

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/235884273>

The Effect of Imagination on Stimulation: The Functional Specificity of Efference Copies in Speech Processing

Article in *Journal of Cognitive Neuroscience* · March 2013

DOI: 10.1162/jocn_a_00381 · Source: PubMed

CITATIONS

37

READS

227

2 authors:



Xing Tian

New York University Shanghai

30 PUBLICATIONS 401 CITATIONS

[SEE PROFILE](#)



David Poeppel

New York University

268 PUBLICATIONS 14,759 CITATIONS

[SEE PROFILE](#)

Some of the authors of this publication are also working on these related projects:



Seeing and Hearing a Word: integration within and across senses [View project](#)

The Effect of Imagination on Stimulation: The Functional Specificity of Efference Copies in Speech Processing

Xing Tian and David Poeppel

Abstract

■ The computational role of efference copies is widely appreciated in action and perception research, but their properties for speech processing remain murky. We tested the functional specificity of auditory efference copies using magnetoencephalography recordings in an unconventional pairing: We used a classical cognitive manipulation (mental imagery—to elicit internal simulation and estimation) with a well-established experimental paradigm (one shot repetition—to assess neuronal specificity). Participants performed tasks that differentially impli-

cated internal prediction of sensory consequences (overt speaking, imagined speaking, and imagined hearing) and their modulatory effects on the perception of an auditory (syllable) probe were assessed. Remarkably, the neural responses to overt syllable probes vary systematically, both in terms of directionality (suppression, enhancement) and temporal dynamics (early, late), as a function of the preceding covert mental imagery adaptor. We show, in the context of a dual-pathway model, that internal simulation shapes perception in a context-dependent manner. ■

INTRODUCTION

The alignment of action and perception is one of the foundational questions in neurobiology and psychology. The concept of “internal forward model” has been proposed to link motor and sensory systems, with one central component being that the (neural, computational, cognitive) system can internally predict the perceptual consequences of planned motor commands by internal simulation of “efference copies” (see Wolpert & Ghahramani, 2000, for a review). This concept traces back to Hermann von Helmholtz (1910) and the biologists Von Holst and Mittelstaedt (1950, 1973), and the importance and utility of the idea now extends to visual perception (Sommer & Wurtz, 2006, 2008; Gauthier, Nommay, & Vercher, 1990a, 1990b), motor control (Todorov & Jordan, 2002; Kawato, 1999; Miall & Wolpert, 1996), cognition (Desmurget & Sirigu, 2009; Grush, 2004; Blakemore & Decety, 2001), speech perception (Poeppel, Idsardi, & Van Wassenhove, 2008; Skipper, van Wassenhove, Nusbaum, & Small, 2007; van Wassenhove, Grant, & Poeppel, 2005), and speech production (Hickok, Houde, & Rong, 2011; Guenther, Ghosh, & Tourville, 2006; Guenther, Hampson, & Johnson, 1998; Guenther, 1995).

Although there exist compelling computational arguments and elegant empirical support, the evidentiary basis for efference copies in speech—and their role and specificity—is typically either indirect or hard to disentangle from the effects of overt production. There are tantalizing hints in the case of speech, but the data remain

sparse. For example, when participants overtly perform tasks such as speech production in neuroimaging studies, it is challenging to isolate putative efference copies, in part because of the temporal resolution and in part because of the overt nature of the stimulation. Furthermore, speech occupies a slightly different role than other aspects of motor control; speech is not “just” an action–perception pairing but a set of operations that interface with cognitive systems (language) in a highly specific manner, providing important further constraints (Poeppel et al., 2008). As such, understanding how ideas from systems and computational neuroscience apply in this domain can illuminate how cognitive and perceptuo-motor systems interact.

Recently, Tian and Poeppel (2010) reported direct electrophysiological evidence for auditory efference copies in the human brain. We argued that auditory efference copies have highly similar neural representations to “real” (exogenous, stimulus-induced) auditory activity patterns, on the basis of examining the temporal and spatial characteristics of neural representations underlying the quasi-perceptual experience elicited during mental imagery of speech production. Importantly, imagery tasks are typically argued to be mediated by internal simulation and prediction processes (Tian & Poeppel, 2010, 2012; Grush, 2004; Miall & Wolpert, 1996; Sirigu et al., 1996; Jeannerod, 1994, 1995).

In the speech domain, efference copies have been argued to lower the sensitivity to normal auditory feedback (Behroozmand & Larson, 2011; Ventura, Nagarajan, & Houde, 2009; Heinks-Maldonado, Nagarajan, & Houde, 2006; Eliades & Wang, 2003, 2005; Heinks-Maldonado, Mathalon, Gray, & Ford, 2005; Houde, Nagarajan, Sekihara, &

Merzenich, 2002; Numminen, Salmelin, & Hari, 1999) and increase sensitivity to perturbed feedback (Behroozmand, Liu, & Larson, 2011; Zheng, Munhall, & Johnsrude, 2010; Behroozmand, Karvelis, Liu, & Larson, 2009; Eliades & Wang, 2008; Katahira, Abla, Masuda, & Okanoya, 2008; Tourville, Reilly, & Guenther, 2008). Such modulation presumably occurs when the putative auditory efference copies overlap with perceptual feedback. In contrast to speech-induced suppression of the speech target (the efference copy decreases the response sensitivity to the speech target during production), Hickok et al. (2011) proposed that an efference copy can enhance the response to an auditory target because of task-dependent attentional effects and hence would benefit detection in subsequent perception, such as in audiovisual–speech integration (e.g., van Atteveldt, Formisano, Goebel, & Blomert, 2004; Calvert, Campbell, & Brammer, 2000), and facilitate detection and correction for unexpected feedback errors (e.g., Behroozmand et al., 2009; Eliades & Wang, 2008). However, the mechanisms underlying sensitivity changes induced by the hypothesized auditory efference copies remain unclear, that is, how such representations generated as part of the internal simulation processes modulate the following perceptual processes. One major open issue concerns the specificity of the activated representations.

The repetition paradigm in neuroscience has become a useful tool to probe functional specificity of neural assemblies. The paradigm takes advantage of repetition and adaptation effects in which the previous experience modulates the response properties of neural populations (Henson, 2003; Grill-Spector & Malach, 2001). Repetition experiments have been implemented within one sensory modality, such as audition (Heinemann, Kaiser, & Altmann, 2011; Dehaene Lambert et al., 2006; Bergerbest, Ghahremani, & Gabrieli, 2004; Belin & Zatorre, 2003) and vision (Mahon et al., 2007; Winston, Henson, Fine-Goulden, & Dolan, 2004; Kourtzi & Kanwisher, 2000). Moreover, repetition designs have now also been successfully implemented across modalities to assess the commonality of neural representations, such as in auditory–visual (Doehrmann, Weigelt, Altmann, Kaiser, & Naumer, 2010) and motor–perceptual domains (Chong, Cunnington, Williams, Kanwisher, & Mattingley, 2008). Here we take advantage of the properties of repetition paradigms and their recent cross-modal extensions. We used the “one-shot repetition paradigm” well established in neuroimaging (e.g., Kourtzi & Kanwisher, 2000) and electrophysiological (e.g., Huber, Tian, Curran, O’Reilly, & Worocho, 2008; Bentin & McCarthy, 1994; Rugg, 1985) research that probes feature-specific neural representations (see Grill-Spector, Henson, & Martin, 2006, for a review). We ask whether an internally generated representation (elicited in a mental imagery task) can act as an adaptor for a subsequent overt probe stimulus—or more colloquially, whether thought will prime perception. Given the nature of repetition/adaptation designs, this can only work if the “format” of the internally generated representation is

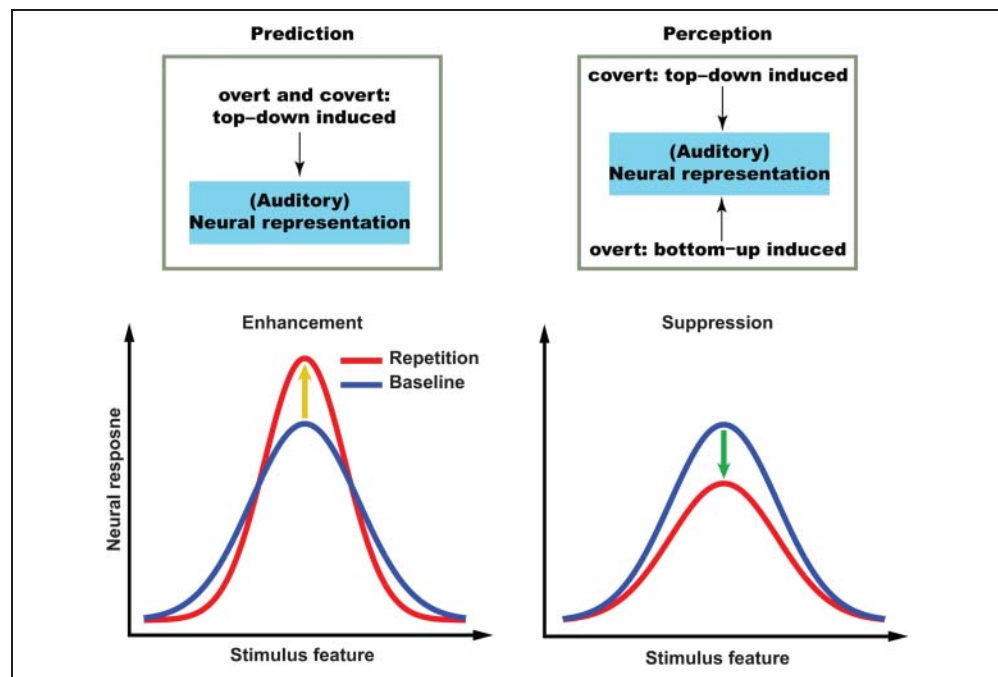
highly similar to or overlapping with the representation generated by an overt stimulus.

Anticipatory/predictive versus perceptual/reactive processes have been suggested to have different consequences for subsequent perception—and produce repetition effects in different directions. The repetition of perceptual processes typically results in repetition suppression. It is hypothesized that the scaling down of perceptual responses results in more efficient processing (see Schacter, Wig, & Stevens, 2007; Grill-Spector et al., 2006, for reviews) and serves as a mechanism to reduce source confusion (Huber, 2008; Huber et al., 2008) and to create relative response increases to signal novelty (Davelaar, Tian, Weidemann, & Huber, 2011; Ulanovsky, Las, & Nelken, 2003; Tiitinen, May, Reinikainen, & Näätänen, 1994). In contrast, active top–down processes can induce feature-specific neural representations in the absence of physical stimuli (Zelano, Mohanty, & Gottfried, 2011; Esterman & Yantis, 2010; Stokes, Thompson, Nobre, & Duncan, 2009), and such feature selection can increase gain for predicted features of an auditory target. This can benefit perception under noisy or challenging conditions in audition (e.g., Elhilali, Xiang, Shamma, & Simon, 2009; van Wassenhove et al., 2005; Grant & Seitz, 2000) and vision (e.g., Peelen, Fei-Fei, & Kastner, 2009; Summerfield et al., 2006; Kastner, Pinsk, De Weerd, Desimone, & Ungerleider, 1999). The relative gain associated with predictive processes would induce repetition enhancement. Figure 1 schematizes such hypothesized outcomes.

