Motor planning modulates neural activity patterns in early human auditory cortex

Daniel J. Gale[1], Corson N. Areshenkoff [1,2], Claire Honda[2], Ingrid S. Johnsrude[5,6,7], J. Randall Flanagan[1,2], & Jason P. Gallivan*[1,2,3]

[1]Centre for Neuroscience Studies, [2]Department of Psychology, and [3]Department of Biomedical and Molecular Sciences, Queen's University, Kingston, Ontario, Canada. [5]Department of Psychology, School of Communication Sciences and Disorders[6], and [7]Brain and Mind Institute, University of Western Ontario

Abbreviated Title: Movement planning modulates auditory cortex

Number of Pages: 43

Number of words in Abstract: 219

Number of words in Introduction: 650

Number of words in Discussion: 1463

Number of Figures: 4

Number of Extended Data Figures: 4

*Correspondence should be addressed to:

Jason Gallivan

Centre for Neuroscience Studies

Queen's University

gallivan@queensu.ca

Abstract

It is well established that movement planning recruits motor-related cortical brain areas in preparation for the forthcoming action. Given that an integral component of the control of action is the processing of sensory information throughout movement, we predicted that movement planning also involves preparing early sensory cortical areas for participation in the impending behaviour. To test this hypothesis, we had human participants perform an object manipulation task wherein we focused on activity in early human auditory cortex, given the role of auditory signals in the sensorimotor control of such tasks and because of its known ipsilateral connections with the motor system. Here we show, using functional MRI and pattern classification methods, that information related to the limb to be used to grasp and lift an object can be decoded, well before movement, from neural activity patterns in early auditory cortex. We further show that the decoding of this motor-related information occurs in a separate subregion of auditory cortex than the decoding of the auditory sensory information used to instruct, and prompt preparation of, the hand actions. Together, this evidence suggests that action planning, in addition to preparing the motor system for movement, involves the task-specific preparation of primary sensory areas, such that they are set up to appropriately process sensory information arising during the unfolding movement.

Keywords

Motor, Planning, Audition, Auditory, Object manipulation, Action

Significance Statement:

The mammalian auditory system is characterized by an extensive, highly interconnected web of feedback projections, an architecture that, in humans, has been primarily implicated in sustaining auditory attention and facilitating working memory. Yet, its role in supporting the planning and control of everyday object-directed movements remains underappreciated and largely unstudied. Here we show, using functional MRI and a naturalistic object manipulation task, that hand movement planning modulates early auditory cortical activity patterns in an automatic, motor-specific fashion. These findings suggest that, prior to movement, the motor system prepares the neural state of early auditory cortex, readying it for the processing of sensory information arising during movement execution.

INTRODUCTION

Most theories of motor control distinguish between the planning of a movement and its subsequent execution. Research examining the neural basis of movement planning has commonly used delayed movement tasks—in which instructions about what movement to perform are separated in time from the instruction to initiate that movement—and has focused on delay period activity in motor-related brain areas. The conventional view has been that planning activity reflects the coding of one or more parameters—e.g., direction, extent, speed, curvature, force—of the forthcoming movement to be executed (e.g., Tanji and Evarts, 1976; Messier and Kalaska, 2000a). However, recent theories, which view neural activity from a dynamical systems perspective, offer a different interpretation; namely that movement planning involves preparing the initial neural state of the motor system, from which point movement execution evolves naturally through intrinsic neural processes (Shenoy et al., 2013a). To date, the focus of this work has been on understanding how changes in these neural activity patterns ultimately shape muscle activity (Churchland et al., 2012).

A critical component of skilled action is the prediction of the sensory consequences of motor commands (Wolpert and Flanagan, 2001). For example, the sensorimotor control of object manipulation tasks involves predicting the sensory signals associated with contact events, which define subgoals of the task (Flanagan et al., 2006). Thus, when reaching towards, lifting, and then replacing an object, the brain predicts sensory signals linked to contact between the digits and the object, the breaking of contact between the object and surface, and contact between the object and surface, respectively. Importantly, these signals can occur in multiple sensory modalities, including tactile, proprioceptive, visual, and auditory (Johansson and Flanagan, 2009). By comparing the predicted to actual sensory outcomes, the brain can monitor task progression, detect performance errors, and quickly launch appropriate, task-protective corrective actions as needed. For example, when lifting an object that is heavier than expected, anticipated tactile events, associated with lift-off, fail to occur at the expected

time, triggering a corrective response. Similar compensatory behaviour has been shown when anticipated auditory events fail to occur at the predicted time (Safstrom et al., 2014). Sensory prediction is also critical in sensory cancellation, the attenuation of predictable sensory events that arise as a consequence of movement. Such attenuation is thought to allow the brain to disambiguate sensory events that arise from movement from events that arise from external sources (Schneider and Mooney, 2018). Given the functional importance of predicting task-specific sensory consequences, we hypothesize that action planning, in addition to preparing motor areas for execution, involves the preparation of primary sensory areas for processing task-specific sensory signals. Given that these sensory signals will depend on the precise action being performed, this hypothesis predicts that neural activity in early sensory areas will represent motor-related information prior to movement, during action planning.

As a critical test of this hypothesis, here we examined, using human functional MRI and a delayed object lifting task, whether the hand used to lift the object can be decoded from delay period activity in auditory cortex. We manipulated the hand used for lifting because the preparatory neural activity associated with left and right hand movements are very different at the level of the motor system (Cisek et al., 2003). We focused our analysis on auditory cortex because of the role of auditory signals in the sensorimotor control of object manipulation tasks (Johansson and Flanagan, 2009), and because of its known ipsilateral connections with the motor system in mammals (Nelson et al., 2013). Finally, although previous work has shown that auditory cortex is modulated by motor inputs during movement execution (Reznik et al., 2015; Schneider et al., 2018), the focus of this work has been on real-time attenuation, during movement execution, of the predictable sensory consequences of action. Here we instead focus on planning-related activity, and the broader function of the motor system in preparing early sensory areas for participation in forthcoming sensorimotor control.

MATERIALS & METHODS

Overview

To test our hypothesis that the motor system modulates the neural state of early auditory cortex during planning, we had 16 participants perform a delayed object lifting task, thereby allowing us to separate motor planning-related modulations from the later motor execution and somatosensory-related modulations. In effect, this delayed movement task allowed us to ask whether the motor command being prepared—but not yet executed—can be decoded from neural activity patterns in early auditory cortex. To examine this, in each trial we had participants first prepare, and then execute (after a jittered delay interval) either a left or right hand object lift-and-replace action, which were cued by two nonsense auditory commands (“Compty” or “Midwig”; see Fig 1). Importantly, halfway throughout each experimental run, participants were required to switch the auditory command-to-hand mapping (i.e., if “Compty” cued a left hand object lift-and-replace action in the first half of the experimental run, then “Compty” would cue a right hand object lift-and-replace action in the second half of the experimental run; see Fig 1B). Critically, this design allowed us to examine early auditory cortex activity during the planning of two distinct hand actions (left versus right hand movements), with invariance to the actual auditory commands (e.g., "Compty") used to instruct those hand actions. As such, any neural differences in auditory cortex prior to movement on these trials are likely to reflect top-down modulations related to motor, and not bottom-up sensory, processing.

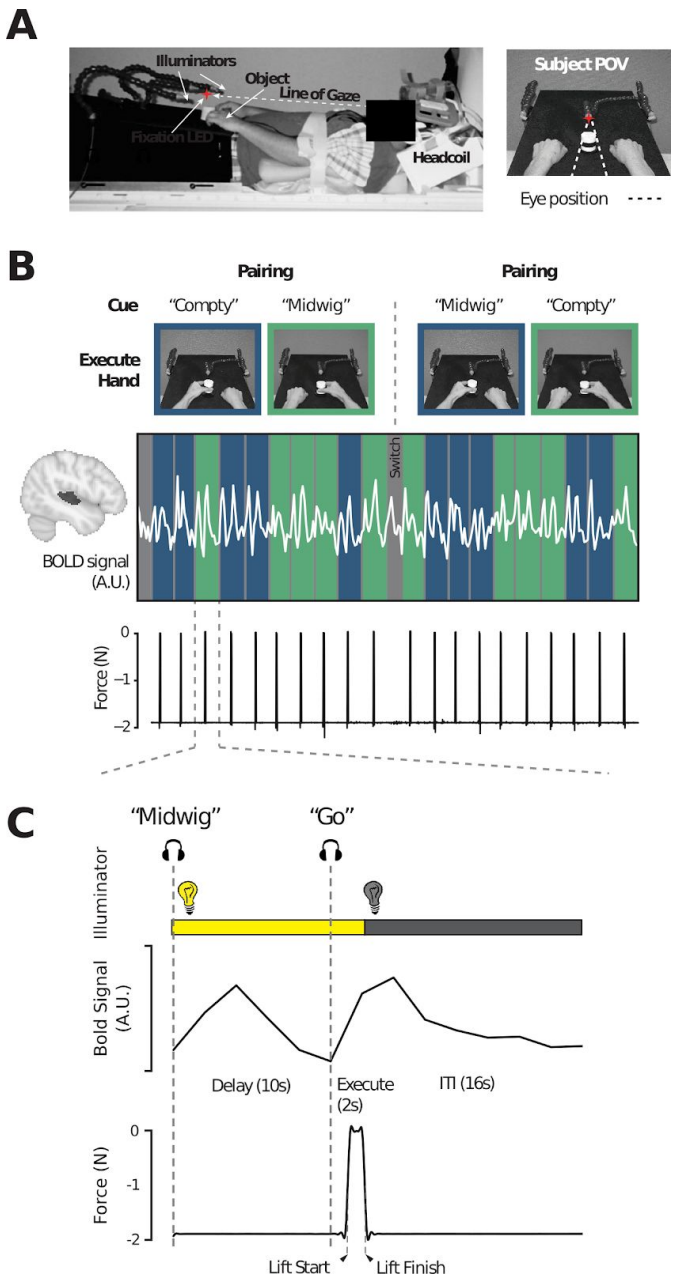

Figure 1. Experimental setup and task overview.
A. MRI setup (left) and subject point-of-view (right) of the experimental workspace. Red star indicates the fixation LED placed above the object. Illuminator LEDs, attached to the flexible stalks, are shown on the left and right. **B.** Example fMRI run of 20 task trials. Color-coded columns (blue = left hand, green = right hand) demarks each trial and the associated time-locked BOLD activity from superior temporal gyrus (STG; shaded in dark gray on cortex, left) of an exemplar subject is indicated by the overlaid white trace. Pairings between auditory cue ("Compty" or "Midwig") and hand (left or right) are indicated in the pictures above, and were reversed halfway through each run following a 'Switch' auditory cue, such that each hand is paired with each auditory cue in every experimental run (see Methods). The corresponding force sensor data, used to track object lifting, is shown below. **C.** Sequence of events and corresponding single-trial BOLD and force sensor data of an exemplar trial from a representative participant in which 'Midwig' cued a right-handed movement. Each trial begins with the hand workspace being illuminated while, simultaneously, participants receive the auditory cue ("Compty" or "Midwig"). This is then followed by a jittered 6-12s Delay interval (10s in this exemplar trial). Next, an auditory "Go" cue initiates the start of the 2s Execute epoch, during which the subject grasp-and-lifts the object (shown by the force trace; arrows indicate the start of the lift and object replacement). Following the 2 s Execute epoch, illumination of the workspace is extinguished and subjects then wait a fixed 16s intertrial interval (ITI) prior to onset of the next trial. See also Fig. 1-1 for a more detailed overview of the trial sequence and the data obtained from a separate behavioural training session.

