neuroimaging studies have shown that common neural substrates mediate mental imagery and the corresponding perceptual processes, such as in visual (e.g., Kosslyn et al., 1999; O'Craven and Kanwisher, 2000), auditory (e.g., Zatorre et al., 1996; Kraemer et al., 2005), somatosensory (e.g., Yoo et al., 2003; Zhang et al., 2004), and olfactory domains (e.g., Bensafi et al., 2003; Djordjevic et al., 2005). The demonstration of activation in corresponding perceptual regions during mental imagery has provided strong evidence to support the claim that the perceptual experience during mental imagery is mediated by modality-specific neural representations (see the review by Kosslyn et al., 2001). However, the top-down "induction mechanism" for the neural activity mediating mental imagery is not well understood.

We focus here on the role of the motor system in the construction of perceptual experience in mental imagery. We propose a motor-based mechanism that is an alternative (additional) mechanism to Kosslyn's memory-attention-based account (Kosslyn, 1994, 2005; Kosslyn et al., 1994): planned action is simulated in motor systems to internally derive the representation of perceptual changes that would be caused by the actual action (perceptual associations). We suggest that the deployment of these two distinct mechanisms depends on task demands and

contextual influence. Studies of mental imagery of speech are summarized to provide evidence for the proposed account—and for the coexistence of both mechanisms. We discuss the motor-to-sensory integration process and propose some working hypotheses regarding certain neural and neuropsychiatric disorders from the perspective of the proposed internal simulation and estimation mechanisms.

DIFFERENT ROUTES FOR INDUCING MENTAL IMAGES

MENTAL IMAGERY OF PERCEPTION AS MEMORY RETRIEVAL (DIRECT SIMULATION)

Mental imagery has been proposed to be essentially a memory retrieval process. That is, perceptual experience is simulated by reconstructing stored perceptual information in modality-specific cortices (Kosslyn, 1994, 2005; Kosslyn et al., 1994). In particular, the process, guided by attention, retrieves object and spatial properties stored in long-term memory to reactivate the topographically organized sensory cortices that represent the object features. Through top-down (re)construction of the neural representation that is similar to the result of bottom-up perceptual processes, the perceptual experience can be re-elicited without the presence of any physical stimuli during mental imagery. This attention-guided memory retrieval process has been demonstrated, for example, in the visual imagery of faces (Ishai et al., 2002).

Mental imagery is further hypothesized to be a predictive process (for future perceptual states), in which the dynamics of perceptual experience can be retrieved/calculated and reconstructed internally (Moulton and Kosslyn, 2009). That is, given an initial point, the series of future perceptual states can be internally simulated by following the regularity (temporal and causal constraints) stored in declarative memory (general knowledge). The mapping between internal simulation and the perception of external stimulation is thought not to be necessarily isomorphic (Goldman, 1989), as only the essential intermediate states are required to have a one-to-one mapping (Fisher, 2006). Because this proposed simulation process is executed entirely within perceptual domains on the basis of memory retrieval—without any representational transformation between motor and perceptual systems—we refer to this account as *direct simulation*.

MENTAL IMAGERY OF MOTOR ACTION AS ESTIMATION DERIVING FROM SIMULATION

Motor imagery is thought to be the process that internally simulates planned actions, by activating similar neural substrates that mediate motor intention and preparation (Jeannerod, 1995, 2001; Decety, 1996). Numerous studies have demonstrated both frontal and parietal activity during motor imagery (Decety et al., 1994; Lotze et al., 1999; Gerardin et al., 2000; Ehrsson et al., 2003; Hanakawa et al., 2003; Dechent et al., 2004; Meister et al., 2004; Nikulin et al., 2008). However, motor system activation does not necessarily link to the kinesthetic feeling generated during motor imagery. The residual neural activity, resulting from the absence of external somatosensory feedback, is thought to mediate the kinesthetic experience during motor imagery (Jeannerod, 1994, 1995). The implicit assumption of the "residual activity account" is that the internal motor simulation during imagery should be transformed into the same representational format as the one resulting from somatosensory feedback. That is, the somatosensory consequences of motor simulation should be estimated. This is consistent with the view that parietal rather than frontal motor regions mediate motor awareness (Desmurget and Sirigu, 2009). In support of the claim that parietal regions mediate somatosensory estimation, direct current stimulation over parietal cortex induces false belief of movement (Desmurget et al., 2009); parietal lesions also impaired the temporal precision of performing motor imagery tasks (Sirigu et al., 1996). Cumulatively, the results suggest that motor simulation in frontal cortex converges in parietal regions to form a kinesthetic representation.

The internal transformation between motor simulation and somatosensory estimation has been proposed in the context of internal forward models in the motor control literature [see the review by Wolpert and Ghahramani (2000)]. The core presupposition is that the neural system can predict the perceptual consequences by internal simulating a copy of a planned action command (the efference copy). Mental imagery has been linked to the concept of internal forward models by the argument that the subjective feeling in mental imagery is the result of the internal estimation of the perceptual consequences following the internal simulation of an action (Grush, 2004). Consistent with
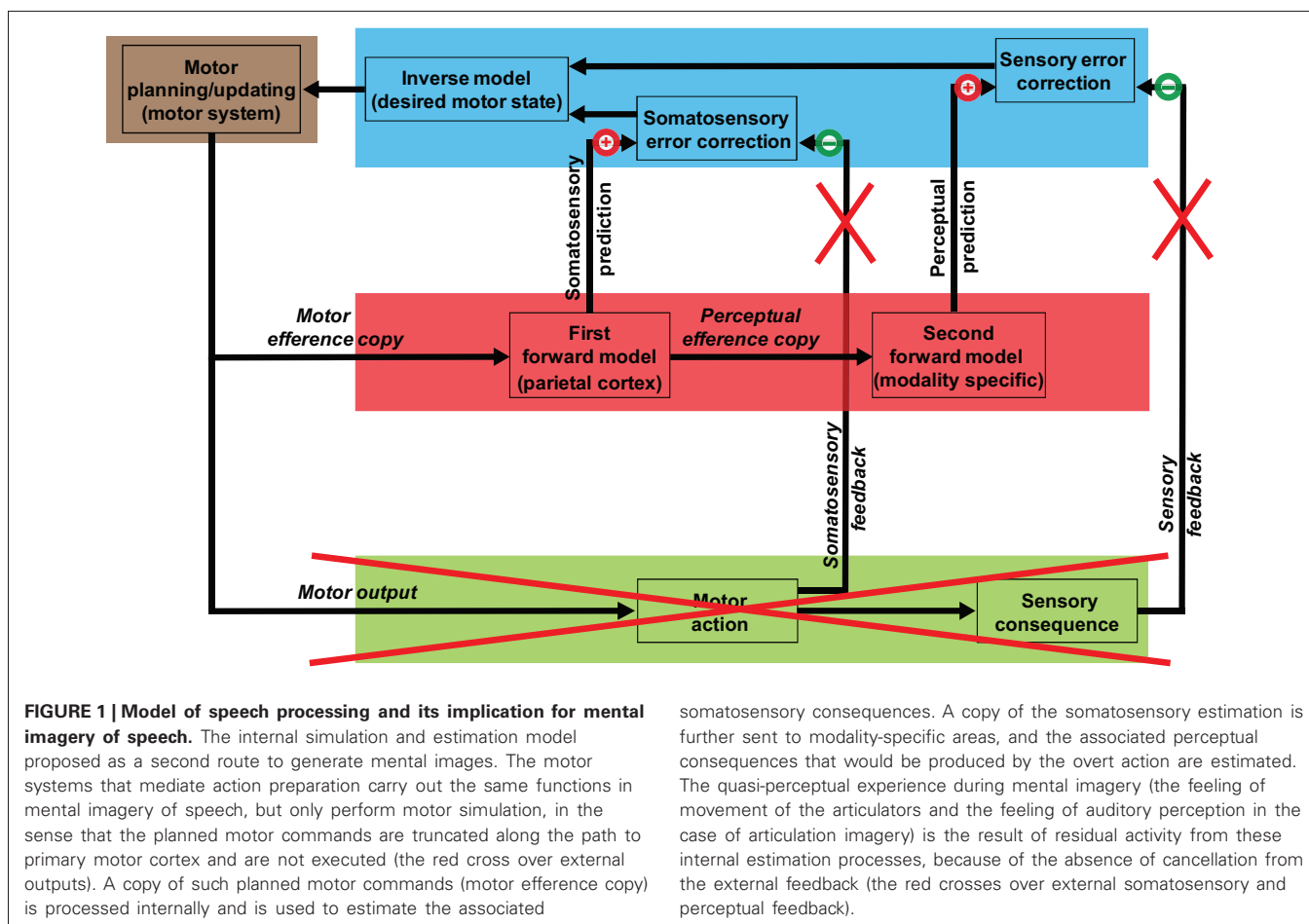
this hypothesis, we propose here that the kinesthetic feeling in motor imagery is the result of somatosensory estimation, derived from internal simulation that closely mimics the dynamics of a motor action. We refer to this account as *motor simulation and estimation*.