In this study, the internal estimation processes in speech were assessed by examining cross-modal modulation effects in a repetition paradigm. We used four different tasks (constituting the factor of Adaptor Type) to induce internal estimation (and therefore auditory efference copies): (i) overt and (ii) covert speech production, (iii) overt auditory perception, and (iv) auditory imagery of speech (covert perception). In the articulation task (*A*), participants were asked to overtly generate a cued syllable. In the articulation imagery task (*AI*), participants were required to imagine saying a syllable without moving the mouth. In the hearing imagery task (*HI*), participants were asked to imagine hearing a cued syllable. In the hearing task (*H*), the adaptor was an overt syllable and the task was passive listening. The factors of Repetition Status (repeated vs. novel) and Adaptor Type were fully crossed, creating eight conditions. Figure 2 schematically summarizes the design.

In a previous article (Tian & Poeppel, 2010), we argued based on the magnetoencephalography (MEG) data that the realization of an auditory efference copy during speech production cannot take longer than ~150–170 msec (which is arguably an overestimate). Here we capitalize on this fact; that is to say, we assume that the rapid generation of such an efference copy underlies the activation of the neuronal population that then is “probed” with the subsequent auditory stimulus (which in the present design occurs within a few hundred milliseconds).

Figure 1. Predictions of repetition effects in different tasks. The auditory efference copy in the internal simulation/prediction process (in overt and covert articulatory tasks, internally top-down induced) is hypothesized to increase the response sensitivity to the feature of repeated stimuli, resulting in neural response enhancement (left, indicated by yellow arrow). The neural representation in perception (in overt perceptual task, externally bottom-up induced and in covert perceptual task, internally top-down induced, constrained by contextual/task demand) is hypothesized to decrease the response sensitivity to the feature of repeated stimuli, resulting in neural response suppression (right, indicated by a green arrow).



Two hypotheses were investigated about how neural activity (with the previous experience formed in distinct ways in the *A*, *AI*, *H*, and *HI* conditions) modulates subsequent responses to the auditory probe. First, we tested whether the activity patterns associated (i) with auditory efference copies (*A* and *AI* conditions) and (ii) auditory mental images (*HI* condition) are the same as activity elicited by overt auditory perception, which would be indicated by the existence of cross-modality repetition effects. If internal simulation processes (and the putative efference copies) are encoded in a largely similar representational format, then cross-modal (covert thought to overt stimulation) repetition should be observed. Second, insofar as similar neural representations are engaged, we investigated the modulatory functions of predictive versus perceptual neural responses on speech perception. We conjectured that the directionality of information flow (top-down in the efference copies [*A*, *AI*] and auditory imagery [*HI*] vs. bottom-up in the overt perception [*H*]) combined with the context (goal of the task: articulation or perception) adaptively shape the generation and subsequent computational role of the auditory neural representation elicited by the adaptors. That is, the effects of repetition are determined by the preceding adaptors. As schematized in Figure 1, we predicted that a “perceptual” task (whether overt as in *H* or covert as in *HI*) will cause a form of local perceptual learning and hence lead to repetition suppression. In contrast, the auditory efference copy in speech production (whether overt *A* or covert *AI*) actively predicts the upcoming auditory consequences; the available efference copy increases the response gain of the following perceptual process, resulting in repetition enhancement. Although the experi-

ment and model we pursue is implemented in the context of speech processing, the nature of the underlying operations is arguably “generic” in the sense that the account we are pursuing generalizes to other instances of repetition suppression and enhancement: The active nature of prediction leads to response gain increases.

We derived specific electrophysiological timing predictions on the basis of literature in linguistics, psycholinguistics, and neurolinguistics. An abstract phonological representation has been hypothesized to underpin the seamless transition between articulatory and acoustic tasks (for discussion of some of these issues, see, e.g., Poeppel et al., 2008; Hickok & Poeppel, 2007). The abstract phonological code is presumably invariant and independent from acoustic features (e.g., Phillips et al., 2000). Human electrophysiological studies on speech suggest that the early auditory components (approximate latencies ~30–100 msec, presumably reflecting the activity in core and belt auditory cortices) underlie the analysis of acoustic-phonetic features, whereas the later auditory components (approximate latencies ~150–250 msec, presumably reflecting activity in associative auditory regions) reflect the abstract phonological representation (e.g., Phillips et al., 2000). We hypothesize that the top-down induced process runs in the opposite direction of the bottom-up process. That is, in the context of a top-down task, the abstract phonological code would be estimated first, followed by the estimation of concrete acoustic features. (Or at least independent from each other and formed in separate processes.) The level of top-down induced auditory neural representations seems to depend on the task demands, where, for example, high demand on recreating concrete acoustic features drives the extension of activation from associative to

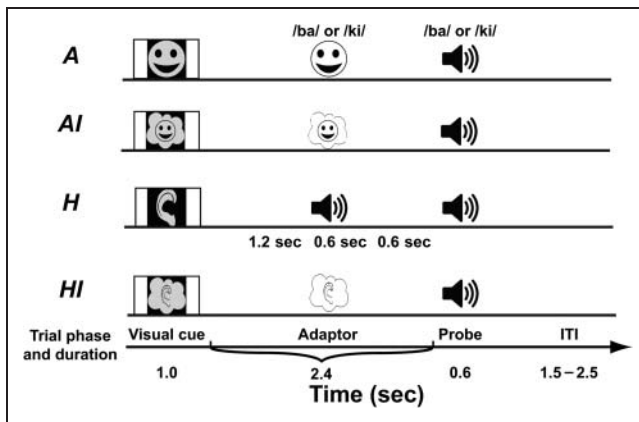


Figure 2. Experimental procedure. The three phases in each trial is presented at the bottom. At the beginning of each trial, a visual cue appears at the center of screen and stays on for 1 sec. Different pictures are used as visual cues to indicate different tasks and a written label, either /ba/ or /ki/, is superimposed on the center of visual cue to inform the content of the task. A 2.4-sec adaptor phase starts at the offset of a visual cue. During the adaptor phase, participants are required to finish different tasks to create an adaptor. Notice that the 2.4-sec adaptor phase is the total duration that participants are allowed to finish the tasks (indicated by the curly bracket). The actual time of forming an adaptor is presumably much shorter. In the articulation task (*A*), participants are asked to overtly pronounce the indicated syllable, whereas in the articulation imagery task (*AI*), participants are required to covertly pronounce the syllable. In the hearing task (*H*), 1.2 sec after the onset of visual cue, participants passively listen to a 0.6-sec syllable sound, followed by a 0.6-sec silent interval. In the hearing imagery task (*HI*), participants are asked to imagine hearing the syllable sound. A 0.6-sec probe sound always follows the adaptor phase and the probe sound is either same (*repeated*) as or different (*novel*) from the content of preceding adaptor. Participants were required to passively listen to the probe sound. Intertrial intervals are randomized between 1.5 and 2.5 sec with 0.25-sec increment.

primary auditory cortices (e.g., Kraemer, Macrae, Green, & Kelley, 2005; Halpern & Zatorre, 1999; for the discussion about level of top-down induced neural representation depending on task demand, see, e.g., Zatorre & Halpern, 2005; Kosslyn, Ganis, & Thompson, 2001). Therefore, together with our manipulation of phonology in this study (same or different syllables, see Methods section for details), we predict that top-down induced estimation would most likely to be strong at the phonological level and hence interact with the subsequent auditory processing in the later components (presumably M200). This contrasts with the bottom-up repetitions that affect both early (presumably M100) and later components.

METHODS

Participants

Fourteen volunteers participated in the experiment. The data from two were excluded: In one participant, the data had extensive artifacts and a high noise level during the recording and another participant failed to perform the tasks. The data from 12 participants (six men, mean

age = 29.1 years, range = 22–43 years) were included in the final analysis. All participants were right-handed and with no history of neurological disorders. The experimental protocol was approved by the New York University Institutional Review Board.

Stimulus Materials

Two 600-msec duration consonant–vowel syllables (/ba/, /ki/) were used as auditory stimuli (female voice; sampling rate of 48 kHz). All sounds were normalized to 70 dB SPL and delivered through plastic air tubes connected to foam ear pieces (E-A-R Tone Gold 3A Insert earphones, Aearo Technologies Auditory Systems). Four images were used as visual cues to indicate four different trial types. Each image was presented foveally, against a black background, and subtended less than 10° visual angle. A label—either “/ba/” or “/ki/”—was superimposed on the center of each picture (<4° visual angle) to indicate the syllable that participants would produce in the following tasks.

The choice of using only female vocalization was motivated by the assumption that we tap the abstract level of phonological representation (Poehpel et al., 2008; Hickok & Poehpel, 2007). The abstract phonological code is presumably invariant and independent from acoustic features (e.g., Phillips et al., 2000). That is, the attributes of this representation are shared across tokens or specific instances of a speech sound (male, female, fast, slow, whispered, etc.). Because the goal of this study is to investigate the neural representation and functional specificity of efference copies at the abstract phonological level and because of the hypothesized invariance of the phonological code, we simplified our experimental design by only presenting a female voice, on the view that it will activate abstract codes for both female and male speakers. Using each individual participant’s vocalizations (an often-used approach that we, too, employ in a related study; Tian & Poehpel, under review) would lead to too long a study in this design, in which tapping to low-level phonetic representations is not as critical as activating the abstract ones.

Procedure

The structure of the trials is schematized in Figure 2. The experiment comprised two classes of trials, overt and covert: articulation (*A*), hearing (*H*), articulation imagery (*AI*), and hearing imagery (*HI*). The timing of trials was consistent across trial types, and trials had three phases. First, a visual cue appeared in the center of the screen at the beginning of each trial and stayed on for 1000 msec. During the following 2400 msec (adaptor phase), participants actively formed a syllable (adaptor) in three of the task conditions (overt and covert production, *A* and *AI*, and covert perception, *HI*; see below for details) or passively perceived an auditory syllable in the overt hearing (*H*) condition, in which a syllable was presented

1200 msec after the offset of visual cue, followed by a 600-msec interval. Notice that the 2.4-sec adaptor phase was the total duration that participants were allowed to finish the tasks (indicated by the curly bracket). The actual time of forming an adaptor was presumably much shorter. Finally, participants were presented the syllable probe sound that always followed the adaptor phase. In summary, the syllable probe stimulus was preceded by one of four different adaptor types. The intertrial interval was 1500–2500 msec (with 250-msec increments). The experiment was run in six blocks with 64 trials in each block.

Two factors were investigated in the experiment, in a 2×4 design. The first factor concerns the relation between the adaptor and probe, Repetition Status, with two levels: The probe syllable /ba/ or /ki/ was either congruent (repeated) or incongruent (novel) with the content of the adaptor. Each syllable was presented equally often as adaptor and probe. The second factor captures the task of forming the adaptor, Adaptor Type, with four levels. In the articulation task (*A*), participants were asked to overtly generate the cued syllable (gently to minimize head movement). In the articulation imagery task (*AI*), participants were required to imagine saying a syllable without any overt movement of the articulators. In the hearing task (*H*), the adaptor was the syllable and the task was passive listening. In the hearing imagery task (*HI*), participants were asked to imagine hearing the cued syllable. The factors Repetition Status and Adaptor Type were fully crossed, creating eight conditions (e.g., *A repeated*, *A novel*, *AI repeated*, *AI novel*, *H repeated*, *H novel*, *HI repeated*, *HI novel*). Eight trials for each condition were included in each of the six recording blocks (pseudorandom presentation order), yielding 48 trials in total for each condition.