Participants

Sixteen healthy right-handed subjects (8 females, 21-25 years of age) participated in one behavioural testing session followed by two fMRI testing sessions (a localizer testing session, and then the experimental testing session, performed on separate days approximately 1-2 weeks apart). Right-handedness was assessed with the Edinburgh handedness questionnaire (Oldfield, 1971). Informed consent and consent to publish were obtained in accordance with ethical standards set out by the Declaration of Helsinki (1964) and with procedures cleared by the Queen's University Health Sciences Research Ethics Board. Subjects were naïve with respect to the hypotheses under evaluation and received monetary compensation for their involvement in the study. Data from one subject were excluded from further analyses in the experimental testing session due to problems in data collection, resulting in a final sample size of 15 subjects. Meanwhile, all 16 subjects were used for the localizer testing session.

Experimental apparatus

The experimental setup for both the localizer and experimental testing sessions consisted of a black platform placed over the waist and tilted away from the horizontal at an angle (~15°) to maximize comfort and target visibility. The MRI head coil was tilted slightly (~20°) and foam cushions were used to give an approximate overall head tilt of 30°. To minimize limb-related artifacts, subjects had the right and left upper-arms braced, limiting movement of the arms to the elbow and thus creating an arc of reachability for each hand. The exact placement of object stimuli on the platform was adjusted to match each subject's arm length such that all required actions were comfortable and ensured that only movement of the forearm, wrist and fingers was required. The platform was illuminated by two bright white Light Emitting Diodes (LEDs) attached to flexible plastic stalks (Loc-Line, Lockwood Products, Lake Oswego, OR, USA) located to the left and right of the platform. To control for eye movements, a small red fixation LED, attached to a flexible plastic stalk, was positioned above the hand workspace and located ~5 cm beyond the target object position (such that the object

appeared in the subject's lower visual field). Experimental timing and lighting were controlled with in-house software created with C++ and MATLAB (The Mathworks, Natnick, MA). Throughout fMRI testing, the subject's hand movements were monitored using an MR-compatible infrared-sensitive camera (MRC Systems GmbH, Heidelberg, Germany), optimally positioned on one side of the platform and facing towards the subject. The videos captured during the experiment were analyzed offline to verify that the subjects were performing the task as instructed and to identify error trials.

Auditory Localizer Task

A separate, block-design localizer task was collected to independently identify auditory cortex and higher-order language regions of interest (ROIs) for use in the analyses of main experimental task. This auditory localizer task included three conditions: (1) Intact speech trials (Intact), which played one of 8 unique excerpts of different speeches; (2) scrambled speech trials (Scrambled), which were incoherent signal-correlated noise versions of the speech excerpts (i.e. applying an amplitude envelope of the speech to uniform Gaussian white noise, ensuring that the noise level was utterance-specific and exactly intense enough at every moment to mask the energy of the spoken words); and (3) rest trials (Rest), in which no audio was played (subjects thus only heard background MRI scanner noise during this time). Trials lasted 20 s each and alternated, in pseudo-random order, between Intact Speech, Scrambled Speech, and Rest for a total of 24 trials in each experimental run. In addition, a 20 s baseline block was placed at the beginning of each experimental run. Each localizer run totaled 500 s and participants completed 2 of these runs during testing (resulting in 16 repetitions per experimental condition per subject). To encourage that participants maintained attention throughout this auditory localizer run, they were required to monitor each of the Intact speeches and let the experimenter know, following the run, whether any of them were repeated. This repeat happened in only one of the experimental runs and each and every subject correctly identified the repeat and non-repeat run (100% accuracy).

Motor Localizer task

Four experimental runs of a motor localizer task were also collected alongside the auditory localizer task, which constituted a block-design protocol that alternated between subtasks designed to localize eight separate motor functions. Task set up and details for all eight conditions are described in Fig. 2-1. The hand grasping condition from this localizer task was used to define dorsal premotor cortex (PMd), which we used as a basis for comparison with our auditory cortex decoding findings (see Results).

The motor and auditory localizer testing session lasted approximately 2 hours and included set-up time, one 7.5 minute high-resolution anatomical scan and 6 functional scanning runs, wherein subjects alternated between performing two runs of the motor localizer task and one run of the auditory localizer, twice. A brief (~10 minute) practice session was carried out before the localizer testing session in the MRI control room in order to familiarize participants with localizer tasks.

Main Experimental Task

In our experimental task (see Fig. 1), we used a delayed movement paradigm wherein, on each individual trial, subjects were first auditorily cued (via headphones) to prepare either a left vs. right hand object grasp-and-lift action on a centrally located cylindrical object (1.9 N weight). Then, following a variable delay period, they were prompted to execute the prepared hand action. At the start of each event-related trial (see Fig. 1C), simultaneously with the LED illuminator lights going on (and the subject's workspace being illuminated), subjects received one of two nonsense speech cues, "Compty" or "Midwig". For a given trial, each nonsense speech cue was paired with a corresponding hand action (e.g., subjects were instructed that, for a predetermined set of trials, "Compty" cued a left hand movement whereas "Midwig" cued a right hand movement). [*Note that nonsense speech commands were chosen because semantically meaningful words such as "left" and "right" would already have strong cognitive associations for participants.*] Following the delivery of the auditory command, there was a jittered delay

interval of 6-12 s (a Gaussian random jitter with a mean of 9 s), after which the verbal auditory command “Go” was delivered, prompting subjects to execute the prepared grasp-and-lift action. For the execution phase of the trial, subjects were required to precision grasp-and-then-lift the object with their thumb and index finger (~2 cm above the platform, via a rotation of the wrist), hold it in midair for ~1 s, and then replace it. Subjects were instructed to keep the timing of each hand action as similar as possible throughout the study. Two seconds following the onset of this “Go” cue, the illuminator lights were extinguished, and subjects then waited 16 s for the next trial to begin (intertrial interval, ITI). Throughout the entire time course of the trial, subjects were required to maintain gaze on the fixation LED.

These event-related trials were completed in two separate blocks per experimental run. At the beginning of each experimental run, the experimenter informed subjects of the auditory-hand mapping to be used for the first 10 event-related trials of the experimental run (e.g. “Compty” for left hand (left hand) movements, “Midwig” for right hand (right hand) movements; 5 intermixed trials of each type). After the 10th trial, the illuminator was turned on (for a duration of 6 s) and subjects simultaneously heard the auditory command “Switch” (following by a 16 s delay). This indicateded that the auditory-hand mapping would now be reversed for the remaining 10 event-related trials (i.e., “Compty” would now cue a right hand movement whereas “Midwig” would now cue a left hand movement). The sequential ordering of this auditory-hand mapping was counterbalanced across runs, and resulted in a total of 4 different auditory-hand mappings (and thus, trial types) per experimental run: Compty-left hand, Compty-right hand, Midwig-left hand, and Midwig-right hand (with 5 repetitions each; 20 trials in total per run). With the exception of the blocked nature of these trials, these trial types were pseudorandomized within a run and counterbalanced across all runs so that each trial type was preceded and followed equally often by every other trial type across the entire experiment.

Separate practice sessions were carried out before the actual fMRI experiment to familiarize subjects with the delayed timing of the task. One of these sessions was conducted before subjects entered the scanner (See *Behavioural Control Experiment* below) and another was conducted during the anatomical scan (collected at the beginning of the experimental testing session). The experimental testing session for each subject lasted approximately 2 hours and included set-up time, one 7.5 minute high-resolution anatomical scan (during which subjects could practice the task) and eight functional scanning runs (for a total of 160 trials; 40 trials for each auditory-motor mapping). Each functional run (an example run shown in Fig. 1B) had a duration of 576 s, with a 30-60 s break in between each run. Lastly, a resting state functional scan, in which subjects lay still (with no task) and only maintained gaze on the fixation LED, was performed for 12 minutes (*data not analyzed here*).

During MRI testing, we also tracked subjects' behaviour using an MRI-compatible force sensor located beneath the object (Nano 17 F/T sensors; ATI Industrial Automation, Garner, NC), and attached to our MRI platform. This force sensor, which was capped with a flat circular disk (diameter of 7.5 cm), supported the object. The force sensor measured the vertical forces exerted by the object (signals sampled at 500 Hz and low-pass filtered using a 5th order, zero-phase lag Butterworth filter with a cutoff frequency of 5 Hz), allowing us to track both subject reaction time (RT), which we define as the time from the onset of the "Go" cue to object contact (Mean = 1601ms, SD = 389ms), and movement time (MT), which we define as the time from object lift to replacement (Mean = 2582ms, SD = 662ms), as well as generally monitor task performance. Note that we did not conduct eye tracking during this or any of the other MRI scan sessions because of the difficulties in monitoring gaze in the head-tilted configuration with standard MRI-compatible eye trackers (due to occlusion from the eyelids)(Gallivan et al., 2014, 2016, 2019).