The *motor simulation and estimation* account differs from the *direct simulation* (memory retrieval) account in that it requires a transformation between motor and somatosensory systems. Our question here, though, extends beyond this: can a motor simulation deliver perceptual consequences that extend to other sensory domains (such as visual and auditory) as well? If so, internal simulation and estimation processes would serve as an additional path to induce modality-specific neural representations similar to the ones induced on the basis of memory retrieval. In the next section, we discuss this possibility in the framework of internal forward models and propose a *sequential simulation and estimation* account. We will use the interaction of motor, somatosensory, and auditory systems in speech production as an example to illustrate such internal cascaded processes, which can generalize to other sensory domains.

MENTAL IMAGERY OF SPEECH AS SEQUENTIAL ESTIMATION

Perception and production systems are functionally connected: perceptual systems analyze the sensory input generated by self actions; the motor system is also regulated by perceptual feedback to perform updates on actions in the future. For example, when people talk, they move their articulators, feel the movement, and hear the self-produced speech that can be used to detect and correct any pronunciation errors. The temporal sequence of physical articulation, proprioception of the articulators, and auditory perception of one's own vocalization makes it possible—on the basis of co-occurrence and associative learning during development—to create internal connections among the neural processes that mediate motor action, somatosensory feedback, and auditory perception. After establishing the connections, motor commands can cycle internally through somatosensory regions and "reach" auditory regions. That is, the estimation in the somatosensory system can serve as a link between motor and auditory systems. Theoretically, such a cascaded estimation architecture has been hypothesized by Hesslow (2002). Anatomically and functionally, the connections between parietal regions and auditory temporal regions have also been demonstrated (Schroeder et al., 2001; Foxe et al., 2002; Fu et al., 2003).

On the basis of recent neurophysiological (MEG) studies, we proposed that a process of auditory inference after somatosensory estimation occurs during *overt speech processing* [**Figure 1**; adapted from Tian and Poeppel (2010)]. Specifically, the estimation of auditory consequences relies on the somatosensory estimation that derives from the simulation of planned action. That is, the internal *auditory* prediction is the result of a coordinate transformation from the somatosensory to the auditory domain. This sequential estimation mechanism (motor plan → somatosensory estimation → auditory prediction/estimation) can derive detailed auditory predictions that are then compared with auditory feedback for self-monitoring and online control.

FIGURE 1 | Model of speech processing and its implication for mental imagery of speech. The internal simulation and estimation model proposed as a second route to generate mental images. The motor systems that mediate action preparation carry out the same functions in mental imagery of speech, but only perform motor simulation, in the sense that the planned motor commands are truncated along the path to primary motor cortex and are not executed (the red cross over external outputs). A copy of such planned motor commands (motor efference copy) is processed internally and is used to estimate the associated somatosensory consequences. A copy of the somatosensory estimation is further sent to modality-specific areas, and the associated perceptual consequences that would be produced by the overt action are estimated. The quasi-perceptual experience during mental imagery (the feeling of movement of the articulators and the feeling of auditory perception in the case of articulation imagery) is the result of residual activity from these internal estimation processes, because of the absence of cancellation from the external feedback (the red crosses over external somatosensory and perceptual feedback).

In the case of the *mental imagery of speech*, we propose that the quasi-perceptual experience of articulator movement and the subsequent auditory percept are induced by the same sequential estimation mechanism. However, the "cancellation" deriving from somatosensory and auditory feedback, which is generated by the overt outputs during production, is absent in the imagery case (**Figure 1**). Therefore, similar to the case of motor imagery (Jeannerod, 1994, 1995), the feeling of articulator movement is the result of residual somatosensory representation resulting from motor simulation; the subsequent auditory perceptual experience, we suggest, is the residual auditory representation from the second estimation stage.

On the basis of sequential estimation account, particular neural activity patterns for the two sequential estimates are predicted to occur in a temporal order. Specifically, an auditory pattern should follow a somatosensory one during mental imagery of speech. Applying a novel multivariate technique (Tian and Huber, 2008; Tian et al., 2011) to MEG data, we observed such a temporal order for somatosensory and auditory estimations during articulation imagery (Tian and Poeppel, 2010), manifested in the sequential activity patterns over modality-specific regions at different latencies (**Figure 2**). A left parietal response pattern was observed during *articulation imagery* at the same latency as when motor responses occurred in the articulation

condition[1]. Following such a left parietal response pattern, a second pattern was identified at a latency of 150–170 ms after the parietal response. This second pattern was very similar to the response elicited by external auditory stimuli. Moreover, in a further experimental condition, *hearing imagery*, we also observed an auditory-like neural response pattern; however, its latency was faster than the same auditory pattern observed in *articulation imagery*. The existence of these two spatially highly similar auditory-like neural representations, with different latencies for *articulation* versus *hearing* imagery tasks, suggests that the same (or strikingly similar) neural representations can be generated either by internal estimation or by memory retrieval, based on contextual variation and task demands.

Note that the auditory estimation is presumably formed along the canonical auditory hierarchy, but the induction process will be in reversed order. That is to say, the abstract representation is (re-)constructed first in higher level associative areas and conveyed to a perceptual-sensory representation in lower areas. The observation of neural activity in the posterior superior temporal