Each participant received training for 15–20 min before the MEG experiment with focus on the timing as well as vividness of imagery. First, only the *H* trials were presented to introduce the relative timing among the visual cue, the auditory adaptor, and the following probe. After participants were familiar with the timing, they were instructed to use the same timing for the other trial types. Next, they practiced on *A* trials while the experimenter observed the overt articulation and provided feedback if needed. It was confirmed that they could execute the task with consistent timing before they moved to the next practice. Subsequently, participants were trained on the imagery conditions. For the *AI* condition, they were told to imagine speaking the syllables “in their mind” without moving any articulators or producing any sounds. They should feel the movement of specific articulators that would associate with actual pronunciation and “hear” their own voice “loud and clear” in their mind. For the *HI* condition, they were asked to retrieve the sounds in the female voice they just heard in the *H* condition. Requirements were to recreate the female voice “loud and clear” in their minds but not to generate any

feeling of movement for any articulators. If needed, the recorded female voice was presented again to form a better memory. The vividness (“loud and clear” as well as the voice distinction) was emphasized, and participants practiced. Participants were asked to generate a movement intention and kinesthetic feeling of articulation in the *AI* condition; in the *HI* condition, such motor-related imagery activity was strongly discouraged. We tried to selectively elicit the motor-induced auditory representation in imagined speaking, while we aimed to target auditory retrieval in imagined hearing. After verbal confirmation of successful distinction of two types of imagery formation process as well as vividly generating the “loud and clear” representations, they further practiced on the *AI* and *HI* tasks to reinforce the vividness of imagery as well as combine the timing requirement in the trials. Lastly, they trained on a practice block in which all four conditions were available. Timing of the *A* condition was monitored by the experimenter and verbal confirmation of vividness during imagery was obtained for each participant before moving to the main experiment.

We monitored whether participants made any overt pronunciation or not throughout the experiment (by microphone adjacent to participants). The observations of overlapping neural networks between covert and overt movement in motor imagery studies (e.g., Dechent, Merboldt, & Frahm, 2004; Meister et al., 2004; Ehrsson, Geyer, & Naito, 2003; Hanakawa et al., 2003; Gerardin et al., 2000; Lotze et al., 1999; Deiber et al., 1998) support that both types of articulator movement would induce a similar motor efference copy, as suggested in several theoretical articles (e.g., Desmurget & Sirigu, 2009; Grush, 2004; Miall & Wolpert, 1996; Jeannerod, 1994, 1995). As long as there is no overt sound, our goal of an internally induced auditory representation from a motor efference copy is valid. Potential subvocal movement is irrelevant to the interpretation.

MEG Recording

Neuromagnetic signals were measured using a 157-channel whole-head axial gradiometer system (KIT, Kanazawa, Japan). Five electromagnetic coils were attached to a participant’s head to monitor head position during MEG recording. The locations of the coils were determined with respect to three anatomical landmarks (nasion, left and right preauricular points) on the scalp using 3-D digitizer software (Source Signal Imaging, Inc., San Diego, CA) and digitizing hardware (Polhemus, Inc., Colchester, VT). The coils were localized to the MEG sensors at both the beginning and the end of the experiment. The MEG data were acquired with a sampling rate of 1000 Hz, filtered on-line between 1 and 200 Hz (2-pole Butterworth low-pass filter, 1-pole high-pass filter), with a notch filter of 60 Hz (1-pole pair band elimination filter).

MEG Analysis

The raw data were noise-reduced off-line using the time-shifted PCA method (de Cheveigné & Simon, 2007). Trials with amplitudes > 2 pT ($\sim 5\%$) were considered artifacts and discarded. In the *H* task, 600-msec epochs of response to the first sound (adaptor), including a 100-msec prestimulus period, were extracted and averaged across *repeated* and *novel* trials. Similarly, 600 msec epochs of response to probes in all eight conditions were extracted and averaged. All averages were baseline-corrected using a 100-msec prestimulus period. The averages were low-pass filtered with a cutoff frequency of 30 Hz (finite impulse response filter, Hamming window with size of 100 points). A typical M100/M200 auditory response complex was observed (Roberts, Ferrari, Stufflebeam, & Poeppel, 2000), and the peak latencies were identified for each individual participant, detailed further below.

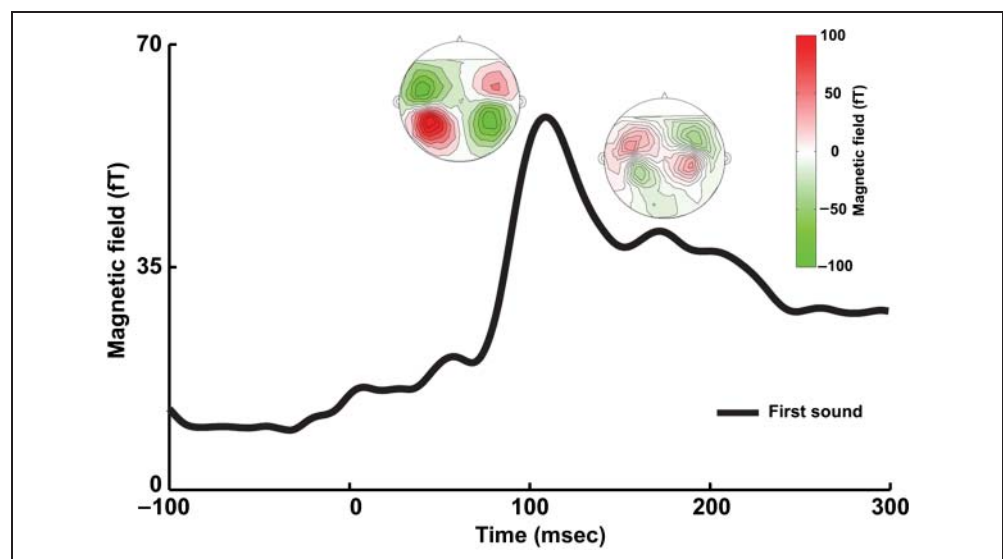
The critical data for this experiment are the neuro-magnetic response amplitudes elicited by the probe syllables and the manner in which these responses are modulated by the preceding adaptor (cf. Figure 1). Because we are testing an electrophysiological hypothesis and aim to stay close to the recorded data, one goal is to analyze sensor-level recordings; however, this provides additional challenges. Because of possible confounds between neural source magnitude change and distribution change in analyses at the sensor level (Tian & Huber, 2008), a multivariate measurement technique (“*angle test* of response similarity”), developed by Tian and Huber (2008) and available as an open-source toolbox (Tian, Poeppel, & Huber, 2011), was implemented to assess the topographic similarity between responses to the *repeated* and *novel* probes. This technique allows the assessment of spatial similarity in electrophysiological

studies regardless of response magnitude and estimation of the similarity of underlying neural source distributions (Davelaar et al., 2011; Tian & Poeppel, 2010; Huber et al., 2008).

After confirming the stability of neural source distributions of auditory perceptual responses, any observed significant changes obtained in the sensor level analyses will be attributed to varying response magnitude as a function of trial type. The root mean square (RMS) of waveforms across 157 channels, indicating the global response power in each condition, was calculated and employed in the following statistical tests. A 25-msec time window centered at individual M100 and M200 latency peaks was applied to obtain the temporal average responses, separately for each condition as well as the first sound in *H*. To aggregate the temporally averaged data across participants, the percent change of response magnitude was calculated. Specifically, the response to the first sound in *H* (reference responses) was subtracted from the responses to the probes and the differences were further divided by the reference responses to convert the absolute differences into percent change.

Distributed source localization of the repetition effects was obtained by using the Minimum Norm Estimation (MNE) software (Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Boston, MA). L2 minimum norm current estimates were constrained on the cortical surface that was reconstructed from individual structural MRI data with Freesurfer software (Martinos Center for Biomedical Imaging, Massachusetts General Hospital, Boston, MA). Current sources were about 5 mm apart on the cortical surface, yielding approximately 2500 locations per hemisphere. Because MEG is sensitive to electromagnetic fields generated from current sources that are in sulci and tangential to the cortical surface (Hämäläinen, Hari, Ilmoniemi, Knuutila, & Lounasmaa, 1993), deeper

Figure 3. Grand average of RMS waveform and M100/M200 topographies of adaptor in *H*. Typical M100 and M200 peaks in the RMS waveform are observed. The topographies of M100 and M200 responses are depicted above each peak. The color in each pair of activity represents the direction of magnetic flux, where red represents the direction of coming out of scalp and green represents going into scalp. The polarities in M100 and M200 responses are reversed indicating the opposite orientation of neural sources.



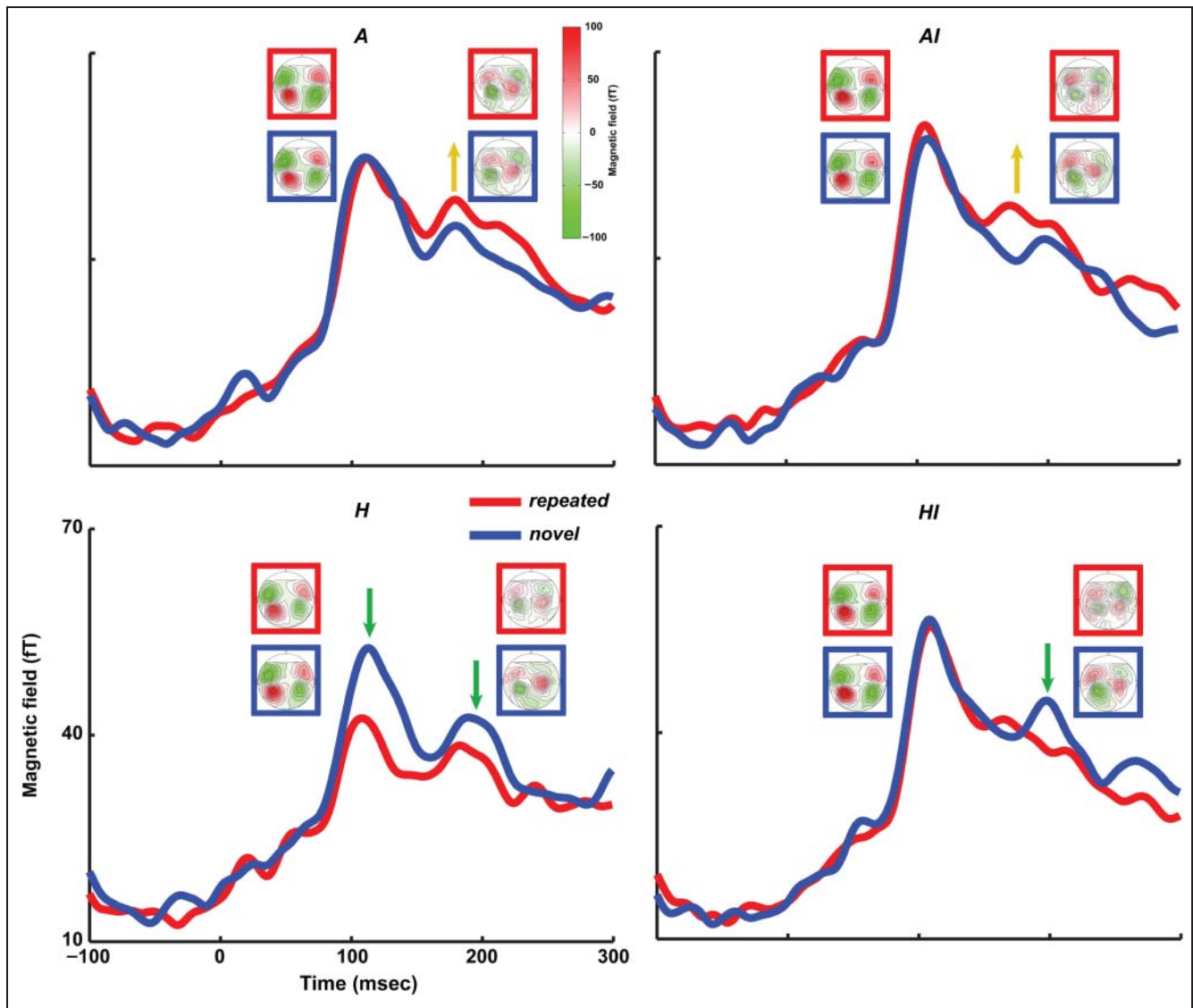


Figure 4. Grand average of RMS waveforms and M100/M200 topographies of all conditions. The red and blue lines depict the RMS waveforms to the *repeated* and *novel* probes in all tasks, respectively. The topographies next to the peaks at 100 and 200 msec present the response patterns of M100 (first column) and M200 (second column) components. The same color scheme of surrounding squares is used to indicate probe types (red for *repeated* and blue for *novel*). Similar pattern distributions are observed within M100 or within M200 responses [comparing first row (*repeated*) with second row (*novel*) in each column]. The yellow and green arrows in each subplot indicate the activity increases and decreases by comparing the responses to the *repeated* probes with the responses to the *novel* probes.