Data Acquisition and Analysis

Subjects were scanned using a 3-Tesla Siemens TIM MAGNETOM Trio MRI scanner located at the Centre for Neuroscience Studies, Queen's University (Kingston, Ontario, Canada). Functional MRI volumes were acquired using a T2*-weighted single-shot gradient-echo echo-planar imaging acquisition sequence (time to repetition = 2000 ms, slice thickness = 4 mm, in-plane resolution = 3 mm x 3 mm, time to echo = 30 ms, field of view = 240 mm x 240 mm, matrix size = 80 x 80, flip angle = 90°, and acceleration factor (integrated parallel acquisition technologies, iPAT) = 2 with generalized auto-calibrating partially parallel acquisitions reconstruction). Each volume comprised 35 contiguous (no gap) oblique slices acquired at a ~30° caudal tilt with respect to the plane of the anterior and posterior commissure (AC-PC), providing whole-brain coverage. Subjects were scanned in a head-tilted configuration, allowing direct viewing of the hand workspace. We used a combination of imaging coils to achieve a good signal to noise ratio and to enable direct object workspace viewing without mirrors or occlusion. Specifically, we tilted (~20° degrees) the posterior half of the 12-channel receive-only head coil (6-channels) and suspended a 4-channel receive-only flex coil over the anterior-superior part of the head. A T1-weighted ADNI MPRAGE anatomical scan was also collected (time to repetition = 1760 ms, time to echo = 2.98 ms, field of view = 192 mm x 240 mm x 256 mm, matrix size = 192 x 240 x 256, flip angle = 9°, 1 mm isotropic voxels).

fMRI data preprocessing

Preprocessing of functional data collected in the localizer and main experimental testing sessions was performed using *fMRIPrep* 1.4.1 (Esteban et al., 2018), which is based on *Nipype* 1.2.0 (Gorgolewski et al., 2011; Esteban et al., 2019).

Anatomical data preprocessing

The T1-weighted (T1w) image was corrected for intensity non-uniformity (INU) with N4BiasFieldCorrection (Tustison et al., 2010), distributed with ANTs 2.2.0 (Avants et al., 2008), and used as T1w-reference throughout the workflow. The T1w-reference was then skull-stripped with a Nipype implementation of the antsBrainExtraction.sh workflow (from ANTs), using OASIS30ANTs as target template. Brain tissue segmentation of cerebrospinal fluid (CSF), white-matter (WM) and gray-matter (GM) was performed on the brain-extracted T1w using fast (FSL 5.0.9, (Zhang et al., 2001)). Brain surfaces were reconstructed using recon-all (FreeSurfer 6.0.1, (Dale et al., 1999)), and the brain mask estimated previously was refined with a custom variation of the method to reconcile ANTs-derived and FreeSurfer-derived segmentations of the cortical gray-matter of Mindboggle (Klein et al., 2017). Volume-based spatial normalization to standard space (voxel size = 2 × 2 × 2 mm) was performed through nonlinear registration with antsRegistration (ANTs 2.2.0), using brain-extracted versions of both T1w reference and the T1w template. The following template was selected for spatial normalization: FSL's MNI ICBM 152 non-linear 6th Generation Asymmetric Average Brain Stereotaxic Registration Model [(Evans et al., 2012); TemplateFlow ID: MNI152NLin6Asym].

Functional data preprocessing

For each of the 14 BOLD runs found per subject (across all tasks and sessions), the following preprocessing was performed. First, a reference volume and its skull-stripped version were generated using a custom methodology of fMRIPrep. The BOLD reference was then co-registered to the T1w reference using bbregister (FreeSurfer) which implements boundary-based registration (Greve and Fischl, 2009). Co-registration was configured with nine degrees of freedom to account for distortions remaining in the BOLD reference. Head-motion parameters with respect to the BOLD reference (transformation matrices, and six corresponding rotation and translation parameters) are estimated before any spatiotemporal filtering using mcflirt (FSL 5.0.9, (Jenkinson et al.,

2002)). BOLD runs were slice-time corrected using 3dTshift from AFNI 20160207 (Cox and Hyde, 1997). The BOLD time-series were normalized by resampling into standard space. All resamplings were performed with a single interpolation step by composing all the pertinent transformations (i.e. head-motion transform matrices, and co-registrations to anatomical and output spaces). Gridded (volumetric) resamplings were performed using antsApplyTransforms (ANTs), configured with Lanczos interpolation to minimize the smoothing effects of other kernels (Lanczos, 1964).

Many internal operations of fMRIPrep use Nilearn 0.5.2 (Abraham et al., 2014), mostly within the functional processing workflow. For more details of the pipeline, see the section corresponding to workflows in fMRIPrep's documentation.

Post-processing

Additional post-processing was performed for specific analyses. Normalized functional scans were temporally filtered using a high-pass filter (cutoff = 0.01 Hz) to remove low-frequency noise (e.g. linear scanner drift), either as part of GLMs (see below) or directly (as in time-point decoding analyses). For the localizer data, normalized functional scans were spatially smoothed (6mm FWHM Gaussian kernel; SPM12) prior to GLM estimation to facilitate subject overlap. [Note that no spatial smoothing was performed on the experimental task data, wherein multi-voxel pattern analyses were performed.]

Error trials

Error trials were identified offline from the videos recorded during the experimental testing session and were excluded from analysis by assigning these trials predictors of no interest. Error trials included those in which the subject performed the incorrect instruction (9 trials, 4 subjects) or contaminated the delay phase data by slightly moving their limb or moving too early (7 trials, 4 subjects). The fact that subjects made so few errors when considering the potentially challenging nature of the task (e.g., having to remember whether "Compty" cued a left hand or right hand movement on the current

trial) speaks to the fact that subjects were fully engaged during experimental testing and very well practiced at the task prior to participating in the experiment.

Statistical Analyses

General Linear Models

For the localizer task analyses, we carried out subject-level analysis using SPM12's first-level analysis toolbox to create general linear models (GLM) for each task (auditory and motor). Each GLM featured condition predictors created from boxcar functions convolved with a double-gamma hemodynamic response function (HRF), which were aligned to the onset of each action/stimulus block with durations dependent on block length (i.e., 10 imaging volumes for both localizer tasks). Temporal derivatives of each predictor and subjects' six motion parameters obtained from motion correction were included as additional regressors. The Baseline/Fixation epochs were excluded from the model; therefore all regression coefficients (betas) were defined relative to the baseline activity during these time points.

In the experimental task, separate GLM models for each individual trial's Delay and Execute epochs were generated in order to extract beta coefficient estimates for decoding analyses (20 trials x 2 epochs x 8 runs = 320 GLMs). The regressor of interest in each model consisted of a boxcar regressor aligned to the start of the epoch, with a duration corresponding to the length of the auditory command cue ('Compty' or 'Midwig') and 'Go' cue (both 0.5 s) at the start of the Delay and Execute epochs, respectively. We included a second regressor that comprised of all remaining trial epochs in the experimental run. Each regressor was then convolved with a double-gamma HRF, and temporal derivatives of both regressors were included along with subjects' six motion parameters obtained from motion correction. Isolating the regressor of interest in this single-trial fashion reduces regressor collinearity, and has been shown to be advantageous in estimating single-trial voxel patterns and for multi-voxel pattern classification (Mumford et al., 2012).

Region of interest (ROI) selection

Regions of interests (ROI) were identified based on second-level (group) analyses of first-level contrast images from each subject. Early auditory cortex ROIs were identified by thresholding a Scrambled Speech > Rest group contrast at an uncorrected voxelwise threshold of $p < 10^{-5}$. This procedure identified tight superior temporal gyrus (STG) activation clusters in left and right Heschl's gyrus (HG), the anatomical landmark for primary (core) auditory cortex (Morosan et al., 2000, 2001; Rademacher et al., 2001; Da Costa et al., 2011), and more posteriorly on the superior temporal plane (Planum Temporale, PT). We verified these locations by intersecting region masks for HG and PT obtained from the Harvard-Oxford anatomical atlas (Desikan et al., 2006) with the masks of left and right STG clusters. This allowed us to define, for each participant, voxels that were active for sound that fell in anatomically defined HG and PT. We considered HG and PT separately since they are at different stages of auditory processing: HG is the location of the core, whereas the PT consists of belt and probably parabelt regions, as well as possibly other types of cortical tissue (Hackett et al., 2014). Since our PT activity is just posterior to HG, we suspect that this is probably in belt or parabelt cortex, one or two stages of processing removed from core. Lastly, a more expansive auditory and speech processing network was obtained using a Intact Speech > Rest contrast with an uncorrected height threshold of $p < .001$ and cluster-extent correction threshold of $p < .05$. Together, these were used as three-dimensional binary masks to constrain our analyses and interpretations of hand-related effects to the auditory system.

Multi-voxel Pattern Analysis (MVPA)

For the experimental task, MVPA was performed with in-house software using Python 3.7.1 with Nilearn v0.6.0 and Scikit-Learn v0.20.1 (Abraham et al., 2014). All analyses implement support vector machine (SVM) binary classifiers (libSVM) using a radial-basis function kernel and with fixed regulation parameter (C = 1) and gamma

parameter (automatically set to the reciprocal of the number of voxels) in order to compute a hyperplane that best separated the trial responses.

Inputs for the pattern classifiers were extracted in two complementary approaches. The first approach used the pattern of voxel beta coefficients from the single-trial GLMs, which provided voxel patterns for each trial's Delay and Execute epochs. The second approach used the pattern of percent signal change values for each time point in the trial (i.e., time-resolved decoding) with respect to the time-course of a run-based averaged baseline value (-1, the imaging volume prior to the start of each trial), for all voxels in the ROI (see Gallivan et al., 2013b). Following the extraction of each trial's beta coefficients and/or time point voxel pattern, these values were standardized across voxels such that each voxel pattern had a mean of 0 and standard deviation of 1. Therefore, the mean univariate signal for each pattern was removed in the ROI. In order to derive main-effects of hand information (i.e., examine decoding of left hand vs. right hand movements) versus auditory cue information (examine decoding of "Compty" vs. "Midwig" cues) and to increase statistical power, we performed separate analyses wherein we collapsed across auditory cue or hand, respectively. Our logic is that, when collapsing across auditory cue (i.e., re-labelling all trials based on the hand used), if we can observe decoding of hand information in auditory cortex during the Delay phase (prior to movement), then this information is represented with invariance to the cue, and thus sensory input (and vice versa).