---

[1] Because previous findings suggest that the time courses for completing execution and imagery are comparable (Decety and Michel, 1989; Decety et al., 1989; Sirigu et al., 1995, 1996), the observed neural responses over parietal regions presumably mediate somatosensory estimation.
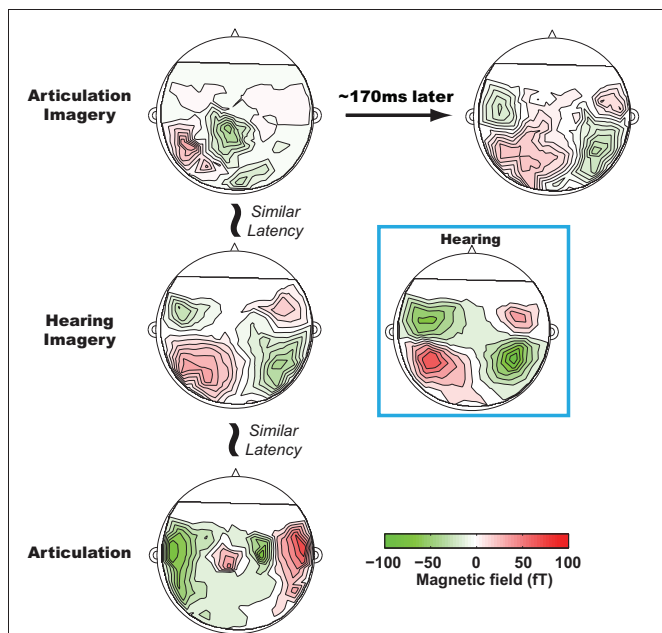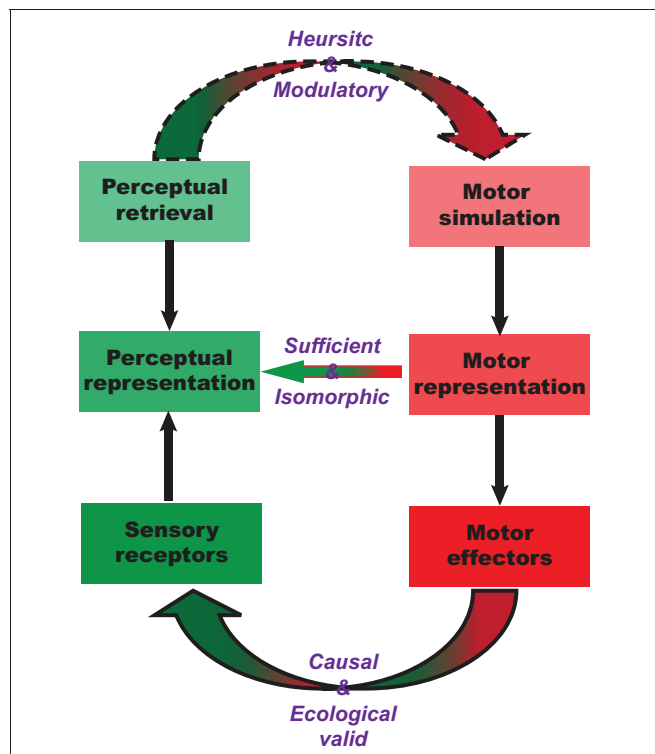
FIGURE 2 | Results from Tian and Poeppel (2010). The sequential estimation during articulation imagery revealed by MEG recordings. All plots are MEG topographies (response patterns) when participants actually speak (lower row), imagine hearing (middle row), and imagine speaking (top row). The activity patterns in the first column are temporally aligned with the onset of articulation movement. At a similar latency, bilateral frontal, bilateral temporal, and left lateralized parietal activity patterns are observed in articulation, hearing imagery, and articulation imagery conditions. In articulation imagery, about 150–170 ms later after the parietal activity, bilateral temporal activity is also observed. All the bilateral temporal activity patterns in *hearing imagery* and *articulation imagery* resemble the topography of the auditory response during actual hearing (highlighted in a blue box, response pattern when participants listen to the same auditory stimuli as in other conditions).

FIGURE 3 | Sufficiency and necessity between motor simulation and perceptual estimation. The characteristics of the proposed motor simulation and perceptual estimation processes, and the nature of motor involvement during perceptual tasks. The internal motor simulation can take a similar path as motor preparation to derive a corresponding motor representation that in turn derives associated perceptual representations in a one-to-one fashion. Such one-to-one mapping is the same as the one in the external connections between the similar motor action and perceptual consequences. In the other direction, when the perceptual representation is needed, different paths can be taken. It can rely on memory retrieval to directly recreate the perceptual representation. It can also take another less demanding path that relies on the motor simulation to derive the associated perceptual representation.

sulcus (pSTS) during silent speaking (Price et al., 2011) could be the result of an earlier reconstruction. Whereas the observations of similarity between responses to mental imagery and to external stimulation, such as in visual (e.g., Kosslyn et al., 1999) and auditory (e.g., **Figure 2**, Tian and Poeppel, 2010) domains, are the results of process continuation to lower perceptual-sensory regions. How much further back the reconstruction process might go seems to depend on the sensory modality and demands of the imagery tasks (Kosslyn and Thompson, 2003; Kraemer et al., 2005; Zatorre and Halpern, 2005).

INTERNAL SIMULATION-ESTIMATION AND RELATION TO SENSORY-MOTOR INTEGRATION

Mental imagery of speech exemplifies a top-down mechanism for sensory-motor integration. The proposal here is motor simulation and sequential estimation. In the first part of this section, we describe the nature of this sequential transformation between motor, somatosensory, and other perceptual systems. We postulate that there is a one-to-one transformation between motor simulation and somatosensory estimation, as well as isomorphic mapping between somatosensory estimation and subsequent

perceptual estimation (**Figure 3**). The entire transformation process is carried out in a continuous manner, beginning with motor simulation, then somatosensory estimation, and ending with modality-specific perceptual estimation. In the second part of this section, we argue that the implementation of motor simulation depends on context and task demands and may only exert modulatory effects on perception.

MOTOR-TO-SENSORY MAPPING: ISOMORPHISM VIA ESTABLISHED CONNECTIONS

The central idea underpinning motor simulation and subsequent perceptual estimation is the conjectured *one-to-one mapping* or *isomorphism* between mental and physical processes. This isomorphism has been proposed for motor simulation (Jeannerod, 1994) and visual mental rotation (Shepard and Cooper, 1986): the intermediate stages of the internal process must have a one-to-one correspondence to intermediate stages of an actualized physical process. We extend this isomorphism to the associations

between the motor simulation and perceptual estimation: the one-to-one mapping between the trajectory of motor simulation and perceptual estimation is a close analog to the causal relation between motor outputs and perceptual changes. That is, not only should the starting and ending points of an action simulation lead to the initiation and results of perceptual estimation, but intermediate points on this action simulation trajectory should result in a sequence of perceptual estimates, even though no external signals are physically presented. Notice that the analogy between internal simulation-estimation and external action-perception does not require the preservation of first-order isomorphism: only the one-to-one relation in the transformation of internal representation from motor to perceptual systems is required, as if the action was actually performed and the percept was actually induced.

The isomorphic transformation from motor to perceptual systems relies on the established internal associations between motor and perceptual representations, which are presumably formed following the causal and ecologically valid sequential occurrence of action-perception pairs, through the mechanisms of associative learning (Mahon and Caramazza, 2008). For example, the movement of articulators can induce somatosensory feedback and subsequent auditory perception of one's own speech. On the basis of the occurrence order (action first, then somatosensory activation, followed by auditory perception), an internal association can be established to link a particular movement trajectory of articulators with the specific somatosensory sensation, followed by a given auditory perception of speech. Note that we do *not* exclude the possible existence of a parallel estimation process that links motor simulation to somatosensory and auditory systems separately (Guenther et al., 2006; Price et al., 2011). Such an additional mechanisms which runs in parallel may mediate the early comparison between auditory estimation from an articulatory plan and intended auditory targets during speech production (Hickok, 2012). The redundancy of the compensation in somatosensory and auditory domains offers a hint for the co-existence of sequential and parallel estimation structures (Lametti et al., 2012). We suggest that the serial updating structure as one of the possible underlying estimation mechanisms naturally follows the biological sequences, providing advantages in learning and plasticity during development as well as online speech control.

Speech-induced suppression and enhancement caused by feedback perturbation provides strong evidence for the one-to-one mapping between motor simulation and estimation of perceptual consequences. When participants speak and listen to their own speech, the evoked auditory responses are smaller compared with the auditory responses to the same speech played back without spoken outputs (Numminen et al., 1999; Houde et al., 2002; Eliades and Wang, 2003, 2005; Ventura et al., 2009). However, when the auditory feedback is perturbed (manipulating, e.g., pitch or format frequencies), the auditory responses during speaking become larger compared with the ones during playback (Eliades and Wang, 2008; Tourville et al., 2008; Zheng et al., 2010; Behroozmand et al., 2011). The suppression caused by articulation demonstrates that an internal signal labels the onset of movement and down-regulates sensitivity to subsequent auditory

perception (general suppression). However, the enhancement caused by feedback perturbation suggests that the internal signal during articulation is not a generic gain control mechanism for all auditory stimuli, but rather provides a precise perceptual prediction and only blocks the feedback that is identical to the prediction. In other words, there is a one-to-one mapping between motor simulation and auditory estimation, and the precise auditory consequence can be predicted based on particular motor trajectory.