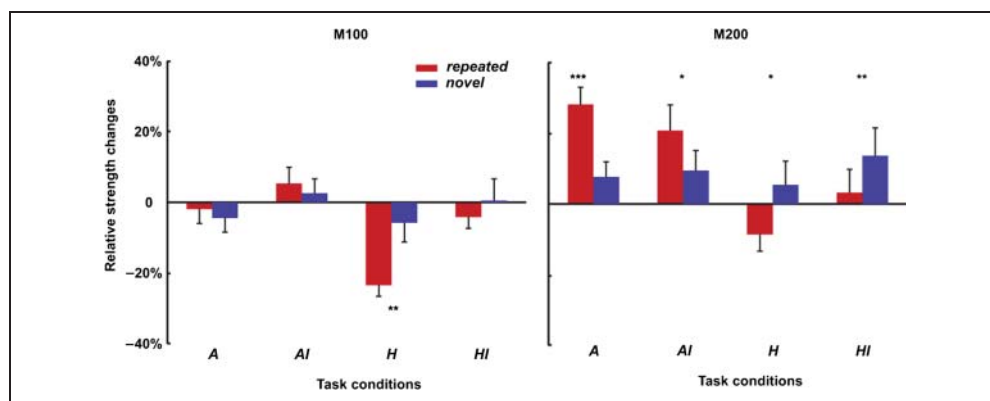
sources are given more weight to overcome the MNE bias toward superficial currents and current estimation favors the sources normal to the local cortical surface (Lin, Belliveau, Dale, & Hämäläinen, 2006). Individual single-compartment boundary element models were used to compute the forward solution. On the basis of the forward solution, the inverse solution was calculated by approximating the current source spatio-temporal distribution that best explains the variance in observed MEG data. Current estimates were normalized by the estimated noise power from the entire epoch to convert into a dynamic parametric map (Dale et al., 2000). To compute and visualize the MNE group results, each participant's cortical surface was inflated and flattened (Fischl, Sereno, & Dale, 1999) and

morphed to a representative surface (Fischl, Sereno, Tootell, & Dale, 1999). Current estimation was first performed within each condition, and the repetition effects in each task were obtained by subtracting the absolute values of estimation of *novel* epoch from the one of *repeated* epoch and then averaged across participants. The same M100 and M200 time windows as used in event-related analysis were then applied.

RESULTS

The canonical response profile to auditory syllables was confirmed—both in terms of temporal profile and

Figure 5. Percent change in temporally averaged auditory M100 and M200 responses of all conditions. The red and blue bars represent the percent change of auditory responses to *repeated* and *novel* probes with respect to the reference responses (auditory responses to the first sound in *H*). Repetition suppression is shown only in *H* in M100 (left plot), demonstrated by the larger decrease for the responses to *repeated* than to *novel* auditory probes. In M200 (right plot), repetition enhancement is found in *A* and *AI*, contrasting with repetition suppression in *H* and *HI*. One, two, three stars respectively indicate $p < .05$, $p < .01$, and $p < .005$.



topography—for the first sound during the overt auditory perception *H* task (reference responses). Figure 3 depicts the grand-averaged (RMS) waveform across channels and participants to that stimulus (and the magnetic field contour maps associated with the respective peaks). Typical M100 and M200 response peaks were observed, with the orientation of the contour map flipped between the response pattern occurring around 100 and 200 msec after stimulus onset, reflecting the underlying source differences. Because no auditory stimuli preceded the adaptor in the *H* trials, the M100 and M200 responses to the adaptor were used as baseline level responses to quantify the relative changes in *repeated* and *novel* probes. The auditory response patterns were also observed for the auditory probe syllables in all eight conditions. Importantly, the *angle test* did not reveal any significant spatial pattern differences (i) between responses to *repeated* and *novel* probes, (ii) between reference responses and responses to *repeated* probes, and (iii) between reference responses and responses to *novel* probes in all conditions. That is, the topographies of auditory responses to probes in all conditions and reference responses were highly similar.

Repetition/Adaptation Response Pattern

Figure 4 shows the RMS waveform responses to the auditory probes in all conditions, separated by tasks. Only in the *H* condition (bottom left) did the amplitude difference between the responses to *repeated* and *novel* probes occur around 100 msec, such that the *novel* probe had a higher amplitude response than the *repeated* one. In the overt *A* and covert *AI* conditions (top row), the M200 responses to the *repeated* probes were larger than the ones to the *novel* probes. In con-

trast, in the overt *H* and covert *HI* conditions, the M200 responses to the *repeated* probes had lower amplitudes compared with the *novel* probes.

The repetition effects observed in the waveform morphologies were further quantified. The percent change of M100 responses (see Procedure section for details) are presented in Figure 5 (left). A repeated-measures two-way ANOVA was carried out on the factors Repetition Status and Adaptor Type. The main effect of Adaptor Type was significant [$F(3, 33) = 8.03, p < .001$] and the interaction was also significant [$F(3, 33) = 2.95, p < .05$]. The planned paired *t* test performed on the significant interaction revealed that, only in condition *H*, responses to *repeated* probes were significantly smaller than the ones to *novel* probes [$t(11) = -3.12, p < .01$]. However, no difference was found between responses to the *repeated* and *novel* probes in *A* or *AI* or *HI* [all t s < 1].

Similar analyses were carried out for the M200 responses. As seen in Figure 5 (right), a differential pattern was observed. A repeated-measures two-way ANOVA shows that the main effect of Adaptor Type was significant [$F(3, 33) = 4.63, p < .01$]. The interaction was also significant [$F(3, 33) = 12.81, p < .001$]. The planned paired *t* test performed on the significant interaction revealed that repetition suppression occurred in conditions *H* [$t(11) = -2.32, p < .05$] and *HI* [$t(11) = -3.30, p < .01$]. In contrast, the repetition effects were associated with robust enhancement in the *A* [$t(11) = 4.12, p < .005$] and *AI* conditions [$t(11) = 2.27, p < .05$].

Analyses by Hemispheres

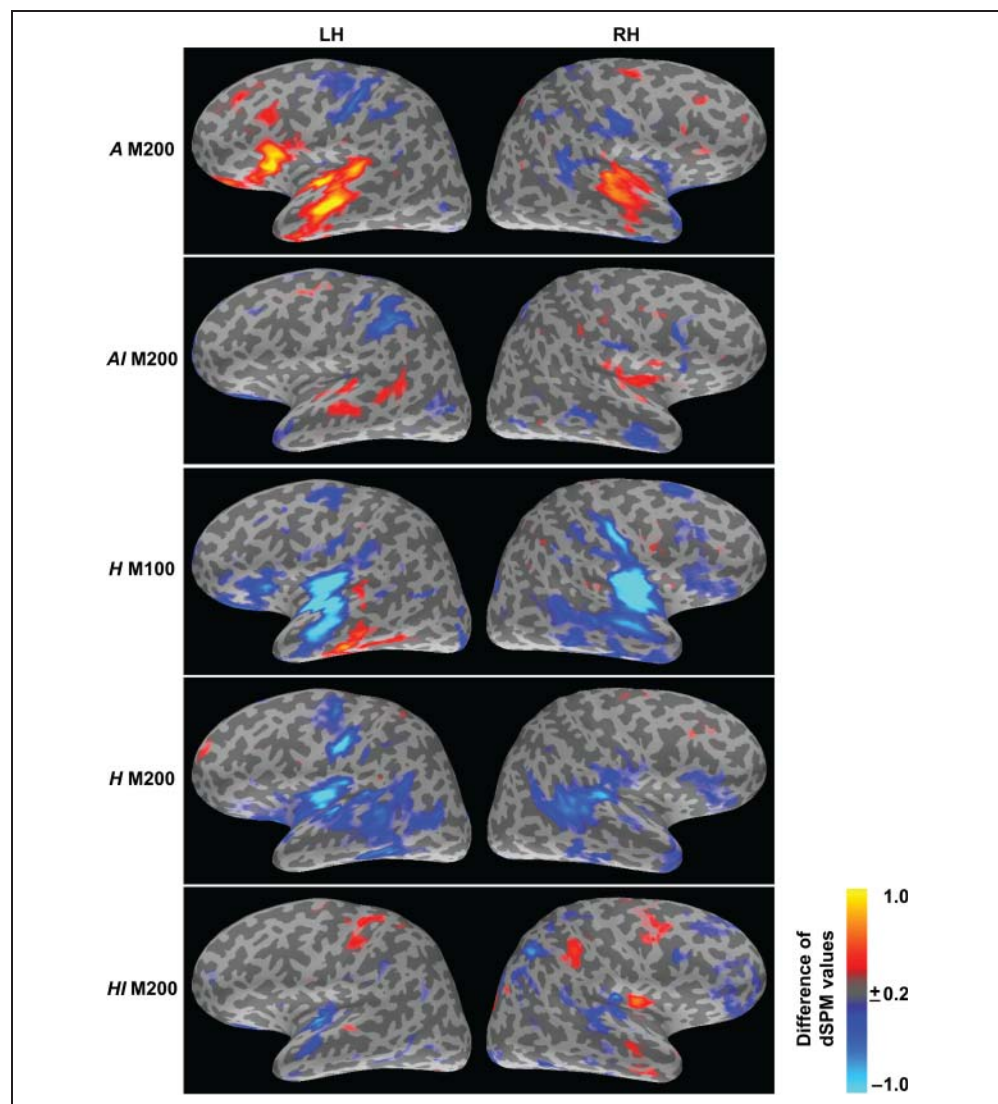
The above analyses were calculated across all channels. To verify that the effects hold within each hemisphere, we

repeated the analyses with restricted sets of channels. This way, potential hemispheric lateralization of repetition effects was further investigated. The same analyses as above were applied to the time averages obtained in the channels over the left and right hemispheres separately. Repetition suppression was obtained in both hemispheres in condition *H* M100 responses (left [$t(11) = -2.52, p < .05$], right [$t(11) = -3.89, p < .005$]), and repetition enhancement was observed in the *A* M200 responses (left [$t(11) = 2.74, p < .05$], right [$t(11) = 2.70, p < .05$]). Leftward lateralization occurred in the M200 responses in *H* and *HI*, with the significant repetition effects only observed in the temporal averages of left hemisphere sensors (*H* [$t(11) = -2.23, p < .05$] and *HI* [$t(11) = -2.40, p < .05$]). Marginal leftward lateralization was observed in *AI* M200 responses, with significant repetition enhancement in the left hemisphere ([$t(11) = 2.29, p < .05$]) and marginally significant enhancement in the right ([$t(11) = 2.00, p = .07$]).