For both decoding approaches, decoding accuracies for each subject were computed as the average classification accuracy across train-and-test iterations using a 'leave-one-run-out' cross-validation procedure. This procedure was performed separately for each ROI, trial epoch/timepoint (Delay and Execute in beta coefficient decoding, each timepoint in time-resolved decoding), and pairwise discrimination (left hand vs right hand movements, "Compty" vs "Midwig"). We statistically assessed decoding significance at the group-level using one-tailed t-tests vs. 50% chance

decoding. To control for the problem of multiple comparisons within each ROI, we applied a Benjamini-Hochberg false-discovery rate (FDR) correction of $q<0.05$. Note that, for the time-resolved decoding approach, the data being used for classification at any single time point (i.e., each TR) are independent, as they are full trial-lengths removed from directly adjacent trials (recall that each trial is, at minimum, equal to 24s), providing more than adequate time for the hemodynamic responses associated with individual TRs to sufficiently uncouple. Furthermore, the trial orders were fully randomized, and so any possible correlations between train and test data is not obvious and should not bias the data towards correct vs. incorrect classification (Misaki et al., 2010; Abraham et al., 2014).

Searchlight Pattern-Information Analyses

To complement our MVPA ROI analyses in the experimental task, we also performed a pattern analysis in each subject using the searchlight approach (Kriegeskorte et al., 2006). Given the scope of this paper (i.e., to examine the top-down modulation of auditory cortex during planning), we constrained this searchlight analysis to the auditory network mask defined by the Intact Speech > Rest contrast using the independent auditory localizer data. In this procedure, the SVM classifier moved through each subjects' localizer-defined auditory network in a voxel-by-voxel fashion whereby, at each voxel, a sphere of surrounding voxels (radius of 4mm; 33 voxels) were extracted, z-scored within pattern (see above), and input into the SVM classifier. The decoding accuracy for that sphere of voxels was then written to the central voxel. This searchlight procedure was performed separately with beta coefficient maps for the Delay and Execute epochs based on the GLM procedure described above, which yielded separate Delay and Execute whole-brain decoding maps. To allow for group-level analyses, the decoding maps were smoothed (6mm FWHM Gaussian kernel) in each subject. Then, for each voxel, we assessed statistical significance using a one-tailed t-test versus 50% chance decoding. Group-level decoding maps for Delay and Execute epochs were

thresholded at $p < .01$ and cluster corrected to $p < .05$ using Monte-Carlo style permutation tests with AFNI's 3dClustSim algorithm (Cox, 1996; Cox et al., 2017).

Behavioural Control Experiment

All subjects participated in a behavioural testing session (performed outside the MRI scanner and before the experimental task) in which their eye fixations and forces corresponding to object grasping and lifting were measured as participants completed the experimental task. This testing session was used for participant screening and to determine, from an analysis of their object lifting and eye-movement behaviour, whether participants were, respectively, (1) maintaining in working memory the instructed hand information over the delay period of each event-related trial and, (2) able to reliably maintain fixation over the duration of an fMRI experimental run (thereby arguing against standard, alternative 'eye-movement confound' interpretations of the fMRI data). In this behavioural testing session, each participant completed 3 experimental runs, identical to those performed in the MRI scanner during the experimental testing session.

The experiment apparatus, equipment and setup was identical to that used in the MRI scanner, but testing was instead performed in the behavioural laboratory. Prior to beginning the behavioural experiment, participants received both verbal instructions and a demonstration by the experimenter as to how to correctly perform the object-directed actions. [Note that force measurements in this behavioural testing session were primarily taken only to provide additional confirmation that participants were capable of performing the task correctly.]

During this behavioural testing, an infrared video-based eye-tracking system (ETL 500 pupil/corneal tracking system, ISCAN Inc. Burlington, MA, USA), mounted below a headband, recorded the gaze position of the left eye at 240 Hz as the participant maintained gaze on the fixation LED. Gaze was calibrated using a two-step procedure: an initial 5-point calibration using ISCAN's Line-of-Sight Plane Intersection Software

followed by a 25-point calibration routine. Twenty-five calibration points (4 mm-diameter circles) were shown on a cardboard frame, presented at the distance of the fixation point, and distributed over a region that incorporated the fixation point, the hand start location, and the location of the central object position. The ISCAN calibration converted raw gaze signals into pixels from the line-of-sight camera and the 25-point calibration converted pixels (i.e., the output of the ISCAN calibration) into the coordinates of hand workspace. Gaze was calibrated at the start of the experimental run and was checked following each block of trials so that, if necessary, gaze could be re-calibrated before starting a new test block.

RESULTS

Delay period decoding of hand information from early auditory cortex

To determine whether signals related to hand movement planning influence early auditory cortex activity, we extracted the trial-related voxel patterns (beta coefficients) associated with the Delay and Execute epochs from early auditory cortex. To this end, we first functionally identified, using the data from an independent auditory localizer task, fMRI activity in the left and right superior temporal gyrus (STG). To provide greater specificity with regards to the localization of potential motor planning-related effects, we further delineated this STG cluster based on its intersections with Heschl's gyrus (HG) and the Planum Temporale (PT), two adjacent human brain regions associated with primary and higher-order cortical auditory processing, respectively (Poeppel et al., 2012)(see Fig. 2A,B for our basic approach). Next, for each of these 3 regions (STG, and its subdivisions into HG and PT) we used their z-scored voxel activity patterns, labelled according to hand (left hand vs. right hand), as inputs to an support vector machine (SVM) binary classifier. This z-scoring procedure, importantly, removed the mean signal from the voxel patterns, which we determined to be an important consideration given that some hand-related and cue-related univariate effects were observed in the data (see Fig. 2D and Fig. 2-2).

Our analysis on the resulting classification accuracies (see Fig. 2C) revealed that information related to the upcoming hand actions to be performed (i.e., during the Delay epoch) was present in bilateral STG (left: t_{14} = 3.55, p = .002; right: t_{14} = 2.34, p = .017) and left HG (t_{14} = 2.43, p = .014). A significant effect was also found in right HG but it did not survive FDR correction (t_{14} = 2.06, p = .029). Meanwhile, no significant decoding was found in left (t_{14} = -.074, p = .529) or right (t_{14} = 1.17, p = .131) PT. During the Execute epoch, however, we found that hand decoding was robust in all three areas in both hemispheres (all p < .001). Because our task did not pair the hand movements to sound generation, and subjects would not have heard the auditory consequences associated with movement (e.g., object lifting and replacement) due to the loud background noise of the scanner, these results suggest that the modulation of auditory cortex activity is automatic and motor-related in nature (Schneider et al., 2014a).

In contrast to our motor-related hand decoding results, our analysis on the resulting classification accuracies for the sensory-related auditory cue ("Compty" vs. "Midwig") revealed that, during the Delay epoch, information related to the delivered verbal cue was present in right STG (t_{14} = 3.71, p = .001). Left PT also showed significant decoding (t_{14} = 1.79, p = .048), although this did not survive FDR correction. No cue decoding was found in the remaining ROIs (all p > .10). Critically, consistent with the fact that this auditory cue information was presented only during the Delay epoch (i.e., participants always received a "Go" cue at the Execute epoch, regardless of trial identity), we also observed no evidence of cue decoding during the Execute epoch (all p > .10). Taken together, this pattern of results suggests that, unlike the representation of auditory cue information, the representation of hand-related information during the Delay epoch is: (1) present across both hemispheres, and (2) present in core (HG) auditory cortex. An additional behavioural control experiment, performed prior to MRI testing, suggests that the emergence of these hand-related effects are unlikely to be driven by systematic differences in eye position across trials (Werner-Reiss et al., 2003), since our trained participants exhibited highly stable fixations throughout the task (Fig. 1-1).