The hypothesized intermediate neurocomputational step of somatosensory estimation that lies between motor simulation and auditory estimation has also been suggested by recent experiments. The sequential neural activity underlying somatosensory and auditory estimation has been observed during articulation imagery using MEG (Tian and Poeppel, 2010), as discussed above (**Figure 2**). Lesions over the left pars opercularis (pOp) in the inferior frontal gyrus (IFG) as well as adjacent to the left supramarginal gyrus (SMG) in parietal cortex correlate with the ability to imagine speech; this demonstrates the possible neural implementation underlying the proposed simulation and (somatosensory) estimation (Geva et al., 2011). Moreover, the causal role of somatosensory feedback in speech perception has also been demonstrated (Ito et al., 2009). There, participants were asked to listen to ambiguous stimuli (e.g., head-had vowel continuum) while their facial skin was manipulated with a robotic device. When the skin at the side of mouth was stretched upward (as in the case of pronouncing "head"), participants were biased toward hearing the ambiguous sound as "head." That is, the somatosensory status affected the auditory perception in a systematic way: there was a one-to-one representational mapping between somatosensory and auditory systems.

THE SIMULATION-ESTIMATION PROCESS IN PERCEPTION

The debates surrounding motor theories of perception and cognition [see the review by Scheerer (1984)] have heated up since the discovery of the putative "mirror neuron system" in monkeys (di Pellegrino et al., 1992; Gallese et al., 1996; see Rizzolatti and Craighero, 2004 for a review) and the observation of motor activity observed during numerous perceptual studies in humans (e.g., Rizzolatti et al., 1996; Iacoboni et al., 1999; Buccino et al., 2001; Wilson et al., 2004). Although these debates are beyond the scope of this review, the proposed mechanism of sequential estimation following motor simulation may provide insight to reconcile some of the observations, providing a top-down perspective.

We propose, building on arguments in the recent literature (Mahon and Caramazza, 2008; Hickok, 2009; Lotto et al., 2009; Rumiati et al., 2010), that the deployment of motor simulation in perceptual tasks is (1) strategy-dependent and (2) exerts modulatory effects on the formation of perceptual representations. That is, the selection of motor involvement in perceptual tasks depends on context and task demands. It is a top-down strategic step to provide modality-specific representations in advance (cf. Moulton and Kosslyn, 2009) and reduce perceptual variance by generating more precise estimation (Mahon and Caramazza, 2008; but also see Pulvermüller

and Fadiga, 2010 for an opposite view from a embodied perspective).

The implementation of motor-to-sensory transformations is strategy-dependent

We describe two types of evidence. First, the recruitment/involvement of motor simulation is influenced by task demands. For example, motor imagery can be performed from a "first person" perspective that relies on kinesthetic feeling, in contrast with when a task is executed from a "third person" perspective in which the action-related visual changes are recreated (Jeannerod, 1994, 1995). Reaction times of hand rotation imagery showed an interaction between imagery perspectives and limb posture: when asked to imagine rotating their hands from first person perspective, participants responded faster when their hands were on the lap but slower when their hands are in the back; the reverse pattern was observed when imagining from third person perspective (Sirigu and Duhamel, 2001). Activation in the motor system was observed when participants were explicitly told to imagine rotating an object with their own hands, but was absent when they were told to imagine rotating the same object with a robotic motor (Kosslyn et al., 2001). Both behavioral and neuroimaging results highlight that the task demands influence the implementation of neural pathways that mediate either direct simulation (memory retrieval) or motor simulation-estimation (transformation between motor and perceptual systems).

Second, motor-to-sensory transformations are influenced by context and the properties of stimuli. For example, neural responses in frontal motor regions have been observed during observation of meaningful actions, contrasted with occipital activity for meaningless actions (Decety et al., 1997). Relatedly, when participants mentally rotated their hands, premotor, primary motor, and posterior parietal cortices were activated. However, frontal motor areas were silent when they mentally rotated objects (Kosslyn et al., 1998). These results suggest that contextual influence and task demands can determine the implementation of motor simulation in a top-down, voluntary, strategic way.

In the context of action observation, understanding/comprehension and imitation could be the result of heuristic engagement of motor simulation. That is, humans can deploy a top-down mechanism that transfers perceptual goals into the motor domain and initiates motor simulation to derive perceptual consequences (**Figure 3**). The strategic and heuristic initiation of motor involvement can be considered as a top-down mental imagery process (possibly exclusive to humans) (cf. Iacoboni et al., 1999; Papeo et al., 2009), wherein the motor action is internally simulated and perceptual consequences estimated thereafter (cf. Tkach et al., 2007).

Modulatory function of motor simulation on perception

The major evidence supporting a modulatory role of motor simulation in perception (rather than a primary causal role) comes from lesion studies. For example, lesions in the frontal lobe only caused deficits in action production, whereas lesions in the parietal lobe caused deficits both during production and

perception of movement (Heilman et al., 1982). A deficit in gesture recognition has also been linked to inferior parietal cortex lesions but not lesions in the frontal lobe (Buxbaum et al., 2005). Action comprehension also relies on a network that includes inferior parietal cortex but not IFG (Saygin et al., 2004). Although patients with IFG lesions demonstrated deficits in action comprehension in the same study, the static stimuli (pictures of pantomimed actions or objects) could require participants to implement the strategy of motor simulation to form the dynamic display of action and to derive the perceptual consequences so that they can fulfill the action-object association task. Such lesion results indicate that a damaged motor system (and the deficits in motor simulation) dissociates from action-perception and comprehension. The abstract meaning of motor action is probably "stored" in parietal regions, and the motor simulation mediated by frontal regions is one of many paths to access the stored representation (in line with our proposed *simulation* over frontal cortex and *estimation* over parietal cortex). Therefore, motor simulation to estimate perceptual consequences is only modulatory and not necessary for perceptual tasks.

Analogous to the advantage of multisensory integration in minimizing perceptual variance (Ernst and Banks, 2002; Alais and Burr, 2004; van Wassenhove et al., 2005; von Kriegstein and Giraud, 2006; Morgan et al., 2008; Poeppel et al., 2008; Fetsch et al., 2009), the modulatory effects of motor simulation convey benefits by providing additional, more detailed information to enrich the perceptual representation using internal sequential estimation mechanism (cf. Mahon and Caramazza, 2008). Human observers can adopt motor strategies to provide more precise perceptual representations and deal with perceptual ambiguity, for example in the case of speech perception. That is, the motor simulation and estimation can provide improved priors to reduce perceptual variance.

In summary, various perceptual tasks can use the motor system to derive perceptual consequences, by implementing the same top-down motor simulation and perceptual estimation mechanism, as in mental imagery of speech. We hypothesize that this motor simulation is modulatory and only serves as one of many possible corridors to induce perceptual representations. Such strategies of sensory-to-motor and motor-to-sensory transformation would be implemented depending on task demands and contextual influence.

IMPLICATIONS FOR THE NEURAL CORRELATES OF SOME DISORDERS

In this section we argue that the internal processes of motor simulation and estimation, revealed originally for the mental imagery of speech, can shed light on possible neural correlates of certain disorders, including auditory hallucinations, stuttering, and phantom limb syndrome. We outline some working hypotheses regarding these disorders, complementing other existing hypotheses. It is suggested that the proposed idea for mental imagery generation, motor simulation, and sequential perceptual estimation, points to the practical value of mental imagery research for understanding the internal mechanisms of such neural disorders.