Source Localization

The neuronal sources of all the observed repetition/adaptation effects at the sensor level were further investigated in MNE (Hämäläinen & Ilmoniemi, 1994; see Procedure section). The cortical activity differences between *repeated* and *novel* probes were averaged across participants and overlaid on a morphed anatomical template (Figure 6). The repetition-induced enhancement was observed over bilateral superior temporal gyrus (STG) and anterior STS in the M200 response of condition *A* (top row). In addition to the auditory cortices, inferior frontal gyrus (IFG), and adjacent premotor cortex also show enhancement, consistent with the hypothesis that an articulation efference copy is generated over these frontal regions (Hickok, 2012; Tian & Poeppel, 2010, 2012; Guenther et al., 2006). A similar enhancement in STG and middle/posterior STS, although more modest in amplitude, was also seen in the M200 responses of *AI* (Figure 6, second row). Repetition suppression was observed

Figure 6. Source localization of repetition effects using MNE. Grand average of difference dSPM activity in all tasks. The difference dSPM values were calculated by subtracting the source responses to the *novel* probes from responses to the *repeated* probes at each source location, superimposed on an inflated and flattened representative cortical surface. The dark and light gray areas represent sulci and gyri, respectively. Notice that because the depicted activity is the difference between the responses to the *repeated* and *novel* probes, the warm color represents the repetition enhancement and the cool color represents the repetition suppression. Refer to the main text for details about the locations of repetition effects.



in the M100 and M200 responses of *H* (third and fourth rows). For the M100, bilateral decreases in the Sylvian fissure, anterior STG and anterior STS were observed; for the M200, bilateral decreases in posterior STS were observed but strong deactivation only presented in left Sylvian fissure and STG. Repetition suppression was also observed in M200 responses in the *HI* condition, with decreased activity in the left anterior part of Sylvian fissure and STS but more posterior in the right hemisphere. These source analyses of the neuromagnetic response to the probe syllables—always the same overt auditory stimulus—underscore the striking extent to which the response direction and spatial pattern are modulated by the adaptor preceding the auditory signal as well as the change over time, indicating that a single metric for “repetition” does not adequately capture the processing elicited by the adaptors.

DISCUSSION

We combined the well-established stimulus repetition (adaptation) design with a classical experimental approach from cognitive psychology, mental imagery. This pairing of techniques is unusual and has not been employed. Insofar as one obtains a repetition effect on a probe stimulus—either systematic suppression or enhancement—one can conclude that the representations underlying the effect are related in some principled way. Internally generated “thought” in mental imagery could then be argued to prime overt stimulation because of the high degree of similarity between the representations. We adapted this unconventional approach to test the question whether efference copies in speech, a concept foundational to numerous current models of production, perception, and their link, display functional specificity (e.g., in the context of internal prediction) or are relatively generic (i.e., any auditory representation will do). Recent work on speech production has yielded strong claims about the existence and role of efference copies (Hickok et al., 2011; Price, Crinion, & MacSweeney, 2011; Tian & Poeppel, 2010; Guenther et al., 2006), but the extent to which such representations are functionally specific has not been approached. In this study, although the representations generated by different adaptors are very similar (hence “generic”), the difference in the modulation effect shows that they are specific in their function and therefore not “generic” but rather “specific” in virtue of being generated and dependent on the task.

Our novel mental imagery paradigm, complementary to the immediate feedback paradigm (e.g., Ford, Roach, & Mathalon, 2010; Houde et al., 2002; Burnett, Freedland, Larson, & Hain, 1998), provided direct neural evidence about internal forward models. The literature suggests that the tasks we employed, mental imagery, are comprising internal simulation and estimation processes that make use of the mechanisms of efference copies (e.g., Tian & Poeppel, 2010, 2012; Davidson & Wolpert, 2005;

Grush, 2004; Miall & Wolpert, 1996; Sirigu et al., 1996; Jeannerod, 1994, 1995)—but without overt muscle and acoustic signals. The absence of movement and auditory input is a feature that overcomes the problems associated with the overlap between the neural processes elicited by external stimuli and internal operations, both in temporal (occurred at same time) and spatial (internal and external induced similar auditory representation) aspects. We therefore first confirmed that the internal simulation/prediction elicited by mental imagery can interact with overt stimulation. The observed “cross-modal” repetition effects (from imagination to stimulation) support the hypothesis of overlapping auditory neural representations between efference copies in production and auditory processing in perception (e.g., Tian & Poeppel, 2010; Ventura et al., 2009; Eliades & Wang, 2003, 2005; Houde et al., 2002; Numminen et al., 1999) and between covert and overt perception (Bunzeck, Wuestenberg, Lutz, Heinze, & Jancke, 2005; Schürmann, Raij, Fujiki, & Hari, 2002; Wheeler, Petersen, & Buckner, 2000; see Hubbard, 2010; Zatorre & Halpern, 2005, for reviews).

Critically, we further uncovered two clear directional and temporal patterns in the data. First, repetition suppression effects are observed in the two hearing conditions, as predicted, whereas the two articulation conditions show repetition enhancement. Second, we show that there is a temporal dynamic underlying the process; only the bottom-up overt perceptual task (*H*) had an early effect at ~100 msec, but all top-down task types (*A*, *AI* and *HI*) principally affected neural responses at ~200 msec.

Because we only used a female voice as a probe stimulus, matching between speaker identities could be an alternative hypothesis to explain the observed repetition effects. In fact, the main effect of enhanced responses to novel probes in all active conditions (Figure 5) could be the effect of mismatching speaker identity. However, crucially, the double dissociation between the *AI* and *HI* conditions in terms of the direction of modulation effects suggests that the mismatch mechanism alone cannot fully explain these findings.

The emphasis on trial timing was to make sure that no temporal overlap would occur between the internally induced representation and the subsequent auditory stimuli. The requirement of consistent timing requires timing judgments, which is not demanding, as demonstrated by the quick learning and consistency in practice. Moreover, *AI* and *HI* arguably require more time to finish, as the mental imagery tasks are more demanding. The timing judgment and task completion time differences could be advanced as possible alternative explanations for the main effects in the *A*, *AI*, and *HI* conditions; but again, it is very hard indeed to explain the double dissociation observed in the modulation effects.

During this experiment, in the *AI* condition participants imagined speaking using their own voice, whereas in *HI* participants imagined hearing using the recorded female voice. The mismatch between the acoustic features could

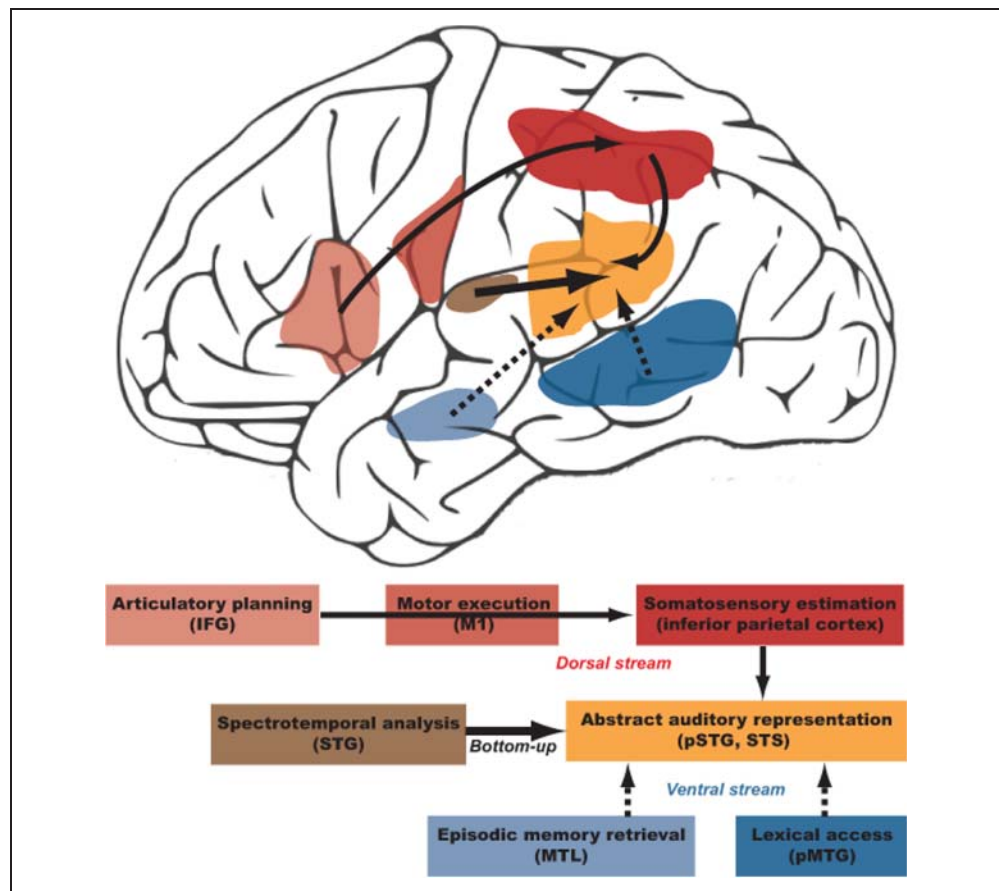
explain the repetition effects. However, as the phonological code is invariant across acoustic features, we believe that we manipulated the phonological level to assess the neural representation and functional specificity of efference copy. In fact, our results provide support for this. The *AI* condition has more mismatch than *HI* (imagining one's own voice in *AI* and imagining the female voice in *HI*, then listening to the female voice). However, there was no main effect between *AI* and *HI*, suggesting the comparison of acoustic features is not the key factor for observed the M200 effects.

Previous studies suggest that the act of articulation affects the perception of subsequent auditory feedback by around 100 msec (Ventura et al., 2009; Houde et al., 2002). The apparent difference between those findings and ours could be caused by two factors. First, in previous work, the manipulations were made on acoustic properties (such as pitch; Behroozmand et al., 2011; Eliades & Wang, 2008); in the present experiment, phonological features were varied, tapping into a different process in the hierarchy of speech perception. In fact, in another study using mental imagery in which pitch was manipulated and the internal simulation was over-

lapped with external feedback, we replicated the common finding of sensitivity around 100 msec (Tian & Poeppel, under review). Second, the duration between the internal simulation and feedback could be an additional factor. Compared with the immediate feedback used in previous studies (Behroozmand et al., 2011; Ventura et al., 2009; Eliades & Wang, 2008), a delay was introduced between the internal simulation and external auditory stimuli. Cognitive control processes may gradually become involved and privilege slightly later, higher-order processes such as feature selection.

The specificity of the timing and directional effects suggests that the neural computations reflected in these response patterns can change flexibly and rapidly depending on task demands and current state. In an effort to pull together different strands of evidence, we outline below a functional anatomic perspective (Figure 7). In particular, we link these findings to the potential differential contribution of the dorsal and ventral speech processing streams (Rauschecker & Scott, 2009; Hickok & Poeppel, 2007) and their putative computational roles: The ventral stream maps acoustic signals to “meaning” (broadly construed) in speech comprehension, and the dorsal stream

Figure 7. Proposed dual stream prediction model (DSPM). Top: approximate anatomical locations of implicated cortical regions in the hypothesized streams. Bottom: schematic functional diagram of the DSPM (color scheme corresponds to the anatomical locations above). The abstract auditory representations (orange) are formed around regions of pSTG and STS. The ventral stream (blue) includes pMTG and MTL, critical for retrieval from long-term lexical memory and episodic memory. The dorsal stream (red) includes IFG and inferior parietal cortex. The articulatory trajectory is planned in IFG (and other premotor structures). If covert production is the goal, the planned articulation signal bypasses M1 and is simulated internally. The somatosensory consequence of the simulated articulation is estimated in inferior parietal cortex, and the auditory consequence in the form of an abstract auditory representation is derived from the subsequent estimation. A



highly specified auditory representation (thick arrow) is obtained in a bottom-up perceptual process that goes through spectrotemporal analysis in STG (brown). The dorsal stream in which the motor simulation and perceptual estimation processes are available can by hypothesis enrich the specificity of predicted auditory representations (solid arrows), compared with ventral stream based memory retrieval (dotted arrows).