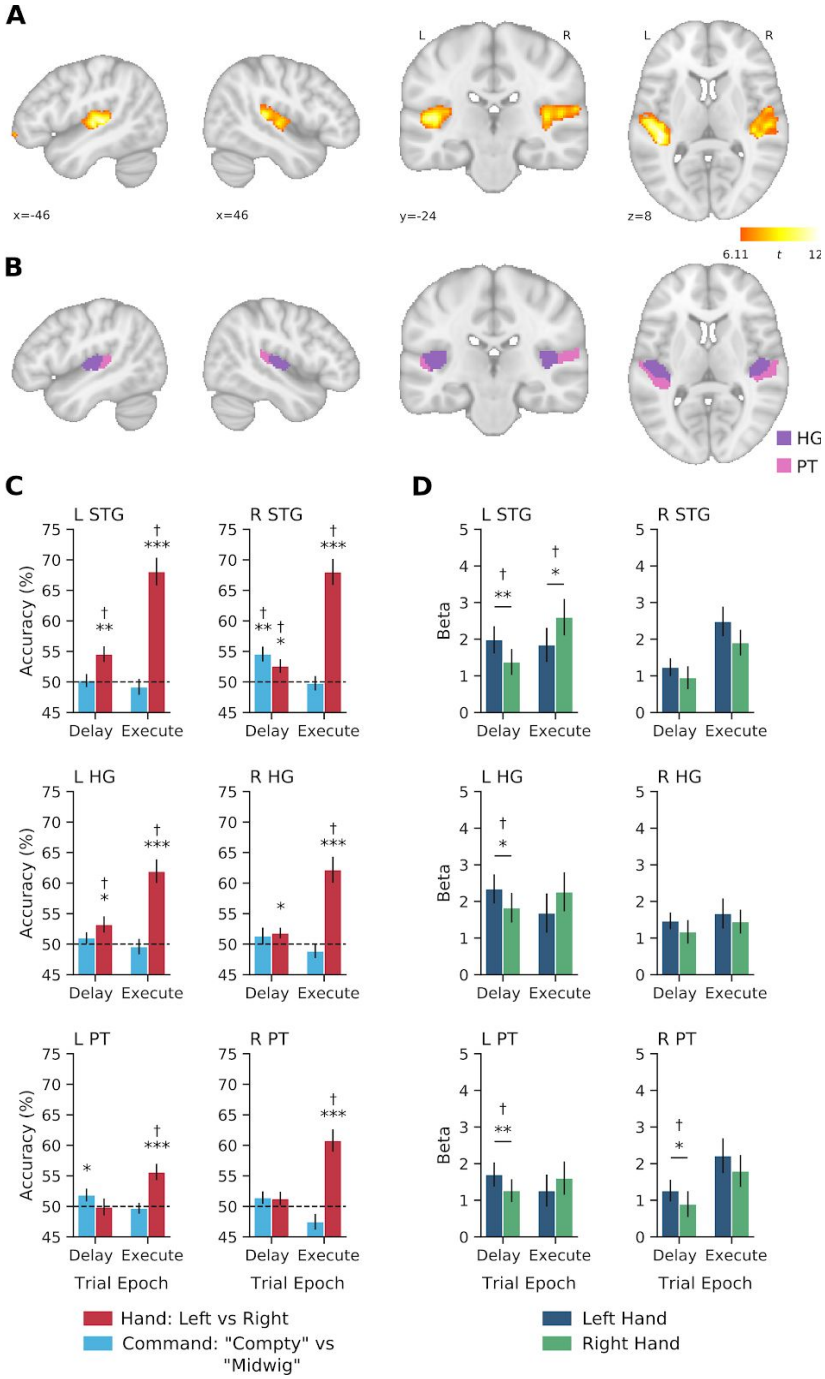

Figure 2. Decoding of hand information in early auditory cortex. **A.** Left and right STG clusters defined at the group-level (N = 16) with the independent auditory localizer task using the contrast of Scrambled Speech > Rest. **B.** Delineation of Heschl's gyrus (HG) and Planum Temporale (PT), within each STG cluster from A. **C.** Decoding accuracies for hand (red) and cue (blue) . Hand and cue decoding accuracies were analyzed separately in each epoch using one-sample t-tests (one-tailed) against chance level (50%). **D.** Mean beta coefficients derived from averaging trialwise betas for left hand and right hand trials. Magnitudes of left hand and right hand trials were compared using paired t-tests at each epoch. Error bars show ± 1 SE of mean. * p < .05, ** p < .01, *** p < .001, † FDR corrected q < .05.

To provide a basis for comparing and interpreting these above hand-related decoding effects in auditory cortex, we also used the data from our experimental task to examine Delay epoch decoding in a positive control region, the dorsal premotor cortex (PMd). This region is well known to be involved in limb-related movement planning in both humans and nonhuman primates (Cisek et al., 2003; Gallivan et al., 2013a) and we independently identified PMd using our motor localizer data. As shown in Fig. 2-1C, we found a remarkably similar profile of limb-specific decoding in this motor-related region to that observed in STG. In fact, direct comparisons between the magnitude of hand-related decoding in STG and PMd were non-significant (see Fig. 2-1D). Taken together, this PMd-result allows for two important observations. First, similar levels of action-related information can be decoded from early auditory cortex as from dorsal premotor cortex, the latter area known to have a well-established role in motor planning (Weinrich et al., 1984; Kaufman et al., 2010; Lara et al., 2018). Second, this Delay epoch-based decoding suggests that the representation of hand-related information evolves prior to movement onset in both STG and PMd.

Time-resolved decoding reveals that hand information in auditory cortex emerges just prior to movement onset

One possible alternative interpretation of our above finding showing the decoding of hand information during the Delay epoch is that, rather than it reflecting a top-down, motor-related modulation, it instead reflects the fact that participants, in our task, must map the auditory command onto a hand instruction. Specifically, upon receiving an auditory cue (e.g., “Compty”), it is plausible that participants immediately translate that cue into the actual instructed action (i.e., think to themselves that “Compty” means “Left Hand”). If so, then it could be the case that it is this auditory conversion or sensory transformation process that is being decoded in auditory cortex, and being interpreted here as the representation of motor effector-specific signals. If this were true, we would then predict that the decoding of hand information should emerge immediately following presentation of the auditory cue (i.e., under the assumption that participants would directly transform the nonsense auditory cue “Compty” into the auditory cue “Left Hand”,

for example). By contrast, if the decoding of hand information is linked to the top-down influence of the motor system, then we would instead predict that the emergence of hand-related information should occur immediately prior to, and at, movement onset. However, because our above pattern decoding analyses were based on beta coefficient measures (reflecting the correspondence between the Delay epoch BOLD time-course signal and a convolved HRF), this approach lacks the necessary time resolution to disentangle these different interpretations. Thus, to directly test this potential alternative explanation of our results, we performed single time-point decoding throughout the trial in our main STG clusters, thereby allowing us to assess not only if, but also when, cue and hand information arises during the Delay epoch (for similar approaches see (Soon et al., 2008; Harrison and Tong, 2009; Gallivan et al., 2013b)).

This single-time point decoding approach in the STG revealed that hand information was decoded in the final imaging volume of the Delay epoch and/or immediately at the onset of the “Go” cue (see Fig. 3). Notably, this was despite the fact that the % BOLD time course signal in STG at those time points was at near-baseline levels. This time course of decoding, as argued above, suggests that the modulation is likely linked to a top-down motor-related influence rather than some simple auditory conversion process. This timing of this hand-related decoding is also noteworthy considering that the earliest latencies we expect to reliably observe an event-induced BOLD effect are at least 2 seconds following that event (Lin et al., 2013); thus, this means that decoding at the final imaging volume of the Delay epoch and immediately at the onset of the “Go” cue likely reflects neural events preceding it by at least 2 seconds. Given that participant reaction times in the MRI scanner were fairly sluggish (Mean = 1601ms, SD = 389ms), it is also unlikely that our hand decoding effects solely reflect some type of pure sensory attenuation effect, as these have been shown to occur in auditory cortical neurons only ~200 ms prior to movement onset (Schneider et al., 2014a). Following the onset of the “Go” cue, we found that decoding accuracies subsequently rose significantly during the

Execute epoch, with the time point associated with the peak decoding accuracy generally corresponding to the peak % BOLD response in the time course signal.

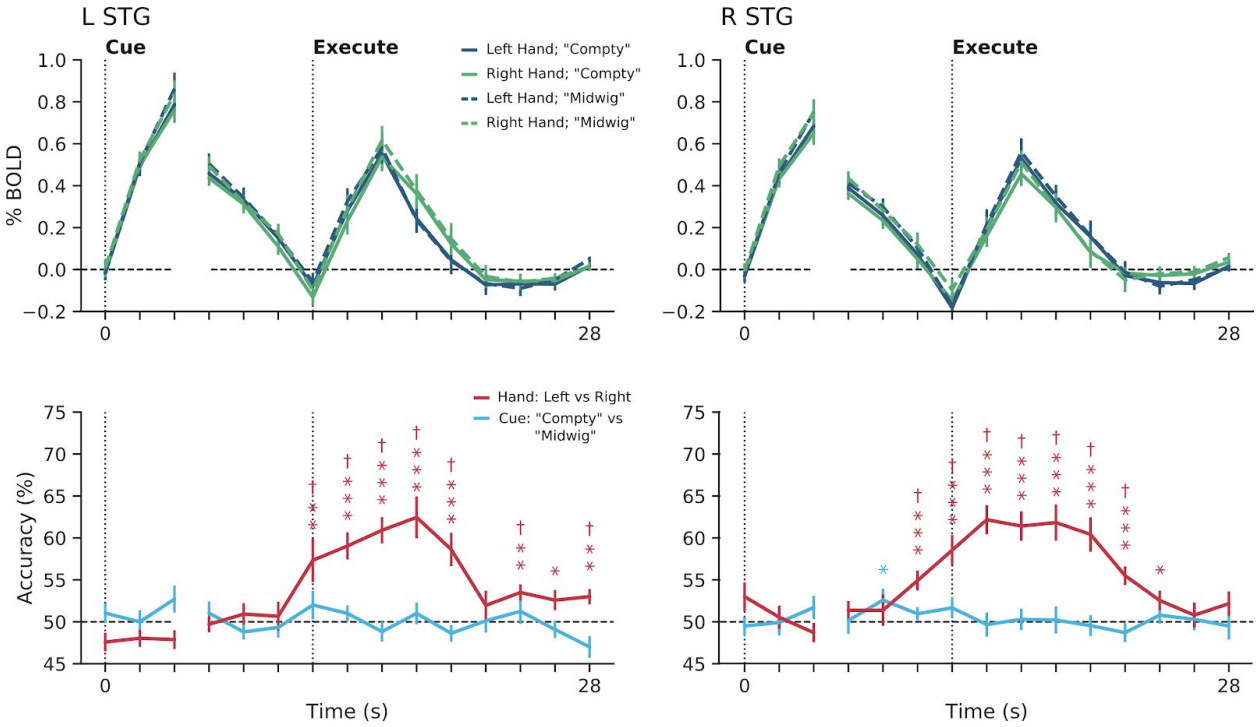

Figure 3. Time-resolved decoding analyses reveal the presence of hand-related information in superior temporal gyrus at time points just prior to movement onset and at the delivery of the "Go" cue. Top, group-averaged percent signal change values of each condition. Bottom, group-level decoding accuracies obtained by separately analyzing hand and auditory cue decoding at each timepoint. Each time point was tested against chance accuracy (50%, horizontal dashed line) using one-sample t-test (one-tailed). The first three time points are time-locked to the onset of the Delay epoch (in which subjects hear the auditory cue), and the remaining 12 time points are time-locked to the onset of the Execute epoch. Error bars show ± 1 SE of mean. Significance markers are colored according to factor (red, hand; blue, cue), * $p < .05$, ** $p < .01$, *** $p < .001$, † FDR corrected $q < .05$.