AUDITORY HALLUCINATIONS: INTACT ESTIMATION VERSUS BROKEN MONITORING

Internal simulation and sequential estimation has been proposed to be a way to distinguish between the perceptual changes caused by self-generated actions and exogenous external events (Blakemore and Frith, 2003; Jeannerod and Pacherie, 2004; Tsakiris and Haggard, 2005). The perceptual consequences of intended movement can be predicted, and the processing of external sensory feedback can be dampened by the internal prediction, such as in the case of speech production (e.g., Houde et al., 2002; Eliades and Wang, 2003, 2005) and somatosensory perception in tickling (e.g., Blakemore et al., 1998). This suggests that the action-induced perceptual signals are identified as self-generated and cancelled by the virtually identical representation generated by internal perceptual prediction. However for patients suffering from auditory hallucinations, deficits of these hypothesized dampening mechanisms for self-induced perceptual changes have been observed in both somatosensory (e.g., Blakemore et al., 2000) and auditory (e.g., Ford et al., 2007; Heinks-Maldonado et al., 2007) domains. These results suggest that patients with auditory hallucinations cannot separate self-induced from external-induced perceptual signals.

Critically, deficits of distinguishing self-induced from externally induced perceptual changes are not enough to account for auditory hallucinations, because the positive symptoms typically occur in the absence of any external stimuli. There must exist an internal mechanism to induce the auditory representations that are then misattributed to an external source/voice. In fact, we face a similar situation during mental imagery: the neural representations mediating perception and mental imagery are very similar, but there is no mechanism in the perceptual system to distinguish them. A *source monitoring function* is required to keep track of the origins of the perceptual neural representation. Therefore, we hypothesize that a higher order function monitors and distinguishes internally versus externally induced neural representations. Such a monitoring operation is functionally independent from the perceptual estimation process that internally reconstructs the perceptual representation. Under this hypothesis, auditory hallucinations are caused by incorrect operation of the monitoring function, resulting in incorrectly labeling the self-induced auditory representation during the intact internal perceptual estimation processes.

Computationally, the independence of the monitoring function versus internal simulation and estimation is demonstrated by the nuanced differences between *corollary discharge* and the *efference copy* [see the review by Crapse and Sommer (2008)]. The efference copy is a duplicate of the planned motor command and provides the dynamics of an action trajectory that can be used to estimate the perceptual consequences (von Holst and Mittelstaedt, 1950, 1973). Corollary discharge is a more general motor related mechanism that can be available at all levels of a motor process. The corollary discharge does not necessarily contain the same representational information as an efference copy; rather, it serves as a generic signal to inform sensory-perceptual systems of the potential occurrence of perceptual changes caused by one's own actions (Sperry, 1950). In the case of speech articulation, these two functions originate at the same stage of

motor simulation, but their functional roles are still separate. The efference copy is used to estimate the detailed perceptual consequences, whereas the corollary discharge labels the internally and externally induced perceptual consequences.

Empirically, the finding that auditory hallucination patients can generate inner speech (e.g., Shergill et al., 2003) demonstrates the relatively intact motor-to-sensory transformation function. The neural responses in IFG and superior temporal gyrus/sulcus (STG/STS) were observed during auditory hallucinations, hinting at the derivation of auditory perceptual consequences from motor simulation during the positive symptom (e.g., McGuire et al., 1993; Shergill et al., 2003). Moreover, the left lateralization during covert speech versus right lateralization during auditory hallucinations offers tantalizing hints about the independence between self-monitoring and the sequential simulation-estimation (Sommer et al., 2008).

We summarize the hypothetical mechanistic account for auditory hallucinations (of this type) as follows: when patients prepare to articulate speech covertly or subvocally (either consciously or unconsciously), the internal motor simulation leads to perceptual estimation (intact efference copy). But the source monitoring process malfunctions (broken corollary discharge). Therefore, the internal prediction of a perceptual consequence, which has the same neural representation as an external perception, is erroneously interpreted as the result of external sources, resulting in an auditory hallucination.

STUTTERING: NOISY ESTIMATION AND CORRECTION PROCESSES

The comparison between internal estimation and external feedback provides information to fine-tune motor control. However, if the internal estimate from motor simulation malfunctions and generates imprecise perceptual predictions, an inaccurate or incorrect feedback control signal would be conveyed. Stuttering could be an example of such erroneous correction. We suggest, along the lines of similar theories (Max et al., 2004; Hickok et al., 2011), that one of the neural mechanisms causing stuttering is a deficit in the motor-to-sensory transformation. That is, the noisy perceptual estimation is mismatched to the external feedback. Such a discrepancy would signal an incorrect error message, and the feedback control system would interpret such an apparent error as the requirement to correct motor action. Hence, unnecessary attempts would be performed to modify the correct articulation, resulting in repetitive/prolonged sound or silent pauses/blocks.

The noise in the estimation process can come both from the somatosensory and auditory domains (since there is sequential estimation). Stutterers showed speed and latency deficits when required to sequentially update articulator movement (Caruso et al., 1988). Smaller magnitude compensation with longer latency adjustment to the perturbation on the jaws was also observed in stutterers (Caruso et al., 1987). In the auditory domain, smaller magnitude compensation to the perturbation of F1 formant in auditory feedback is observed (Cai et al., 2012). The inaccurate compensation to external perturbation in both somatosensory and auditory domains (with intact somatosensory and auditory processes) demonstrates that inaccurate prediction in both domains could be causal for stuttering.

Interestingly, dramatically altering auditory feedback (e.g., by delaying feedback onset or shifting frequency) can enhance speech fluency in people who stutter (Martin and Haroldson, 1979; Stuart et al., 1997, 2008). The improvement could be because the magnitude of error signals is scaled down when the distance between feedback and prediction is beyond some threshold, so that fewer correction attempts are made.

PHANTOM LIMBS: MISMATCH BETWEEN INTERNAL ESTIMATION AND EXTERNAL FEEDBACK

The mismatch between internal prediction and external feedback could also be caused by an acute change of conditions leading to the absence of feedback. One such example is the phantom limb phenomenon, where amputees feel control over a lost limb (phantom limb) accompanied with chronic and sometimes acute pain. We hypothesize that the apparent awareness and control of a lost limb occurs as follows: the missing somatosensory feedback is "replaced" by the results of internal estimation (cf. Frith et al., 2000; Fotopoulou et al., 2008). Such a hypothesis is similar to the mislabeling of the internal estimation as an external perception (due to the malfunction of source monitoring) in auditory hallucinations.

The causes of pain in phantom limbs are more intriguing. The most significant physical changes are loss of proprioception, or somatosensory afference, after lost limbs. Because motor control as well as motor simulation of the lost limb are still in some sense valid (e.g., Raffin et al., 2012), we hypothesize that a mismatch between the intact internal estimation and absent external somatosensory feedback can cause the pain associated with phantom limbs. In fact, consistent with our hypothesis, limb pain can be induced in normal participants by mismatching visual and proprioceptive feedback (McCabe et al., 2005) and spinal cord injured patients report that neuropathic pain increases while they imagine moving their ankles (Gustin et al., 2008).

This mismatch hypothesis may represent an intermediate step between cortical reorganization and pain induction. Lost limbs cause reorganization in both motor (Maihöfner et al., 2007) and somatosensory (Maihöfner et al., 2003) cortices, and pain reduction has been demonstrated to correlate with more granular organization in the same areas (MacIver et al., 2008). Motor imagery can lead to cortical reorganization that correlates with pain reduction in phantom limbs (Moseley, 2006). Seeing the movement of the opposite functioning arm in a mirror can reduce the pain associated with the phantom limb (Ramachandran et al., 1995). Such behavioral and psychological training can provide more precise topographic maps in both motor and somatosensory cortices and hence reduce the inaccurate motor firing caused by the "take over" effect (e.g., cortex of lip movement expand to the cortex mediated a lost hand), as well as erroneous somatosensory estimation. The internal estimation hypothesis offers a new perspective on pain induction. However, there is neither a clear pain center (Mazzola et al., 2012) nor a mechanistic pain induction account (Flor, 2002). Further research is needed to understand how the proposed mismatch hypothesis could underpin pain induction.