Abbreviations: pSTG= posterior STG; pMTG= posterior middle temporal gyrus; MTL= middle temporal lobe; M1= primary motor cortex.

underpins the coordinate transformations and transfer of phonological codes from temporal regions to frontal articulation networks. In the context of internal prediction in speech production, we hypothesize that the information flow is in an opposite direction in the dual streams. Specifically, in the dorsal stream, the articulatory code is transformed to a phonological code via somatosensory estimation (corresponding to phonemic coding); whereas in the ventral stream, episodic and semantic memory is retrieved in a conserved fashion compared with comprehension. That is, articulation imagery is hypothesized as simulation and estimation in the dorsal stream, whereas the hearing imagery is hypothesized as memory retrieval in the ventral stream. This new hypothesis provides an unforeseen and provocative new framework to analyze the processing of prediction in the perception and production of speech.

Indisputably, when the perceptual processes are largely bottom-up, one observes suppression effects. Therefore, the effects we report here must be because of top-down factors. Two such factors are particularly relevant: (i) attention/pre-cueing versus (ii) featural prediction. A recent theory by Kok, Rahnev, Jehee, Lau, and de Lange (2011) argues that a cognitive control function (termed “precision”) actively weights and scales the magnitude of the following perceptual responses. That is, attention is hypothesized to increase the precision of a prediction and scale up the responses (enhancement effects). On the basis of the hypothesized dorsal and ventral differences, we conjecture that the motor simulation and perceptual estimation (during articulation imagery) deriving from the dorsal stream leads to more precise prediction than the hearing imagery condition that derives from the ventral stream. The more precise prediction and attentional process would then scale up the weighting function and lead to the observed enhancement effect. In short, we propose that the task demands and contextual influences determine which pathway is preferentially activated. This provides a new mechanistic perspective on how dorsal and ventral stream structures interact with on-line tasks.

Motor simulation or articulator movement would enrich the detail of the auditory representation. As in our proposed sequential estimation model (Tian & Poeppel, 2010, 2012 and here in the dorsal stream), somatosensory estimation precedes auditory estimation. As in the recent model proposed by Hickok (2012), there is a motor phoneme estimation stage, which is consistent with our proposed somatosensory estimation. Such somatosensory estimation will provide detailed motor-to-sensory transformation dynamics that enrich the details of the representation, leading perhaps from phonemic to phonetic levels of detail, which is not available in the memory retrieval route. This is consistent with Oppenheim and Dell’s (2008, 2010) proposal that motor engagement can enrich the concreteness of the content during speech imagery.

The proposed dual stream prediction model is a strong case that distinguishes simulation-based and memory retrieval-based top-down routes that generate similar auditory representation. Speech imagery that at least includes articulation and hearing imagery could involve both production and perception. Interpreted in the context of our dual stream prediction model, both articulation imagery and hearing imagery could involve motor simulation. But we speculate that hearing imagery is the result of the combination of motor simulation and memory retrieval processes. That is, hearing imagery is the intermediate stage that balances between simulation and memory retrieval. Different weights giving to simulation and memory retrieval routes for inducing auditory representation could lead to the observed distinct modulation effects of articulation and hearing imagery. Future studies should test the anatomical and functional hypotheses generated by the proposed dual stream prediction model. The hypothesis that motor simulation available in the dorsal prediction stream can enrich the auditory representation will also be investigated. Finally, it is important to evaluate to what extent information that is preferentially processed in the two streams is parallel and independent versus concurrent but interactive.

The directional differences in our repetition effects underscore the adaptive nature of the underlying computations. Overt perception (*H*) induces suppression in subsequent responses to repeated auditory stimuli, replicating the well-established repetition suppression in perception using fMRI (Altmann, Doehrmann, & Kaiser, 2007; Dehaene Lambertz et al., 2006; Bergerbest et al., 2004; Belin & Zatorre, 2003) and EEG/MEG (Altmann et al., 2008; Ahveninen et al., 2006; Jääskeläinen et al., 2004; Rosburg, 2004). Interestingly, covert perception also leads to repetition suppression, which suggests that covert “perceptual” processes, much like in overt perception, scale down the sensitivity of subsequent activity. But remarkably, covert and overt production induce repetition enhancement, suggesting that actively formed neural representations and, by consequence efference copies, can specifically enhance the sensitivity of predicted upcoming perceptual processes.

The repetition suppression observed in our perceptual conditions supports predictive coding theory (Winkler, Denham, & Nelken, 2009; Bar, 2007; Friston, 2005) in which the current input is used in a Bayesian fashion to presensitize relevant representations and minimize the prediction error in subsequent perception. Conversely, the repetition enhancement observed in the articulation conditions agrees with feature-based attention theory (Summerfield & Egner, 2009), in which the expectation prioritizes the preselected features and boosts the sensitivity of particular features during subsequent perception. Our results demonstrate that these two competing theories—that predict opposite aftereffects—can be tentatively reconciled by considering the contextual

influence and task relevance, in the context of two anatomically distinct processing streams.

The adaptive, plastic nature of what must be considered highly similar neural representations can be understood by considering the direction of the component processing steps (bottom-up versus top-down) and how they interact with the context of processing (the task demands). The ability to detect unusual or unanticipated stimuli is essential (Kohonen, 1988; Sokolov & Vinogradova, 1975; James, 1890). Repetition suppression may provide one mechanism to block resources for unnecessary, redundant information (Jääskeläinen et al., 2004), and hence facilitate the efficient detection of ecologically relevant novel stimuli (Tiitinen et al., 1994). On the other hand, when the perception of upcoming (auditory or other) stimuli is the goal of the task, such as understanding speech in noisy conditions using visual cues (Grant & Seitz, 2000) and expecting to perceive stimuli with particular features in noisy and challenging environments (Stokes et al., 2009; Summerfield et al., 2006), more weight would be given to predicted features of stimuli, increasing the sensitivity to the repeated features. Indeed, the nature of task can lead to switching between different neural mechanisms (Scolari & Serences, 2009; Jazayeri & Movshon, 2007), and task demands have been demonstrated to balance between enhancement and suppression in auditory receptive fields (Neelon, Williams, & Garell, 2011).

We assume that top-down processes (that can be driven by dorsal or ventral structures; Figure 7) create a template (from memory) based on the task demands. Should the results of the sensory process fit the template, perceptual “success” is established. To exemplify, in the cases of Esterman and Yantis (2010), Elhilali et al. (2009), Eger, Henson, Driver, and Dolan (2007), and Dolan et al. (1997), goal-directed attention provides such a template (target frequency in the Elhilali study and specific category in the Dolan, Eger, and Esterman studies), and the template—kept in working memory during the task—induces the enhancement to the predicted features.

Additional evidence from other studies supports the plausibility of repetition enhancement. A compelling example comes from electrophysiological recordings in macaque visual area V4, where Rainer, Lee, and Logothetis (2004) reported such an effect for visual stimuli presented in a one-shot repetition design very much like ours. This is a closely related example, although there are other instances of repetition enhancement, for instance, in the studies of longer-term plasticity, both using fMRI (Kourtzi, Betts, Sarkheil, & Welchman, 2005) and EEG (Chandrasekaran, Hornickel, Skoe, Nicol, & Kraus, 2009). The repetition enhancement, compared with repetition suppression, suggests that the direction of modulation effect could be switched on the basis of content, task demand, and distinct neural pathways that reverse the repetition effect from “dampening” to “amplifying” the repeated representation (e.g., Thoma & Henson, 2011;

Nakamura, Dehaene, Jobert, Le Bihan, & Kouider, 2007; Turk-Browne, Yi, Leber, & Chun, 2007; Henson, Shallice, & Dolan, 2000).

In summary, we observe the “classical” repetition suppression effect in cases of (overt or imagined) perception but observe—in sharp contrast—a repetition enhancement effect in the case of (overt or covert) production. This means that simply repeating a stimulus is not captured by the most straightforward model. The details of the task demands matter greatly and in fact alter the neuronal processing in temporally (M100 vs. M200) and directionally (suppression vs. enhancement) precise ways. These findings provide a new way to think about repetition, both as a phenomenon but also as a tool to study neuronal representation.

We draw three conclusions. First, because we have demonstrated cross-modal repetition effects between (overt and covert) speech production and perception, we suggest that highly similar neural populations underlie the representation of auditory efference copies (production related), auditory memory (covert), and “real” (overt) perception. Second, the different temporal characteristics are consistent with the view that top-down (internally generated) and bottom-up (stimulus driven) representations activate different levels of a processing hierarchy. Third, the direction of repetition effects and their possible association with dorsal and ventral processing streams suggest a high degree of functional specificity, depending on task demands and contextual requirements. Thus, the MEG evidence is compelling that internally generated representations such as efference copies guide subsequent perception in a functionally specific manner.

Acknowledgments

We thank Jeff Walker for his excellent technical support and invaluable comments by Jean Mary Zarate, Luc Arnal, and Nai Ding. This study was supported by MURI ARO 54228-LS-MUR and NIH 2R01DC 05660.

Reprint requests should be sent to Xing Tian, Department of Psychology, NYU, 6 Washington Place, New York, NY 10003, or via e-mail: xing.tian@nyu.edu.

REFERENCES

- Ahveninen, J., Jääskeläinen, I. P., Raij, T., Bonmassar, G., Devore, S., Hämäläinen, M., et al. (2006). Task-modulated “what” and “where” pathways in human auditory cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *103*, 14608.
- Altmann, C. F., Doehrmann, O., & Kaiser, J. (2007). Selectivity for animal vocalizations in the human auditory cortex. *Cerebral Cortex*, *17*, 2601.
- Altmann, C. F., Nakata, H., Noguchi, Y., Inui, K., Hoshiyama, M., Kaneoke, Y., et al. (2008). Temporal dynamics of adaptation to natural sounds in the human auditory cortex. *Cerebral Cortex*, *18*, 1350.