Searchlight analyses reveal the representation of hand information in early auditory cortex during planning

To complement our above ROI analyses, we also performed a group-level searchlight analysis within the wider auditory processing network, localized using our independent auditory localizer data (see Methods, Fig. 4-1). During the Delay epoch (see Fig. 4), two hand-related decoding clusters were identified in left STG, which includes a cluster

centered on HG (212 voxels; peak, x=-48, y=-24, z=10, t_{14}=8.31) and a cluster spanning anterolateral portions of left STG (321 voxels; peak, x = -60, y = -2, z = -2, t_{14} = 5.72). In the right hemisphere, one large cluster was revealed, which broadly spans across STG and superior temporal sulcus (519 voxels; peak, x = -50, y = -22, z = 8, t_{14} = 6.06). Notably, when we examined cue-related decoding during the Delay epoch (i.e., decoding the auditory command 'Compty' vs. 'Midwig'), we found one cluster in right STG (272 voxels; peak, x = 58, y = -6, z = 8, t_{14} = 8.27), which did not overlap with the hand decoding clusters. This suggests that separate subregions of the auditory system are modulated by motor-related (hand) versus sensory-related (auditory cue) information, and further bolsters our motor-related interpretations of the time-resolved decoding analyses presented in Fig. 3. In addition, the overlap of hand decoding clusters on bilateral HG and STG, as well as a cue decoding cluster in right STG, replicate our basic pattern of ROI-based results presented in Fig. 2.

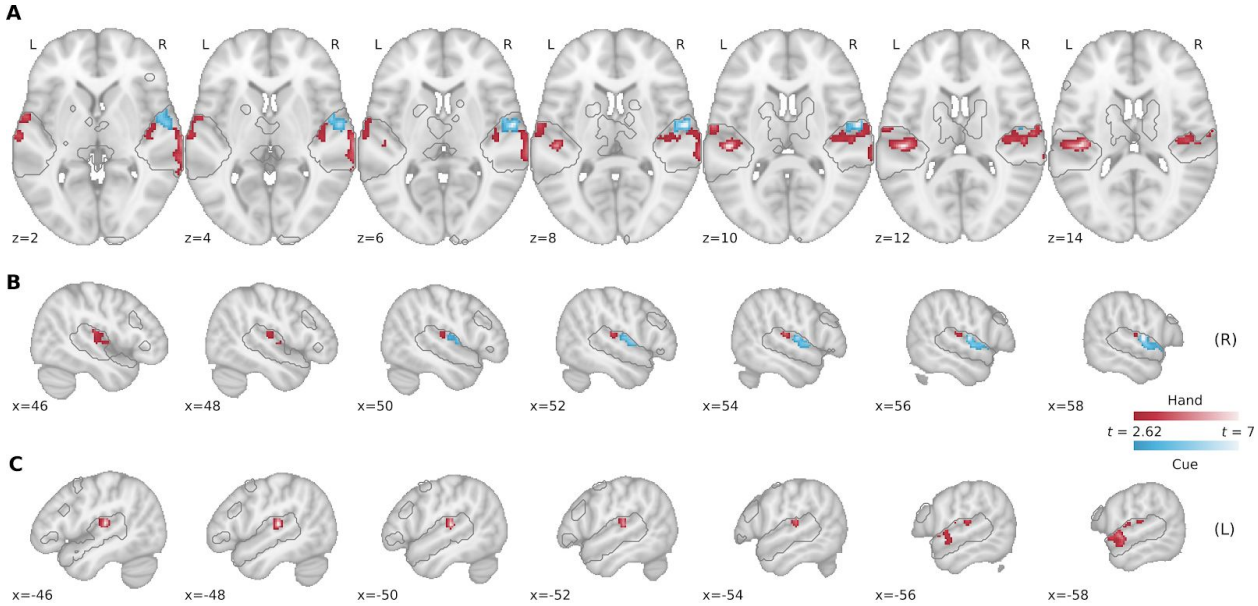

Figure 4. The representation of motor-related (hand) information during the Delay epoch is spatially distinct from the representation of sensory-related (auditory cue) information within superior temporal gyrus. Searchlight analyses were restricted to a mask defined by significant voxels in a Intact Speech > Rest contrast using the independent auditory localizer task (gray traced regions; see Methods). Group-level searchlight maps were thresholded at t > 2.62 (one-tailed p < .01) and

cluster-corrected at $p < .05$. **A.** Transverse slices of significant decoding clusters for hand (red) and auditory cue (blue) in early auditory areas during the Delay epoch. **B.** Sagittal slices of the right hemisphere. **C.** Sagittal slices of left hemisphere.

A searchlight analysis using the Execute epoch data revealed a far more extensive pattern of hand decoding throughout the auditory network, with significant trial classification extending all along the superior and middle temporal gyri bilaterally, and even into the basal ganglia and medial temporal cortex (see Fig. 4-1B). By contrast, and in line with our ROI-based results, no cue decoding clusters were detected during the Execute epoch. These searchlight findings, when considered jointly with our ROI-based results, provide strong supporting evidence that movement planning selectively modulates neural activity patterns in early auditory cortex.

DISCUSSION

Here we have shown, using fMRI and a delayed object lifting task, that the hand used for lifting can be decoded from pre-movement (i.e., delay period) neural activity patterns in early auditory cortex. Importantly, this decoding was invariant to the auditory cue used to instruct the participant on which hand to use. In addition, with our searchlight analyses, we found that the decoding of this hand-related motor information occured in a separate and earlier subregion of auditory cortex than the decoding of the auditory cue that instructed the motor action. Together, these findings suggest that a critical component of action planning, beyond preparing motor areas for the forthcoming movement, is preparing early sensory areas. Such preparation may enable these areas to more effectively participate in the filtering and processing of task-specific sensory signals that arise during the unfolding movement itself.

Several hypotheses have been proposed about the role of planning-related activity in motor areas, including the primary, premotor, and supplementary motor cortices. Several researchers have suggested that planning activity encodes a variety of different movement parameters (e.g., direction), with a view that it represents some subthreshold

version of the forthcoming movement (Tanji and Evarts, 1976; Riehle and Requin, 1989; Hocherman and Wise, 1991; Shen and Alexander, 1997; Messier and Kalaska, 2000b; Churchland et al., 2006; Pesaran et al., 2006; Batista et al., 2007). Recent work, examining the dynamics of populations of neurons in motor areas, has instead suggested that movement planning involves setting up the initial state of the population, such that movement execution can unfold naturally through transitory neural dynamics (Churchland et al., 2010, 2012; Shenoy et al., 2013b; Sussillo et al., 2015; Pandarinath et al., 2017; Lara et al., 2018). Within this framework, our results suggest that motor planning may also involve preparing the initial state of primary sensory cortical areas. Whereas the neural activity patterns that unfold during movement execution in motor areas are thought to regulate the timing and nature of descending motor commands (Churchland et al., 2012; Shenoy et al., 2013a), such activity in primary sensory areas may instead regulate the timing and nature of the filtering of incoming sensory signals.

Motor planning signals could, in principle, tune early sensory areas for participation in sensory prediction. Prediction of the sensory consequences of movement is essential for the accurate sensorimotor control of movement, per se, and also provides a mechanism for distinguishing between self-generated and externally generated sensory information (Wolpert and Flanagan, 2001). The critical role of prediction in sensorimotor control has been well documented in the context of object manipulation tasks (Flanagan et al., 2006; Johansson and Flanagan, 2009). The control of such tasks centers around contact events, which give rise to discrete sensory signals in multiple modalities (e.g., auditory, tactile) and represent subgoals of the overall task. Thus, in the grasp, lift, and replace task that our participants performed, the brain predicts the timing and nature of discrete sensory signals associated with contact between the digits and object, as well as the breaking, and subsequent making, of contact between the object and surface; events that signify successful object grasp, lift-off, and replacement, respectively. By comparing predicted and actual sensory signals associated with these events, the brain can monitor task progression and launch rapid corrective actions if mismatches occur

(Wolpert et al., 2011). These corrective actions are themselves quite sophisticated and depend on both the phase of the task and the nature of the mismatch (Flanagan et al., 2006). Thus, the planning of manipulation tasks clearly involves forming what could be referred to as a ‘sensory plan’; i.e., a series of sensory events—linked to contact events—that, during subsequent movement execution, can be predicted based on knowledge of object properties and information related to outgoing motor commands (Johansson and Flanagan, 2009).

The disambiguation of self- and externally generated sensory information is thought to rely on cancelling, or attenuating, the predictable sensory consequences of movements (Wolpert and Flanagan, 2001). Such 'sensory cancellation’ has been investigated in the context of tasks involving manual interactions with objects. It has been shown, for example, that when participants use one index finger to tap a disk mounted above the other index finger, the perceived force—acting on the lower finger—is attenuated relative to when the same tap is delivered by an electric motor (Shergill et al., 2003). Importantly, this attenuation is not linked to movement, *per se*, but is temporally tuned to the timing of the predicted contact event (Bays et al., 2005, 2006).

Sensory predictions for use in *sensory cancellation* are generally thought to be represented in primary sensory areas. According to this view, an efference copy of descending motor commands, associated with movement execution, is transmitted in a top-down fashion to early sensory cortices in order to attenuate self-generated sensory information (Holst et al., 1950; Crapse and Sommer, 2008). In contrast, sensory predictions, for use in *sensorimotor control*, are thought to be represented in the same frontoparietal circuits involved in movement planning and control (Scott, 2012, 2016). According to this view, incoming sensory information, associated with movement execution, is transmitted in a bottom-up fashion from early sensory areas to frontoparietal circuits for mismatch detection and movement correction (Desmurget et al., 1999; Tunik et al., 2005; Jenmalm et al., 2006). However, our understanding of

sensory prediction in the context of sensorimotor control remains quite limited, and whether it involves the top-down modulation of sensory areas has remained an open question.

Critically, the modulation of early auditory cortex we observed occurred well in advance of movement execution, and is thus highly unlikely to solely reflect sensory cancellation or on-line sensorimotor control. Indeed, sensory attenuation responses in primary auditory cortex have been shown to occur about two hundred milliseconds prior to movement onset (Schneider et al., 2014a, 2018), whereas the hand-specific modulation of auditory cortex we report occurred, at minimum, several *seconds* prior to movement onset. Thus, one possible explanation for the pre-movement modulation of auditory cortex is that it arises from the motor system preparing the state of auditory cortex to process auditory inputs in a task- and phase-dependent manner. One line of support for this notion comes from the fact that, in the scanner environment, our participants could not have heard the auditory consequences of their actions; e.g., sounds associated with contacting, lifting, and replacing the object. This argues that our reported pre-movement modulation of auditory cortex has a motor-related origin, and is not linked to sensory reafference or attentional processes (see Otazu et al., 2009; Schneider et al., 2014b). Moreover, it suggests that this modulation arises from automatic processes that occur during movement planning. A second line of support for this notion comes from the fact that our searchlight analyses identified separate clusters in STG that decoded hand information (in early auditory cortex) from those that decoded auditory cue information (in higher order auditory cortex). This suggests that top-down signals from the motor system modulate different subregions of auditory cortex than those that discriminate sensory information.