CONCLUSION

In this perspective, we argued that mental imagery is an internal predictive process. Using mental imagery of speech as an example, we demonstrated a variety of principles underlying how the mechanism of motor simulation and sequential perceptual estimation in mental imagery works. We conclude that the simulation-estimation mechanism provides a novel conceptual and practical perspective that allows for new types of research on predictive functions and sensory-motor integration, as well as stimulating some new insights into several neural disorders. Typically, mental imagery has been studied in cognitive psychology and cognitive neuroscience, while the concepts of internal forward models (and sensory-motor integration) are the focus of motor control research from an engineering perspective. Our atypical pairing of internal models as an additional source for mental imagery yields, in our view, some provocative new angles on mental imagery in both basic research and applied contexts.

REFERENCES

Alais, D., and Burr, D. (2004). The ventriloquist effect results from near-optimal bimodal integration. *Curr. Biol.* 14, 257–262.

Behroozmand, R., Liu, H., and Larson, C. R. (2011). Time-dependent neural processing of auditory feedback during voice pitch error detection. *J. Cogn. Neurosci.* 23, 1205–1217.

Bensafi, M., Porter, J., Pouliot, S., Mainland, J., Johnson, B., Zelano, C., et al. (2003). Olfactomotor activity during imagery mimics that during perception. *Nat. Neurosci.* 6, 1142–1144.

Blakemore, S. J., and Frith, C. (2003). Self-awareness and action. *Curr. Opin. Neurobiol.* 13, 219–224.

Blakemore, S. J., Smith, J., Steel, R., Johnstone, E., and Frith, C. (2000). The perception of self-produced sensory stimuli in patients with auditory hallucinations and passivity experiences: evidence for a breakdown in self-monitoring. *Psychol. Med.* 30, 1131–1139.

Blakemore, S. J., Wolpert, D. M., and Frith, C. D. (1998). Central cancellation of self-produced tickle sensation. *Nat. Neurosci.* 1, 635–640.

Buccino, G., Binkofski, F., Fink, G. R., Fadiga, L., Fogassi, L., Gallese, V., et al. (2001). Action observation activates premotor and parietal areas in a somatotopic manner: an fMRI study. *Eur. J. Neurosci.* 13, 400–404.

Buxbaum, L. J., Kyle, K. M., and Menon, R. (2005). On beyond mirror neurons: internal representations subserving imitation and recognition of skilled object-related actions in humans. *Cogn. Brain Res.* 25, 226–239.

Cai, S., Beal, D. S., Ghosh, S. S., Tiede, M. K., Guenther, F. H., and Perkell, J. S. (2012). Weak responses to auditory feedback perturbation during articulation in persons who stutter: evidence for abnormal auditory-motor transformation. *PLoS ONE* 7:e41830. doi: 10.1371/journal.pone.0041830

Caruso, A. J., Abbs, J. H., and Gracco, V. L. (1988). Kinematic analysis of multiple movement coordination during speech in stutterers. *Brain* 111, 439.

Caruso, A. J., Gracco, V., and Abbs, J. H. (1987). "A speech motor control perspective on stuttering: preliminary observations," in *Speech Motor Dynamics in Stuttering,* eds H. F. M. Peters and W. Hulstijn (Wien, Austria: Springer-Verlag), 245–258.

Crapse, T. B., and Sommer, M. A. (2008). Corollary discharge across the animal kingdom. *Nat. Rev. Neurosci.* 9, 587–600.

Decety, J. (1996). The neurophysiological basis of motor imagery. *Behav. Brain Res.* 77, 45–52.

Decety, J., Grezes, J., Costes, N., Perani, D., Jeannerod, M., Procyk, E., et al. (1997). Brain activity during

observation of actions. Influence of action content and subject's strategy. *Brain* 120, 1763.

Decety, J., Jeannerod, M., and Prablanc, C. (1989). The timing of mentally represented actions. *Behav. Brain Res.* 34, 35–42.

Decety, J., and Michel, F. (1989). Comparative analysis of actual and mental movement times in two graphic tasks. *Brain Cogn.* 11, 87–97.

Decety, J., Perani, D., Jeannerod, M., Bettinardi, V., Tadary, B., Woods, R., et al. (1994). Mapping motor representations with positron emission tomography. *Nature* 371, 600–602.

Dechent, P., Merboldt, K. D., and Frahm, J. (2004). Is the human primary motor cortex involved in motor imagery? *Cogn. Brain Res.* 19, 138–144.

Desmurget, M., Reilly, K. T., Richard, N., Szathmari, A., Mottolese, C., and Sirigu, A. (2009). Movement intention after parietal cortex stimulation in humans. *Science* 324, 811.

Desmurget, M., and Sirigu, A. (2009). A parietal-premotor network for movement intention and motor awareness. *Trends Cogn. Sci.* 13, 411–419.

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

Djordjevic, J., Zatorre, R., Petrides, M., Boyle, J., and Jones-Gotman, M. (2005). Functional neuroimaging of odor imagery. *Neuroimage* 24, 791–801.

Ehrsson, H. H., Geyer, S., and Naito, E. (2003). Imagery of voluntary movement of fingers, toes, and tongue activates corresponding body-part-specific motor representations. *J. Neurophysiol.* 90, 3304–3316.

Eliades, S. J., and Wang, X. (2003). Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *J. Neurophysiol.* 89, 2194.

Eliades, S. J., and Wang, X. (2005). Dynamics of auditory–vocal interaction in monkey auditory cortex. *Cereb. Cortex* 15, 1510.

Eliades, S. J., and Wang, X. (2008). Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature* 453, 1102–1106.

Ernst, M. O., and Banks, M. S. (2002). Humans integrate visual and haptic information in a statistically optimal fashion. *Nature* 415, 429–433.

Fetsch, C. R., Turner, A. H., DeAngelis, G. C., and Angelaki, D. E. (2009). Dynamic reweighting of visual and vestibular cues during self-motion perception. *J. Neurosci.* 29, 15601.

Fisher, J. C. (2006). Does simulation theory really involve simulation? *Philos. Psychol.* 19, 417–432.

Flor, H. (2002). Phantom-limb pain: characteristics, causes, and treatment. *Lancet Neurol.* 1, 182–189.

Ford, J. M., Gray, M., Faustman, W. O., Roach, B. J., and Mathalon, D. H. (2007). Dissecting corollary discharge dysfunction in schizophrenia. *Psychophysiology* 44, 522–529.

Fotopoulou, A., Tsakiris, M., Haggard, P., Vagopoulou, A., Rudd, A., and Kopelman, M. (2008). The role of motor intention in motor awareness: an experimental study on anosognosia for hemiplegia. *Brain* 131, 3432–3442.

Foxe, J. J., Wylie, G. R., Martinez, A., Schroeder, C. E., Javitt, D. C., Guilfoyle, D., et al. (2002). Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. *J. Neurophysiol.* 88, 540.

Frith, C. D., Blakemore, S. J., and Wolpert, D. M. (2000). Abnormalities in the awareness and control of action. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 355, 1771–1788.

Fu, K. M. G., Johnston, T. A., Shah, A. S., Arnold, L., Smiley, J., Hackett, T. A., et al. (2003). Auditory cortical neurons respond to somatosensory stimulation. *J. Neurosci.* 23, 7510–7515.

Gallese, V., Fadiga, L., Fogassi, L., and Rizzolatti, G. (1996). Action recognition in the premotor cortex. *Brain* 119, 593.

Gerardin, E., Sirigu, A., Lehericy, S., Poline, J.-B., Gaymard, B., Marsault, C., et al. (2000). Partially overlapping neural networks for real and imagined hand movements. *Cereb. Cortex* 10, 1093–1104.

Geva, S., Jones, P. S., Crinion, J. T., Price, C. J., Baron, J. C., and Warburton, E. A. (2011). The neural correlates of inner speech defined by voxel-based lesion–symptom mapping. *Brain* 134, 3071–3082.

Goldman, A. I. (1989). Interpretation psychologized. *Mind Lang.* 4, 161–185.