- Bar, M. (2007). The proactive brain: Using analogies and associations to generate predictions. *Trends in Cognitive Sciences*, *11*, 280–289.
- Behroozmand, R., Karvelis, L., Liu, H., & Larson, C. R. (2009). Vocalization-induced enhancement of the auditory cortex responsiveness during voice F0 feedback perturbation. *Clinical Neurophysiology*, *120*, 1303–1312.
- Behroozmand, R., & Larson, C. R. (2011). Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. *BMC Neuroscience*, *12*, 54.
- Behroozmand, R., Liu, H., & Larson, C. R. (2011). Time-dependent neural processing of auditory feedback during voice pitch error detection. *Journal of Cognitive Neuroscience*, *23*, 1205–1217.
- Belin, P., & Zatorre, R. J. (2003). Adaptation to speaker's voice in right anterior temporal lobe. *NeuroReport*, *14*, 2105.
- Bentin, S., & McCarthy, G. (1994). The effects of immediate stimulus repetition on reaction time and event-related potentials in tasks of different complexity. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *20*, 130.
- Bergerbest, D., Ghahremani, D. G., & Gabrieli, J. D. E. (2004). Neural correlates of auditory repetition priming: Reduced fMRI activation in the auditory cortex. *Journal of Cognitive Neuroscience*, *16*, 966–977.
- Blakemore, S. J., & Decety, J. (2001). From the perception of action to the understanding of intention. *Nature Reviews Neuroscience*, *2*, 561–567.
- Bunzeck, N., Wuestenberg, T., Lutz, K., Heinze, H. J., & Jancke, L. (2005). Scanning silence: Mental imagery of complex sounds. *NeuroImage*, *26*, 1119–1127.
- Burnett, T. A., Freedland, M. B., Larson, C. R., & Hain, T. C. (1998). Voice F0 responses to manipulations in pitch feedback. *Journal of the Acoustical Society of America*, *103*, 3153.
- Calvert, G. A., Campbell, R., & Brammer, M. J. (2000). Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Current Biology*, *10*, 649–658.
- Chandrasekaran, B., Hornickel, J., Skoe, E., Nicol, T., & Kraus, N. (2009). Context-dependent encoding in the human auditory brainstem relates to hearing speech in noise: Implications for developmental dyslexia. *Neuron*, *64*, 311–319.
- Chong, T. T. J., Cunnington, R., Williams, M. A., Kanwisher, N., & Mattingley, J. B. (2008). fMRI adaptation reveals mirror neurons in human inferior parietal cortex. *Current Biology*, *18*, 1576–1580.
- 'de Cheveigné, A., & Simon, J. Z. (2007). Denoising based on time-shift PCA. *Journal of Neuroscience Methods*, *165*, 297–305.
- Dale, A. M., Liu, A. K., Fischl, B. R., Buckner, R. L., Belliveau, J. W., Lewine, J. D., et al. (2000). Dynamic statistical parametric mapping: Combining fMRI and MEG for high-resolution imaging of cortical activity. *Neuron*, *26*, 55–67.
- Davelaar, E. J., Tian, X., Weidemann, C. T., & Huber, D. E. (2011). A habituation account of change detection in same/different judgments. *Cognitive, Affective & Behavioral Neuroscience*, *11*, 608–626.
- Davidson, P. R., & Wolpert, D. M. (2005). Widespread access to predictive models in the motor system: A short review. *Journal of Neural Engineering*, *2*, S313.
- Dechent, P., Merboldt, K. D., & Frahm, J. (2004). Is the human primary motor cortex involved in motor imagery? *Cognitive Brain Research*, *19*, 138–144.
- Dehaene Lambertz, G., Dehaene, S., Anton, J. L., Campagne, A., Ciucu, P., Dehaene, G. P., et al. (2006). Functional segregation of cortical language areas by sentence repetition. *Human Brain Mapping*, *27*, 360–371.
- Deiber, M. P., Ibañez, V., Honda, M., Sadato, N., Raman, R., & Hallett, M. (1998). Cerebral processes related to visuomotor imagery and generation of simple finger movements studied with positron emission tomography. *NeuroImage*, *7*, 73–85.
- Desmurget, M., & Sirigu, A. (2009). A parietal-premotor network for movement intention and motor awareness. *Trends in Cognitive Sciences*, *13*, 411–419.
- Doehrmann, O., Weigelt, S., Altmann, C. F., Kaiser, J., & Naumer, M. J. (2010). Audiovisual functional magnetic resonance imaging adaptation reveals multisensory integration effects in object-related sensory cortices. *Journal of Neuroscience*, *30*, 3370.
- Dolan, R., Fink, G., Rolls, E., Booth, M., Holmes, A., Frackowiak, R., et al. (1997). How the brain learns to see objects and faces in an impoverished context. *Nature*, *389*, 596–598.
- Eger, E., Henson, R., Driver, J., & Dolan, R. (2007). Mechanisms of top-down facilitation in perception of visual objects studied by fMRI. *Cerebral Cortex*, *17*, 2123–2133.
- Ehrsson, H. H., Geyer, S., & Naito, E. (2003). Imagery of voluntary movement of fingers, toes, and tongue activates corresponding body-part-specific motor representations. *Journal of Neurophysiology*, *90*, 3304–3316.
- Elhilali, M., Xiang, J., Shamma, S. A., & Simon, J. Z. (2009). Interaction between attention and bottom-up saliency mediates the representation of foreground and background in an auditory scene. *PLoS Biology*, *7*, e1000129.
- Eliades, S. J., & Wang, X. (2003). Sensory-motor interaction in the primate auditory cortex during self-initiated vocalizations. *Journal of Neurophysiology*, *89*, 2194.
- Eliades, S. J., & Wang, X. (2005). Dynamics of auditory-vocal interaction in monkey auditory cortex. *Cerebral Cortex*, *15*, 1510.
- Eliades, S. J., & Wang, X. (2008). Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature*, *453*, 1102–1106.
- Esterman, M., & Yantis, S. (2010). Perceptual expectation evokes category-selective cortical activity. *Cerebral Cortex*, *20*, 1245.
- Fischl, B., Sereno, M. I., & Dale, A. M. (1999). Cortical surface-based analysis. II: Inflation, flattening, and a surface-based coordinate system. *NeuroImage*, *9*, 195–207.
- Fischl, B., Sereno, M. I., Tootell, R. B. H., & Dale, A. M. (1999). High-resolution intersubject averaging and a coordinate system for the cortical surface. *Human Brain Mapping*, *8*, 272–284.
- Ford, J. M., Roach, B. J., & Mathalon, D. H. (2010). Assessing corollary discharge in humans using noninvasive neurophysiological methods. *Nature Protocols*, *5*, 1160–1168.
- Friston, K. (2005). A theory of cortical responses. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *360*, 815.
- Gauthier, G. M., Nommay, D., & Vercher, J. L. (1990a). Ocular muscle proprioception and visual localization of targets in man. *Brain*, *113*, 1857–1871.
- Gauthier, G. M., Nommay, D., & Vercher, J. L. (1990b). The role of ocular muscle proprioception in visual localization of targets. *Science*, *249*, 58–61.
- 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. *Cerebral Cortex*, *10*, 1093–1104.
- Grant, K. W., & Seitz, P. F. (2000). The use of visible speech cues for improving auditory detection of spoken sentences. *Journal of the Acoustical Society of America*, *108*, 1197.
- Grill-Spector, K., Henson, R., & Martin, A. (2006). Repetition and the brain: Neural models of stimulus-specific effects. *Trends in Cognitive Sciences*, *10*, 14–23.

- Grill-Spector, K., & Malach, R. (2001). fMR-adaptation: A tool for studying the functional properties of human cortical neurons. *Acta Psychologica*, *107*, 293–321.
- Grush, R. (2004). The emulation theory of representation: Motor control, imagery, and perception. *Behavioral and Brain Sciences*, *27*, 377–396.
- Guenther, F. H. (1995). Speech sound acquisition, coarticulation, and rate effects in a neural network model of speech production. *Psychological Review*, *102*, 594–620.
- Guenther, F. H., Ghosh, S. S., & Tourville, J. A. (2006). Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and Language*, *96*, 280–301.
- Guenther, F. H., Hampson, M., & Johnson, D. (1998). A theoretical investigation of reference frames for the planning of speech movements. *Psychological Review*, *105*, 611–633.
- Halpern, A. R., & Zatorre, R. J. (1999). When that tune runs through your head: A PET investigation of auditory imagery for familiar melodies. *Cerebral Cortex*, *9*, 697–704.
- Hämäläinen, M. S., Hari, R., Ilmoniemi, R. J., Knuutila, J., & Lounasmaa, O. V. (1993). Magnetoencephalography-theory, instrumentation, and applications to noninvasive studies of the working human brain. *Reviews of Modern Physics*, *65*, 413–497.
- Hämäläinen, M. S., & Ilmoniemi, R. (1994). Interpreting magnetic fields of the brain: Minimum norm estimates. *Medical & Biological Engineering & Computing*, *32*, 35–42.
- Hanakawa, T., Immisch, I., Toma, K., Dimyan, M. A., Van Gelderen, P., & Hallett, M. (2003). Functional properties of brain areas associated with motor execution and imagery. *Journal of Neurophysiology*, *89*, 989–1002.
- Heinemann, L. V., Kaiser, J., & Altmann, C. F. (2011). Auditory repetition enhancement at short interstimulus intervals for frequency-modulated tones. *Brain Research*, *1411*, 65–75.
- Heinks-Maldonado, T. H., Mathalon, D. H., Gray, M., & Ford, J. M. (2005). Fine-tuning of auditory cortex during speech production. *Psychophysiology*, *42*, 180–190.
- Heinks-Maldonado, T. H., Nagarajan, S. S., & Houde, J. F. (2006). Magnetoencephalographic evidence for a precise forward model in speech production. *NeuroReport*, *17*, 1375–1379.
- Henson, R. (2003). Neuroimaging studies of priming. *Progress in Neurobiology*, *70*, 53–81.
- Henson, R., Shallice, T., & Dolan, R. (2000). Neuroimaging evidence for dissociable forms of repetition priming. *Science*, *287*, 1269–1272.
- Hickok, G. (2012). Computational neuroanatomy of speech production. *Nature Reviews Neuroscience*, *13*, 135–145.
- Hickok, G., Houde, J., & Rong, F. (2011). Sensorimotor integration in speech processing: Computational basis and neural organization. *Neuron*, *69*, 407–422.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature Reviews Neuroscience*, *8*, 393–402.
- Houde, J. F., Nagarajan, S. S., Sekihara, K., & Merzenich, M. M. (2002). Modulation of the auditory cortex during speech: An MEG study. *Journal of Cognitive Neuroscience*, *14*, 1125–1138.
- Hubbard, T. L. (2010). Auditory imagery: Empirical findings. *Psychological Bulletin*, *136*, 302.
- Huber, D. E. (2008). Immediate priming and cognitive aftereffects. *Journal of Experimental Psychology: General*, *137*, 324.
- Huber, D. E., Tian, X., Curran, T., O'Reilly, R. C., & Worocho, B. (2008). The dynamics of integration and separation: ERP, MEG, and neural network studies of immediate repetition effects. *Journal of Experimental Psychology: Human Perception and Performance*, *34*, 1389–1416.
- Jääskeläinen, I. P., Ahveninen, J., Bonmassar, G., Dale, A. M., Ilmoniemi, R. J., Levänen, S., et al. (2004). Human posterior auditory cortex gates novel sounds to consciousness. *Proceedings of the National Academy of Sciences, U.S.A.*, *101*, 6809.
- James, W. (1890). *The principles of psychology* (Vol. 1). New York: Henry Holt and Co.
- Jazayeri, M., & Movshon, J. A. (2007). A new perceptual illusion reveals mechanisms of sensory decoding. *Nature*, *446*, 912–915.
- Jeannerod, M. (1994). The representing brain: Neural correlates of motor intention and imagery. *Behavioral and Brain Sciences*, *17*, 187–202.
- Jeannerod, M. (1995). Mental imagery in the motor context. *Neuropsychologia*, *33*, 1419–1432.
- Kastner, S., Pinsk, M. A., De Weerd, P., Desimone, R., & Ungerleider, L. G. (1999). Increased activity in human visual cortex during directed attention in the absence of visual stimulation. *Neuron*, *22*, 751–761.
- Katahira, K., Abta, D., Masuda, S., & Okanoya, K. (2008). Feedback-based error monitoring processes during musical performance: An ERP study. *Neuroscience Research*, *61*, 120–128.
- Kawato, M. (1999). Internal models for motor control and trajectory planning. *Current Opinion in Neurobiology*, *9*, 718–727.
- Kohonen, T. (1988). *Self-organization and associative memory* (Vol. 8). Berlin: Springer-Verlag.
- Kok, P., Rahnev, D., Jehee, J. F. M., Lau, H. C., & de Lange, F. P. (2011). Attention reverses the effect of prediction in silencing sensory signals. *Cerebral Cortex*, *22*, 2197–2206.
- Kosslyn, S. M., Ganis, G., & Thompson, W. L. (2001). Neural foundations of imagery. *Nature Reviews Neuroscience*, *2*, 635–642.
- Kourtzi, Z., Betts, L. R., Sarkheil, P., & Welchman, A. E. (2005). Distributed neural plasticity for shape learning in the human visual cortex. *PLoS Biol*, *3*, e204.
- Kourtzi, Z., & Kanwisher, N. (2000). Cortical regions involved in perceiving object shape. *Journal of Neuroscience*, *20*, 3310.
- Kraemer, D. J. M., Macrae, C. N., Green, A. E., & Kelley, W. M. (2005). Musical imagery sound of silence activates auditory cortex. *Nature*, *434*, 158.
- Lin, F. H., Belliveau, J. W., Dale, A. M., & Hämäläinen, M. S. (2006). Distributed current estimates using cortical orientation constraints. *Human Brain Mapping*, *27*, 1–13.
- 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. *Journal of Cognitive Neuroscience*, *11*, 491–501.
- Mahon, B. Z., Milleville, S. C., Negri, G. A. L., Rumiati, R. I., Caramazza, A., & Martin, A. (2007). Action-related properties shape object representations in the ventral stream. *Neuron*, *55*, 507–520.
- 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. *Cognitive Brain Research*, *19*, 219–228.
- Miall, R. C., & Wolpert, D. M. (1996). Forward models for physiological motor control. *Neural Networks*, *9*, 1265–1279.
- Nakamura, K., Dehaene, S., Jobert, A., Le Bihan, D., & Kouider, S. (2007). Task-specific change of unconscious neural priming in the cerebral language network. *Proceedings of the National Academy of Sciences*, *104*, 19643–19648.
- Neelon, M. F., Williams, J. C., & Garell, P. C. (2011). Elastic attention: Enhanced, then sharpened response to auditory input as attentional load increases. *Frontiers in Human Neuroscience*, *5*, doi:10.3389/fnhum.2011.00041.