Prior work has also demonstrated that tactile input alone is capable of driving auditory cortex activity (Foxe et al., 2002; Kayser et al., 2005; Schürmann et al., 2006; Lakatos et al., 2007), indicating a potential role for auditory cortex in multisensory integration. As

noted above, the control of object manipulation tasks involves accurately predicting discrete sensory events that arise in multiple modalities, including tactile and auditory (Johansson and Flanagan, 2009). It is plausible then that some portion of the pre-movement auditory cortex modulation described here reflects the predicted tactile events arising from our task (e.g., object contact, lift-off and replacement), which we would also expect to be linked to the acting hand. Though we cannot disentangle this possibility in the current study, it is a direction for future work and does not undercut our main observation that early sensory cortex is modulated as a function of the movement being prepared; nor does it undercut our interpretation—that such modulation is likely linked to sensorimotor prediction.

In summary, here we show that, prior to movement, neural activity patterns in early auditory cortex carry information about the hand to be used in the upcoming action. This result supports the hypothesis that 'motor' planning, which is critical in preparing neural states ahead of movement execution (Lara et al., 2018), not only occurs in motor areas but also in primary sensory areas. Further work is required to establish the precise role of this movement-related modulation. Our findings add to a growing line of evidence indicating that early sensory systems are directly modulated by sensorimotor computations performed in higher-order cortex (Chapman et al., 2011; Steinmetz and Moore, 2014; Gutteling et al., 2015; Gallivan et al., 2019) and not merely low-level relayers of incoming sensory information (Scheich et al., 2007; Matyas et al., 2010; Weinberger, 2011; Huang et al., 2019).

Acknowledgements

This work was supported by operating grants from the Canadian Institutes of Health Research (CIHR) awarded to J.R.F. and J.P.G. (MOP126158). J.P.G. was supported by a Natural Sciences and Engineering Research Council (NSERC) Discovery Grant, as well as funding from the Canadian Foundation for Innovation. D.J.G. was supported by a R.S. McLaughlin Fellowship and an NSERC graduate fellowship; C.N.A. was supported by an NSERC graduate fellowship. The authors would like to thank Adam McLean, Haitao Yang, Derek Quinlan, Martin York, Sean Hickman, and Don O'Brien for technical assistance.

Author Contributions

J.R.F., I.S.J and J.P.G. designed the experiment. D.J.G. and J.P.G. performed research. D.J.G., C.N.A., C.H., J.R.F and J.P.G., analyzed and interpreted data. D.J.G., J.R.F and J.P.G wrote the paper. All authors provided edits and feedback on the final version of the paper.

Competing Interests Statement

The authors declare no competing financial interests.

REFERENCES

Abraham A, Pedregosa F, Eickenberg M, Gervais P, Mueller A, Kossaifi J, Gramfort A, Thirion B, Varoquaux G (2014) Machine learning for neuroimaging with scikit-learn. Front Neuroinform 8:14.

Avants BB, Epstein CL, Grossman M, Gee JC (2008) Symmetric diffeomorphic image registration with cross-correlation: evaluating automated labeling of elderly and neurodegenerative brain. Med Image Anal 12:26–41.

Batista AP, Santhanam G, Yu BM, Ryu SI, Afshar A, Shenoy KV (2007) Reference Frames for Reach Planning in Macaque Dorsal Premotor Cortex. Journal of Neurophysiology 98:966–983 Available at: http://dx.doi.org/10.1152/jn.00421.2006.

Bays PM, Flanagan JR, Wolpert DM (2006) Attenuation of self-generated tactile sensations is predictive, not postdictive. PLoS Biol 4:e28.

Bays PM, Wolpert DM, Flanagan JR (2005) Perception of the consequences of self-action is temporally tuned and event driven. Curr Biol 15:1125–1128.

Chapman CS, Gallivan JP, Culham JC, Goodale MA (2011) Mental blocks: fMRI reveals top-down modulation of early visual cortex when obstacles interfere with grasp planning. Neuropsychologia 49:1703–1717.

Churchland MM, Cunningham JP, Kaufman MT, Foster JD, Nuyujukian P, Ryu SI, Shenoy KV (2012) Neural population dynamics during reaching. Nature 487:51–56.

Churchland MM, Cunningham JP, Kaufman MT, Ryu SI, Shenoy KV (2010) Cortical preparatory activity: representation of movement or first cog in a dynamical machine? Neuron 68:387–400.

Churchland MM, Santhanam G, Shenoy KV (2006) Preparatory activity in premotor and motor cortex reflects the speed of the upcoming reach. J Neurophysiol 96:3130–3146.

Cisek P, Crammond DJ, Kalaska JF (2003) Neural activity in primary motor and dorsal premotor cortex in reaching tasks with the contralateral versus ipsilateral arm. J Neurophysiol 89:922–942.

Cox RW (1996) AFNI: software for analysis and visualization of functional magnetic resonance neuroimages. Comput Biomed Res 29:162–173.

Cox RW, Chen G, Glen DR, Reynolds RC, Taylor PA (2017) FMRI Clustering in AFNI: False-Positive Rates Redux. Brain Connect 7:152–171.

Cox RW, Hyde JS (1997) Software tools for analysis and visualization of fMRI data. NMR Biomed 10:171–178.

Crapse TB, Sommer MA (2008) Corollary discharge across the animal kingdom. Nat Rev Neurosci 9:587–600.

Da Costa S, van der Zwaag W, Marques JP, Frackowiak RSJ, Clarke S, Saenz M (2011) Human primary auditory cortex follows the shape of Heschl's gyrus. J Neurosci 31:14067–14075.

Dale AM, Fischl B, Sereno MI (1999) Cortical surface-based analysis. I. Segmentation and surface reconstruction. Neuroimage 9:179–194.

Desikan RS, Ségonne F, Fischl B, Quinn BT, Dickerson BC, Blacker D, Buckner RL, Dale AM, Maguire RP, Hyman BT, Albert MS, Killiany RJ (2006) An automated labeling system for subdividing the human cerebral cortex on MRI scans into gyral based regions of interest. Neuroimage 31:968–980.

Desmurget M, Epstein CM, Turner RS, Prablanc C, Alexander GE, Grafton ST (1999) Role of the posterior parietal cortex in updating reaching movements to a visual target. Nat Neurosci 2:563–567.

Esteban O et al. (2019) nipy/nipype: 1.4.0. Available at: https://zenodo.org/record/3588470.

Esteban O, Markiewicz CJ, Blair RW, Moodie CA, Ilkay Isik A, Erramuzpe A, Kent JD, Goncalves M, DuPre E, Snyder M, Oya H, Ghosh SS, Wright J, Durnez J, Poldrack RA, Gorgolewski KJ (2018) fMRIPrep: a robust preprocessing pipeline for functional MRI. Nat Methods:1.

Evans AC, Janke AL, Collins DL, Baillet S (2012) Brain templates and atlases. Neuroimage 62:911–922.

Flanagan JR, Bowman MC, Johansson RS (2006) Control strategies in object manipulation tasks. Curr Opin Neurobiol 16:650–659.

Foxe JJ, Wylie GR, Martinez A, Schroeder CE, Javitt DC, Guilfoyle D, Ritter W, Murray MM (2002) Auditory-somatosensory multisensory processing in auditory association cortex: an fMRI study. J Neurophysiol 88:540–543.

Gallivan JP, Cant JS, Goodale MA, Flanagan JR (2014) Representation of object weight in human ventral visual cortex. Curr Biol 24:1866–1873.

Gallivan JP, Chapman CS, Gale DJ, Flanagan JR, Culham JC (2019) Selective Modulation of Early Visual Cortical Activity by Movement Intention. Cereb Cortex Available at: http://dx.doi.org/10.1093/cercor/bhy345.

Gallivan JP, Johnsrude IS, Flanagan JR (2016) Planning Ahead: Object-Directed Sequential Actions Decoded from Human Frontoparietal and Occipitotemporal Networks. Cereb Cortex 26:708–730.

Gallivan JP, McLean DA, Flanagan JR, Culham JC (2013a) Where one hand meets the other: limb-specific and action-dependent movement plans decoded from preparatory signals in single human frontoparietal brain areas. J Neurosci 33:1991–2008.

Gallivan JP, McLean DA, Valyear KF, Culham JC (2013b) Decoding the neural mechanisms of human tool use. Elife 2:e00425.

Gorgolewski K, Burns CD, Madison C, Clark D, Halchenko YO, Waskom ML, Ghosh SS (2011) Nipype: a flexible, lightweight and extensible neuroimaging data processing framework in python. Front Neuroinform 5:13.

Greve DN, Fischl B (2009) Accurate and robust brain image alignment using boundary-based registration. Neuroimage 48:63–72.

Gutteling TP, Petridou N, Dumoulin SO, Harvey BM, Aarnoutse EJ, Kenemans JL, Neggers SFW (2015) Action preparation shapes processing in early visual cortex. J Neurosci 35:6472–6480.

Hackett TA, de la Mothe LA, Camalier CR, Falchier A, Lakatos P, Kajikawa Y, Schroeder CE (2014) Feedforward and feedback projections of caudal belt and parabelt areas of auditory cortex: refining the hierarchical model. Front Neurosci 8:72.

Harrison SA, Tong F (2009) Decoding reveals the contents of visual working memory in early visual areas. Nature 458:632–635 Available at: http://dx.doi.org/10.1038/nature07832.

Hocherman S, Wise SP (1991) Effects of hand movement path on motor cortical activity in awake, behaving rhesus monkeys. Exp Brain Res 83:285–302.

Holst E von, von Holst E, Mittelstaedt H (1950) Das Reafferenzprinzip. Naturwissenschaften 37:464–476 Available at: http://dx.doi.org/10.1007/bf00622503.

Huang Y, Heil P, Brosch M (2019) Associations between sounds and actions in early auditory cortex of nonhuman primates. Elife 8 Available at: http://dx.doi.org/10.7554/eLife.43281.