Grush, R. (2004). The emulation theory of representation: motor control, imagery, and perception. *Behav. Brain Sci.* 27, 377–396.

Guenther, F. H., Ghosh, S. S., and Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain Lang.* 96, 280–301.

Gustin, S. M., Wrigley, P. J., Gandevia, S. C., Middleton, J. W., Henderson, L. A., and Siddall, P. J. (2008). Movement imagery increases pain in people with neuropathic pain following complete thoracic spinal cord injury. *Pain* 137, 237–244.

Hanakawa, T., Immisch, I., Toma, K., Dimyan, M. A., Van Gelderen, P., and Hallett, M. (2003). Functional properties of brain areas associated with motor execution and imagery. *J. Neurophysiol.* 89, 989–1002.

Heilman, K. M., Rothi, L. J., and Valenstein, E. (1982). Two forms of ideomotor apraxia. *Neurology* 32, 342–342.

Heinks-Maldonado, T. H., Mathalon, D. H., Houde, J. F., Gray, M., Faustman, W. O., and Ford, J. M. (2007). Relationship of imprecise corollary discharge in schizophrenia to auditory hallucinations. *Arch. Gen. Psychiatry* 64, 286.

Hesslow, G. (2002). Conscious thought as simulation of behaviour and perception. *Trends Cogn. Sci.* 6, 242–247.

Hickok, G. (2009). Eight problems for the mirror neuron theory of action understanding in monkeys and humans. *J. Cogn. Neurosci.* 21, 1229–1243.

Hickok, G. (2012). Computational neuroanatomy of speech production. *Nat. Rev. Neurosci.* 13, 135–145.

Hickok, G., Houde, J., and Rong, F. (2011). Sensorimotor integration in speech processing: computational basis and neural organization. *Neuron* 69, 407–422.

Houde, J. F., Nagarajan, S. S., Sekikara, K., and Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: an MEG study. *J. Cogn. Neurosci.* 14, 1125–1138.

Iacoboni, M., Woods, R. P., Brass, M., Bekkering, H., Mazziotta, J. C., and Rizzolatti, G. (1999). Cortical mechanisms of human imitation. *Science* 286, 2526.

Ishai, A., Haxby, J. V., and Ungerleider, L. G. (2002). Visual imagery of famous faces: effects of memory and attention revealed by fMRI. *Neuroimage* 17, 1729–1741.

Ito, T., Tiede, M., and Ostry, D. J. (2009). Somatosensory function in speech perception. *Proc. Natl. Acad. Sci. U.S.A.* 106, 1245.

Jeannerod, M. (1994). The representing brain: neural correlates of motor intention and imagery. *Behav. Brain Sci.* 17, 187–202.

Jeannerod, M. (1995). Mental imagery in the motor context. *Neuropsychologia* 33, 1419–1432.

Jeannerod, M. (2001). Neural simulation of action: a unifying mechanism for motor cognition. *Neuroimage* 14, S103–S109.

Jeannerod, M., and Pacherie, E. (2004). Agency, simulation and self-identification. *Mind Lang.* 19, 113–146.

Kosslyn, S. M. (1994). *Image and Brain: The Resolution of the Imagery Debate.* Cambridge, MA: MIT Press.

Kosslyn, S. M. (2005). Mental images and the brain. *Cogn. Neuropsychol.* 22, 333–347.

Kosslyn, S. M., Alpert, N. M., Thompson, W. L., Chabris, C. F., Rauch, S. L., and Anderson, A. K. (1994). Identifying objects seen from different viewpoints A PET investigation. *Brain* 117, 1055.

Kosslyn, S. M., Digirolamo, G. J., Thompson, W. L., and Alpert, N. M. (1998). Mental rotation of objects versus hands: neural mechanisms revealed by positron emission tomography. *Psychophysiology* 35, 151–161.

Kosslyn, S. M., Ganis, G., and Thompson, W. L. (2001). Neural foundations of imagery. *Nat. Rev. Neurosci.* 2, 635–642.

Kosslyn, S. M., Pascual-Leone, A., Felician, O., Camposano, S., Keenan, J., Ganis, G., et al. (1999). The role of area 17 in visual imagery: convergent evidence from PET and rTMS. *Science* 284, 167.

Kosslyn, S. M., and Thompson, W. L. (2003). When is early visual cortex activated during visual mental imagery? *Psychol. Bull.* 129, 723–746.

Kosslyn, S. M., Thompson, W. L., Wraga, M., and Alpert, N. M. (2001). Imagining rotation by endogenous versus exogenous forces: distinct neural mechanisms. *Neuroreport* 12, 2519.

Kraemer, D. J. M., Macrae, C. N., Green, A. E., and Kelley, W. M. (2005). Musical imagery: sound of silence activates auditory cortex. *Nature* 434, 158.

Lametti, D. R., Nasir, S. M., and Ostry, D. J. (2012). Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. *J. Neurosci.* 32, 9351–9358.

Lotto, A. J., Hickok, G. S., and Holt, L. L. (2009). Reflections on mirror neurons and speech perception. *Trends Cogn. Sci.* 13, 110–114.

Lotze, M., Montoya, P., Erb, M., Hulsmann, E., Flor, H., Klose, U., et al. (1999). Activation of cortical and cerebellar motor areas during executed and imagined hand movements: an fMRI study. *J. Cogn. Neurosci.* 11, 491–501.

MacIver, K., Lloyd, D., Kelly, S., Roberts, N., and Nurmikko, T. (2008). Phantom limb pain, cortical reorganization and the therapeutic effect of mental imagery. *Brain* 131, 2181.

Mahon, B. Z., and Caramazza, A. (2008). A critical look at the embodied cognition hypothesis and a new proposal for grounding conceptual content. *J. Physiol. Paris* 102, 59–70.

Maihöfner, C., Baron, R., DeCol, R., Binder, A., Birklein, F., Deuschl, G., et al. (2007). The motor system shows adaptive changes in complex regional pain syndrome. *Brain* 130, 2671.

Maihöfner, C., Handwerker, H. O., Neundörfer, B., and Birklein, F. (2003). Patterns of cortical reorganization in complex regional pain syndrome. *Neurology* 61, 1707–1715.

Martin, R., and Haroldson, S. K. (1979). Effects of five experimental treatments on stuttering. *J. Speech Hear. Res.* 22, 132.

Max, L., Guenther, F. H., Gracco, V. L., Ghosh, S. S., and Wallace, M. E. (2004). Unstable or insufficiently activated internal models and feedback-biased motor control as sources of dysfluency: a theoretical model of stuttering. *Contemp. Issues Commun. Sci. Disord.* 31, 105–122.

Mazzola, L., Isnard, J., Peyron, R., and Mauguière, F. (2012). Stimulation of the human cortex and the experience of pain: Wilder Penfield's observations revisited. *Brain* 135, 631–640.

McCabe, C., Haigh, R., Halligan, P., and Blake, D. (2005). Simulating sensory–motor incongruence in healthy volunteers: implications for a cortical model of pain. *Rheumatology* 44, 509.

McGuire, P., Murray, R., and Shah, G. (1993). Increased blood flow in Broca's area during auditory hallucinations in schizophrenia. *Lancet* 342, 703–706.

Meister, I. G., Krings, T., Foltys, H., Boroojerdi, B., Müller, M., Töpper, R., et al. (2004). Playing piano in the mind: an fMRI study on music imagery and performance in pianists. *Cogn. Brain Res.* 19, 219–228.

Morgan, M. L., DeAngelis, G. C., and Angelaki, D. E. (2008). Multisensory integration in macaque visual cortex depends on cue reliability. *Neuron* 59, 662–673.

Moseley, G. L. (2006). Graded motor imagery for pathologic pain. *Neurology* 67, 2129–2134.

Moulton, S. T., and Kosslyn, S. M. (2009). Imagining predictions: mental imagery as mental emulation. *Philos. Trans. R. Soc. B Biol. Sci.* 364, 1273–1280.