- Numminen, J., Salmelin, R., & Hari, R. (1999). Subject's own speech reduces reactivity of the human auditory cortex. *Neuroscience Letters*, *265*, 119–122.
- Oppenheim, G. M., & Dell, G. S. (2008). Inner speech slips exhibit lexical bias, but not the phonemic similarity effect. *Cognition*, *106*, 528–537.
- Oppenheim, G. M., & Dell, G. S. (2010). Motor movement matters: The flexible abstractness of inner speech. *Memory & Cognition*, *38*, 1147–1160.
- Peelen, M. V., Fei-Fei, L., & Kastner, S. (2009). Neural mechanisms of rapid natural scene categorization in human visual cortex. *Nature*, *460*, 94–97.
- Phillips, C., Pellathy, T., Marantz, A., Yellin, E., Wexler, K., Poeppel, D., et al. (2000). Auditory cortex accesses phonological categories: An MEG mismatch study. *Journal of Cognitive Neuroscience*, *12*, 1038–1055.
- Poeppel, D., Idsardi, W. J., & Van Wassenhove, V. (2008). Speech perception at the interface of neurobiology and linguistics. *Philosophical Transactions of the Royal Society of London, Series B, Biological Sciences*, *363*, 1071.
- Price, C. J., Crinion, J. T., & MacSweeney, M. (2011). A generative model of speech production in Broca's and Wernicke's areas. *Frontiers in Psychology*, *2*, doi:10.3389/fpsyg.2011.00237.
- Rainer, G., Lee, H., & Logothetis, N. K. (2004). The effect of learning on the function of monkey extrastriate visual cortex. *PLoS Biology*, *2*, e44.
- Rauschecker, J. P., & Scott, S. K. (2009). Maps and streams in the auditory cortex: Nonhuman primates illuminate human speech processing. *Nature Neuroscience*, *12*, 718–724.
- Roberts, T. P. L., Ferrari, P., Stufflebeam, S. M., & Poeppel, D. (2000). Latency of the auditory evoked neuromagnetic field components: Stimulus dependence and insights toward perception. *Journal of Clinical Neurophysiology*, *17*, 114–129.
- Rosburg, T. (2004). Effects of tone repetition on auditory evoked neuromagnetic fields. *Clinical Neurophysiology*, *115*, 898–905.
- Rugg, M. D. (1985). The effects of semantic priming and word repetition on event-related potentials. *Psychophysiology*, *22*, 642–647.
- Schacter, D. L., Wig, G. S., & Stevens, W. D. (2007). Reductions in cortical activity during priming. *Current Opinion in Neurobiology*, *17*, 171–176.
- Schürmann, M., Raij, T., Fujiki, N., & Hari, R. (2002). Mind's ear in a musician: Where and when in the brain. *Neuroimage*, *16*, 434–440.
- Scolari, M., & Serences, J. T. (2009). Adaptive allocation of attentional gain. *Journal of Neuroscience*, *29*, 11933.
- Sirigu, A., Duhamel, J.-R., Cohen, L., Pillon, B., Dubois, B., & Agid, Y. (1996). The mental representation of hand movements after parietal cortex damage. *Science*, *273*, 1564–1568.
- Skipper, J. I., van Wassenhove, V., Nusbaum, H. C., & Small, S. L. (2007). Hearing lips and seeing voices: How cortical areas supporting speech production mediate audiovisual speech perception. *Cerebral Cortex*, *17*, 2387–2399.
- Sokolov, E. N., & Vinogradova, O. S. (1975). *Neuronal mechanisms of the orienting reflex*. New York: Halsted Press.
- Sommer, M. A., & Wurtz, R. H. (2006). Influence of the thalamus on spatial visual processing in frontal cortex. *Nature*, *444*, 374–377.
- Sommer, M. A., & Wurtz, R. H. (2008). Brain circuits for the internal monitoring of movements. *Annual Review of Neuroscience*, *31*, 317–338.
- Stokes, M., Thompson, R., Nobre, A. C., & Duncan, J. (2009). Shape-specific preparatory activity mediates attention to targets in human visual cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *106*, 19569.
- Summerfield, C., & Egner, T. (2009). Expectation (and attention) in visual cognition. *Trends in Cognitive Sciences*, *13*, 403–409.
- Summerfield, C., Egner, T., Greene, M., Koechlin, E., Mangels, J., & Hirsch, J. (2006). Predictive codes for forthcoming perception in the frontal cortex. *Science*, *314*, 1311.
- Thoma, V., & Henson, R. N. (2011). Object representations in ventral and dorsal visual streams: fMRI repetition effects depend on attention and part-whole configuration. *Neuroimage*, *57*, 513–525.
- Tian, X., & Huber, D. E. (2008). Measures of spatial similarity and response magnitude in MEG and scalp EEG. *Brain Topography*, *20*, 131–141.
- Tian, X., & Poeppel, D. (2010). Mental imagery of speech and movement implicates the dynamics of internal forward models. *Frontiers in Psychology*, *1*, doi:10.3389/fpsyg.2010.00166.
- Tian, X., & Poeppel, D. (2012). Mental imagery of speech: Linking motor and perceptual systems through internal simulation and estimation. *Frontiers in Human Neuroscience*, *6*, doi: 10.3389/fnhum.2012.00314.
- Tian, X., Poeppel, D., & Huber, D. E. (2011). TopoToolbox: Using sensor topography to calculate psychologically meaningful measures from event-related EEG/MEG. *Computational Intelligence and Neuroscience*, *2011*, doi:10.1155/2011/674605.
- Tiitinen, H., May, P., Reinikainen, K., & Näätänen, R. (1994). Attentive novelty detection in humans is governed by pre-attentive sensory memory. *Nature*, *372*, 90–92.
- Todorov, E., & Jordan, M. I. (2002). Optimal feedback control as a theory of motor coordination. *Nature Neuroscience*, *5*, 1226–1235.
- Tourville, J. A., Reilly, K. J., & Guenther, F. H. (2008). Neural mechanisms underlying auditory feedback control of speech. *Neuroimage*, *39*, 1429–1443.
- Turk-Browne, N., Yi, D., Leber, A., & Chun, M. (2007). Visual quality determines the direction of neural repetition effects. *Cerebral Cortex*, *17*, 425.
- Ulanovsky, N., Las, L., & Nelken, I. (2003). Processing of low-probability sounds by cortical neurons. *Nature Neuroscience*, *6*, 391–398.
- van Atteveldt, N., Formisano, E., Goebel, R., & Blomert, L. (2004). Integration of letters and speech sounds in the human brain. *Neuron*, *43*, 271–282.
- van Wassenhove, V., Grant, K. W., & Poeppel, D. (2005). Visual speech speeds up the neural processing of auditory speech. *Proceedings of the National Academy of Sciences, U.S.A.*, *102*, 1181.
- Ventura, M., Nagarajan, S., & Houde, J. (2009). Speech target modulates speaking induced suppression in auditory cortex. *BMC Neuroscience*, *10*, 58.
- von Helmholtz, H. (1910). *Handbuch der physiologischen Optik* (3rd ed.). In A. Gullstrand, J. v. Kries, & W. N. Voss (Eds.). Hamburg: Leopold Voss.
- Von Holst, E., & Mittelstaedt, H. (1950). Das Reafferenzprinzip. Wechselwirkungen zwischen Zentralnervensystem und Peripherie. *Naturwissenschaften*, *37*, 467–476.
- Von Holst, E., & Mittelstaedt, H. (1973). The reafference principle. In R. Martin (Trans.), *The behavioral physiology of animals and man: The collected papers of Erich von Holst* (pp. 139–173). Coral Gables, FL: University of Miami Press.
- Wheeler, M. E., Petersen, S. E., & Buckner, R. L. (2000). Memory's echo: Vivid remembering reactivates sensory-specific cortex. *Proceedings of the National Academy of Sciences, U.S.A.*, *97*, 11125–11129.

- Winkler, I., Denham, S. L., & Nelken, I. (2009). Modeling the auditory scene: Predictive regularity representations and perceptual objects. *Trends in Cognitive Sciences*, *13*, 532–540.
- Winston, J. S., Henson, R., Fine-Goulden, M. R., & Dolan, R. J. (2004). fMRI-adaptation reveals dissociable neural representations of identity and expression in face perception. *Journal of Neurophysiology*, *92*, 1830.
- Wolpert, D. M., & Ghahramani, Z. (2000). Computational principles of movement neuroscience. *Nature Neuroscience*, *3*, 1212–1217.
- Zatorre, R. J., & Halpern, A. R. (2005). Mental concerts: Musical imagery and auditory cortex. *Neuron*, *47*, 9–12.
- Zelano, C., Mohanty, A., & Gottfried, J. A. (2011). Olfactory predictive codes and stimulus templates in piriform cortex. *Neuron*, *72*, 178–187.
- Zheng, Z. Z., Munhall, K. G., & Johnsrude, I. S. (2010). Functional overlap between regions involved in speech perception and in monitoring one's own voice during speech production. *Journal of Cognitive Neuroscience*, *22*, 1770–1781.