Jenkinson M, Bannister P, Brady M, Smith S (2002) Improved optimization for the robust and accurate linear registration and motion correction of brain images. Neuroimage 17:825–841.

Jenmalm P, Schmitz C, Forssberg H, Ehrsson HH (2006) Lighter or heavier than predicted: neural correlates of corrective mechanisms during erroneously programmed lifts. J Neurosci 26:9015–9021.

Johansson, Flanagan JR (2009) Coding and use of tactile signals from the fingertips in object manipulation tasks. Nat Rev Neurosci 10:345–359.

Kaufman MT, Churchland MM, Santhanam G, Yu BM, Afshar A, Ryu SI, Shenoy KV (2010) Roles of monkey premotor neuron classes in movement preparation and execution. J Neurophysiol 104:799–810.

Kayser C, Petkov CI, Augath M, Logothetis NK (2005) Integration of touch and sound in auditory cortex. Neuron 48:373–384.

Klein A, Ghosh SS, Bao FS, Giard J, Häme Y, Stavsky E, Lee N, Rossa B, Reuter M, Chaibub Neto E, Keshavan A (2017) Mindboggling morphometry of human brains. PLoS Comput Biol 13:e1005350.

Kriegeskorte N, Goebel R, Bandettini P (2006) Information-based functional brain mapping.

Proc Natl Acad Sci U S A 103:3863–3868.

Lakatos P, Chen C-M, O'Connell MN, Mills A, Schroeder CE (2007) Neuronal oscillations and multisensory interaction in primary auditory cortex. Neuron 53:279–292.

Lanczos C (1964) Evaluation of Noisy Data. Journal of the Society for Industrial and Applied Mathematics Series B Numerical Analysis 1:76–85.

Lara, Elsayed GF, Zimnik AJ, Cunningham JP, Churchland MM (2018) Conservation of preparatory neural events in monkey motor cortex regardless of how movement is initiated. Elife 7 Available at: http://dx.doi.org/10.7554/eLife.31826.

Lin F-H, Witzel T, Raij T, Ahveninen J, Tsai KW-K, Chu Y-H, Chang W-T, Nummenmaa A, Polimeni JR, Kuo W-J, Hsieh J-C, Rosen BR, Belliveau JW (2013) fMRI hemodynamics accurately reflects neuronal timing in the human brain measured by MEG. NeuroImage 78:372–384 Available at: http://dx.doi.org/10.1016/j.neuroimage.2013.04.017.

Matyas F, Sreenivasan V, Marbach F, Wacongne C, Barsy B, Mateo C, Aronoff R, Petersen CCH (2010) Motor control by sensory cortex. Science 330:1240–1243.

Messier J, Kalaska JF (2000a) Covariation of primate dorsal premotor cell activity with direction and amplitude during a memorized-delay reaching task. J Neurophysiol 84:152–165.

Messier J, Kalaska JF (2000b) Covariation of primate dorsal premotor cell activity with direction and amplitude during a memorized-delay reaching task. J Neurophysiol 84:152–165.

Misaki M, Kim Y, Bandettini PA, Kriegeskorte N (2010) Comparison of multivariate classifiers and response normalizations for pattern-information fMRI. Neuroimage 53:103–118.

Morosan P, Mohlberg H, Amunts K, Schleicher A, Rademacher J, Schormann T, Zilles K (2000) Population maps of cytoarchitectonically defined human auditory areas. NeuroImage 11:S304 Available at: http://dx.doi.org/10.1016/s1053-8119(00)91236-9.

Morosan P, Rademacher J, Schleicher A, Amunts K, Schormann T, Zilles K (2001) Human primary auditory cortex: cytoarchitectonic subdivisions and mapping into a spatial reference system. Neuroimage 13:684–701.

Mumford JA, Turner BO, Ashby FG, Poldrack RA (2012) Deconvolving BOLD activation in event-related designs for multivoxel pattern classification analyses. Neuroimage 59:2636–2643.

Nelson A, Schneider DM, Takatoh J, Sakurai K, Wang F, Mooney R (2013) A circuit for motor cortical modulation of auditory cortical activity. J Neurosci 33:14342–14353.

Oldfield RC (1971) The assessment and analysis of handedness: the Edinburgh inventory. Neuropsychologia 9:97–113.

Otazu GH, Tai L-H, Yang Y, Zador AM (2009) Engaging in an auditory task suppresses responses in auditory cortex. Nat Neurosci 12:646–654.

Pandarinath C, Nuyujukian P, Blabe CH, Sorice BL, Saab J, Willett FR, Hochberg LR, Shenoy

KV, Henderson JM (2017) High performance communication by people with paralysis using an intracortical brain-computer interface. Elife 6 Available at: http://dx.doi.org/10.7554/eLife.18554.

Pesaran B, Nelson MJ, Andersen RA (2006) Dorsal premotor neurons encode the relative position of the hand, eye, and goal during reach planning. Neuron 51:125–134.

Poeppel D, Overath T, Popper AN, Fay RR (2012) The Human Auditory Cortex. Springer Science & Business Media.

Rademacher J, Morosan P, Schormann T, Schleicher A, Werner C, Freund HJ, Zilles K (2001) Probabilistic mapping and volume measurement of human primary auditory cortex. Neuroimage 13:669–683.

Reznik D, Ossmy O, Mukamel R (2015) Enhanced auditory evoked activity to self-generated sounds is mediated by primary and supplementary motor cortices. J Neurosci 35:2173–2180.

Riehle A, Requin J (1989) Monkey primary motor and premotor cortex: single-cell activity related to prior information about direction and extent of an intended movement. J Neurophysiol 61:534–549.

Safstrom D, Johansson RS, Flanagan JR (2014) Gaze behavior when learning to link sequential action phases in a manual task. Journal of Vision 14:3–3 Available at: http://dx.doi.org/10.1167/14.4.3.

Scheich H, Brechmann A, Brosch M, Budinger E, Ohl FW (2007) The cognitive auditory cortex: task-specificity of stimulus representations. Hear Res 229:213–224.

Schneider DM, Mooney R (2018) How Movement Modulates Hearing. Annual Review of Neuroscience 41:553–572 Available at: http://dx.doi.org/10.1146/annurev-neuro-072116-031215.

Schneider DM, Nelson A, Mooney R (2014a) A synaptic and circuit basis for corollary discharge in the auditory cortex. Nature 513:189–194.

Schneider DM, Nelson A, Mooney R (2014b) A synaptic and circuit basis for corollary discharge in the auditory cortex. Nature 513:189–194 Available at: http://dx.doi.org/10.1038/nature13724.

Schneider DM, Sundararajan J, Mooney R (2018) A cortical filter that learns to suppress the acoustic consequences of movement. Nature 561:391–395.

Schürmann M, Caetano G, Hlushchuk Y, Jousmäki V, Hari R (2006) Touch activates human auditory cortex. Neuroimage 30:1325–1331.

Scott SH (2012) The computational and neural basis of voluntary motor control and planning. Trends Cogn Sci 16:541–549.

Scott SH (2016) A Functional Taxonomy of Bottom-Up Sensory Feedback Processing for Motor

Actions. Trends Neurosci 39:512–526.

Shen L, Alexander GE (1997) Preferential representation of instructed target location versus limb trajectory in dorsal premotor area. J Neurophysiol 77:1195–1212.

Shenoy KV, Sahani M, Churchland MM (2013a) Cortical control of arm movements: a dynamical systems perspective. Annu Rev Neurosci 36:337–359.

Shenoy KV, Sahani M, Churchland MM (2013b) Cortical control of arm movements: a dynamical systems perspective. Annu Rev Neurosci 36:337–359.

Shergill SS, Bays PM, Frith CD, Wolpert DM (2003) Two eyes for an eye: the neuroscience of force escalation. Science 301:187.

Soon CS, Brass M, Heinze H-J, Haynes J-D (2008) Unconscious determinants of free decisions in the human brain. Nature Neuroscience 11:543–545 Available at: http://dx.doi.org/10.1038/nn.2112.

Steinmetz NA, Moore T (2014) Eye Movement Preparation Modulates Neuronal Responses in Area V4 When Dissociated from Attentional Demands. Neuron 83:496–506 Available at: http://dx.doi.org/10.1016/j.neuron.2014.06.014.

Sussillo D, Churchland MM, Kaufman MT, Shenoy KV (2015) A neural network that finds a naturalistic solution for the production of muscle activity. Nat Neurosci 18:1025–1033.

Tanji J, Evarts EV (1976) Anticipatory activity of motor cortex neurons in relation to direction of an intended movement. J Neurophysiol 39:1062–1068.

Tunik E, Frey SH, Grafton ST (2005) Virtual lesions of the anterior intraparietal area disrupt goal-dependent on-line adjustments of grasp. Nat Neurosci 8:505–511.

Tustison NJ, Avants BB, Cook PA, Zheng Y, Egan A, Yushkevich PA, Gee JC (2010) N4ITK: improved N3 bias correction. IEEE Trans Med Imaging 29:1310–1320.

Weinberger NM (2011) Reconceptualizing the Primary Auditory Cortex: Learning, Memory and Specific Plasticity. The Auditory Cortex:465–491 Available at: http://dx.doi.org/10.1007/978-1-4419-0074-6_22.

Weinrich M, Wise SP, Mauritz KH (1984) A neurophysiological study of the premotor cortex in the rhesus monkey. Brain 107 (Pt 2):385–414.

Werner-Reiss U, Kelly KA, Trause AS, Underhill AM, Groh JM (2003) Eye Position Affects Activity in Primary Auditory Cortex of Primates. Current Biology 13:554–562 Available at: http://dx.doi.org/10.1016/s0960-9822(03)00168-4.

Wolpert DM, Diedrichsen J, Flanagan JR (2011) Principles of sensorimotor learning. Nat Rev Neurosci 12:739–751.

Wolpert DM, Flanagan JR (2001) Motor prediction. Current Biology 11:R729–R732 Available at: http://dx.doi.org/10.1016/s0960-9822(01)00432-8.

Zhang Y, Brady M, Smith S (2001) Segmentation of brain MR images through a hidden Markov random field model and the expectation-maximization algorithm. IEEE Trans Med Imaging 20:45–57.