Nikulin, V. V., Hohlefeld, F. U., Jacobs, A. M., and Curio, G. (2008). Quasi-movements: a novel motor-cognitive phenomenon. *Neuropsychologia* 46, 727–742.

Numminen, J., Salmelin, R., and Hari, R. (1999). Subject's own speech reduces reactivity of the human auditory cortex. *Neurosci. Lett.* 265, 119–122.

O'Craven, K. M., and Kanwisher, N. (2000). Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J. Cogn. Neurosci.* 12, 1013–1023.

Papeo, L., Vallesi, A., Isaja, A., and Rumiati, R. I. (2009). Effects of TMS on different stages of motor and non-motor verb processing in the primary motor cortex. *PLoS ONE* 4:e4508. doi: 10.1371/journal.pone.0004508

Poeppel, D., Idsardi, W. J., and van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 363, 1071.

Price, C. J., Crinion, J. T., and MacSweeney, M. (2011). A generative model of speech production in Broca's and Wernicke's areas. *Front. Psychology* 2:237. doi: 10.3389/fpsyg.2011.00237

Pulvermüller, F., and Fadiga, L. (2010). Active perception: sensorimotor circuits as a cortical basis for language. *Nat. Rev. Neurosci.* 11, 351–360.

Raffin, E., Mattout, J., Reilly, K. T., and Giraux, P. (2012). Disentangling motor execution from motor imagery with the phantom limb. *Brain* 135, 582–595.

Ramachandran, V. S., Rogers-Ramachandran, D., and Cobb, S. (1995). Touching the phantom limb. *Nature* 377, 489.

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

Rizzolatti, G., Fadiga, L., Matelli, M., Bettinardi, V., Paulesu, E., Perani, D., et al. (1996). Localization of grasp representations in humans by PET: 1. Observation versus execution. *Exp. Brain Res.* 111, 246–252.

Rumiati, R. I., Papeo, L., and Corradi-Dell'Acqua, C. (2010). Higher-level motor processes. *Ann. N.Y. Acad. Sci.* 1191, 219–241.

Saygin, A. P., Wilson, S. M., Dronkers, N. F., and Bates, E. (2004). Action comprehension in aphasia: linguistic and non-linguistic deficits and their lesion correlates. *Neuropsychologia* 42, 1788–1804.

Scheerer, E. (1984). "Motor theories of cognitive structure: a historical review," in *Cognition and Motor Processes*, eds W. Prinz and A. Sanders (Berlin: Springer-Verlag), 77–98.

Schroeder, C. E., Lindsley, R. W., Specht, C., Marcovici, A., Smiley, J. F., and Javitt, D. C. (2001). Somatosensory input to auditory association cortex in the macaque monkey. *J. Neurophysiol.* 85, 1322.

Shepard, R. N., and Cooper, L. A. (1986). *Mental Images and Their Transformations*. Cambridge, MA: The MIT Press.

Shergill, S. S., Brammer, M. J., Fukuda, R., Williams, S. C. R., Murray, R. M., and McGuire, P. K. (2003). Engagement of brain areas implicated in processing inner speech in people with auditory hallucinations. *Br. J. Psychiatry* 182, 525–531.

Sirigu, A., Cohen, L., Duhamel, J. R., Pillon, B., Dubois, B., Agid, Y., et al. (1995). Congruent unilateral impairments for real and imagined hand movements. *Neuroreport* 6, 997–1001.

Sirigu, A., and Duhamel, J. (2001). Motor and visual imagery as two complementary but neurally dissociable mental processes. *J. Cogn. Neurosci.* 13, 910–919.

Sirigu, A., Duhamel, J., Cohen, L., Pillon, B., Dubois, B., and Agid, Y. (1996). The mental representation of hand movements after parietal cortex damage. *Science* 273, 1564–1568.

Sommer, I. E. C., Diederen, K. M. J., Blom, J. D., Willems, A., Kushan, L., Slotema, K., et al. (2008). Auditory verbal hallucinations predominantly activate the right inferior frontal area. *Brain* 131, 3169.

Sperry, R. (1950). Neural basis of the spontaneous optokinetic response produced by visual inversion. *J. Comp. Physiol. Psychol.* 43, 482.

Stuart, A., Frazier, C. L., Kalinowski, J., and Vos, P. W. (2008). The effect of frequency altered feedback on stuttering duration and type. *J. Speech Lang. Hear. Res.* 51, 889.

Stuart, A., Kalinowski, J., and Rastatter, M. P. (1997). Effect of monaural and binaural altered auditory feedback on stuttering frequency. *J. Acoust. Soc. Am.* 101, 3806.

Tian, X., and Huber, D. E. (2008). Measures of spatial similarity and response magnitude in MEG and scalp EEG. *Brain Topogr.* 20, 131–141.

Tian, X., and Poeppel, D. (2010). Mental imagery of speech and movement implicates the dynamics of internal forward models. *Front. Psychology* 1:166. doi: 10.3389/fpsyg.2010.00166

Tian, X., Poeppel, D., and Huber, D. E. (2011). TopoToolbox: using sensor topography to calculate psychologically meaningful measures from event-related EEG/MEG. *Comput. Intell. Neurosci.* 2011, 8.

Tkach, D., Reimer, J., and Hatsopoulos, N. G. (2007). Congruent activity during action and action observation in motor cortex. *J. Neurosci.* 27, 13241.

Tourville, J. A., Reilly, K. J., and Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *Neuroimage* 39, 1429–1443.

Tsakiris, M., and Haggard, P. (2005). Experimenting with the acting self. *Cogn. Neuropsychol.* 22, 387–407.

van Wassenhove, V., Grant, K. W., and Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proc. Natl. Acad. Sci. U.S.A.* 102, 1181.

Ventura, M., Nagarajan, S., and Houde, J. (2009). Speech target modulates speaking induced suppression in auditory cortex. *BMC Neurosci.* 10:58. doi: 10.1186/1471-2202-10-58

von Holst, E., and Mittelstaedt, H. (1950). Daz Reafferenzprinzip. Wechselwirkungen zwischen Zentralnerven-system und Peripherie. *Naturwissenschaften* 37, 467–476.

von Holst, E., and Mittelstaedt, H. (1973). *The Reafference Principle (R. Martin, Trans.). The Behavioral Physiology of Animals and Man: The Collected Papers of Erich von Holst*. (Coral Gables, FL: University of Miami Press), 139–173.

von Kriegstein, K., and Giraud, A. L. (2006). Implicit multisensory associations influence voice recognition. *PLoS Biol.* 4:e326. doi: 10.1371/journal.pbio.0040326

Wilson, S. M., Saygin, A. P., Sereno, M. I., and Iacoboni, M. (2004). Listening to speech activates motor areas involved in speech production. *Nat. Neurosci.* 7, 701–702.

Wolpert, D. M., and Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nat. Neurosci.* 3, 1212–1217.

Yoo, S. S., Freeman, D. K., McCarthy, J. J. 3rd., and Jolesz, F. A. (2003). Neural substrates of tactile imagery: a functional MRI study. *Neuroreport* 14, 581.

Zatorre, R. J., and Halpern, A. R. (2005). Mental concerts: musical imagery and auditory cortex. *Neuron* 47, 9–12.

Zatorre, R. J., Halpern, A. R., Perry, D. W., Meyer, E., and Evans, A. C. (1996). Hearing in the mind's ear: a PET investigation of musical imagery and perception. *J. Cogn. Neurosci.* 8, 29–46.

Zhang, M., Weisser, V. D., Stilla, R., Prather, S., and Sathian, K. (2004). Multisensory cortical processing of object shape and its relation to mental imagery. *Cogn. Affect. Behav. Neurosci.* 4, 251–259.

Zheng, Z. Z., Munhall, K. G., and Johnsrude, I. S. (2010). Functional overlap between regions involved in speech perception and in monitoring one's own voice during speech production. *J. Cogn. Neurosci.* 22, 1770–1781.