

# The Fusion of Mental Imagery and Sensation in the Temporal Association Cortex

Christopher C. Berger and H. Henrik Ehrsson

Department of Neuroscience, Karolinska Institutet, 171 77 Stockholm, Sweden

It is well understood that the brain integrates information that is provided to our different senses to generate a coherent multisensory percept of the world around us (Stein and Stanford, 2008), but how does the brain handle concurrent sensory information from our mind and the external world? Recent behavioral experiments have found that mental imagery—the internal representation of sensory stimuli in one’s mind—can also lead to integrated multisensory perception (Berger and Ehrsson, 2013); however, the neural mechanisms of this process have not yet been explored. Here, using functional magnetic resonance imaging and an adapted version of a well known multisensory illusion (i.e., the ventriloquist illusion; Howard and Templeton, 1966), we investigated the neural basis of mental imagery-induced multisensory perception in humans. We found that simultaneous visual mental imagery and auditory stimulation led to an illusory translocation of auditory stimuli and was associated with increased activity in the left superior temporal sulcus (L. STS), a key site for the integration of real audiovisual stimuli (Beauchamp et al., 2004a, 2010; Driver and Noesselt, 2008; Ghazanfar et al., 2008; Dahl et al., 2009). This imagery-induced ventriloquist illusion was also associated with increased effective connectivity between the L. STS and the auditory cortex. These findings suggest an important role of the temporal association cortex in integrating imagined visual stimuli with real auditory stimuli, and further suggest that connectivity between the STS and auditory cortex plays a modulatory role in spatially localizing auditory stimuli in the presence of imagined visual stimuli.

**Key words:** auditory cortex; mental imagery; multisensory; superior temporal sulcus; ventriloquism

## Introduction

Imagining something in one’s mind and perceiving something in the external world are phenomenologically similar experiences, and there is mounting evidence that these two experiences are represented similarly in the brain (Kosslyn et al., 2001). For example, it has been found that visual imagery activates the primary visual cortex (Kosslyn et al., 2001; Kamitani and Tong, 2005), and that visual imagery of objects from different categories selectively activates the corresponding parts of the visual cortex associated with perceiving objects from those categories (O’Craven and Kanwisher, 2000; Cichy et al., 2012). Similar findings exist for tactile (Anema et al., 2012), motor (Roth et al., 1996; Ehrsson et al., 2003), and auditory imagery (Bunzeck et al., 2005; Oh et al., 2013) and for the corresponding sensory or motor cortices. However, research investigating the relationship between imagery and perception has focused on similarities and interactions within a given sensory modality, and the possibility of multisensory interactions between imagery and perception has been largely ignored.

Recent evidence favoring this possibility arises from a series of behavioral experiments demonstrating that mental imagery can induce multisensory perceptual illusions (Berger and Ehrsson, 2013). In one experiment, it was observed that imagined visual stimuli alter the perceived location of sounds in the same manner as real visual stimuli in the well known ventriloquist illusion (Berger and Ehrsson, 2013). Although previous neuroimaging studies have linked perceptual multisensory integration to neural activity in the frontal, parietal, and temporal association cortices, and to subcortical structures, such as the superior colliculus and putamen (Calvert et al., 2000; Bushara et al., 2003; Macaluso and Driver, 2005; Bischoff et al., 2007; Stein and Stanford, 2008), it remains unknown whether imagery-related multisensory integration relies on similar neural mechanisms.

To investigate this possibility, we used functional magnetic resonance imaging (fMRI) and an adapted version of a ventriloquism paradigm. Consistent with previous studies that have manipulated the temporal correspondence between audiovisual stimuli to investigate the neural correlates of multisensory integration (Calvert et al., 2000; Ehrsson et al., 2004; Bischoff et al., 2007; Noesselt et al., 2007; Driver and Noesselt, 2008; Marchant et al., 2012; Gentile et al., 2013), we asked participants to vividly imagine the appearance of a white circle in synchrony or asynchrony with a spatially disparate auditory stimulus (Fig. 1; see Materials and Methods). To quantify the ventriloquist effect, the participants reported whether they perceived the sound to come from the left, right, or center after each trial.

Received March 10, 2014; revised Aug. 27, 2014; accepted Sept. 1, 2014.

Author contributions: C.C.B. and H.H.E. designed research; C.C.B. performed research; C.C.B. analyzed data; C.C.B. and H.H.E. wrote the paper.

This work was supported by the European Research Council, the Swedish Foundation for Strategic Research, the James S. McDonnell Foundation, the Swedish Research Council, and Söderbergska Stiftelsen. We thank Giovanni Gentile for practical assistance with parts of the fMRI analyses.

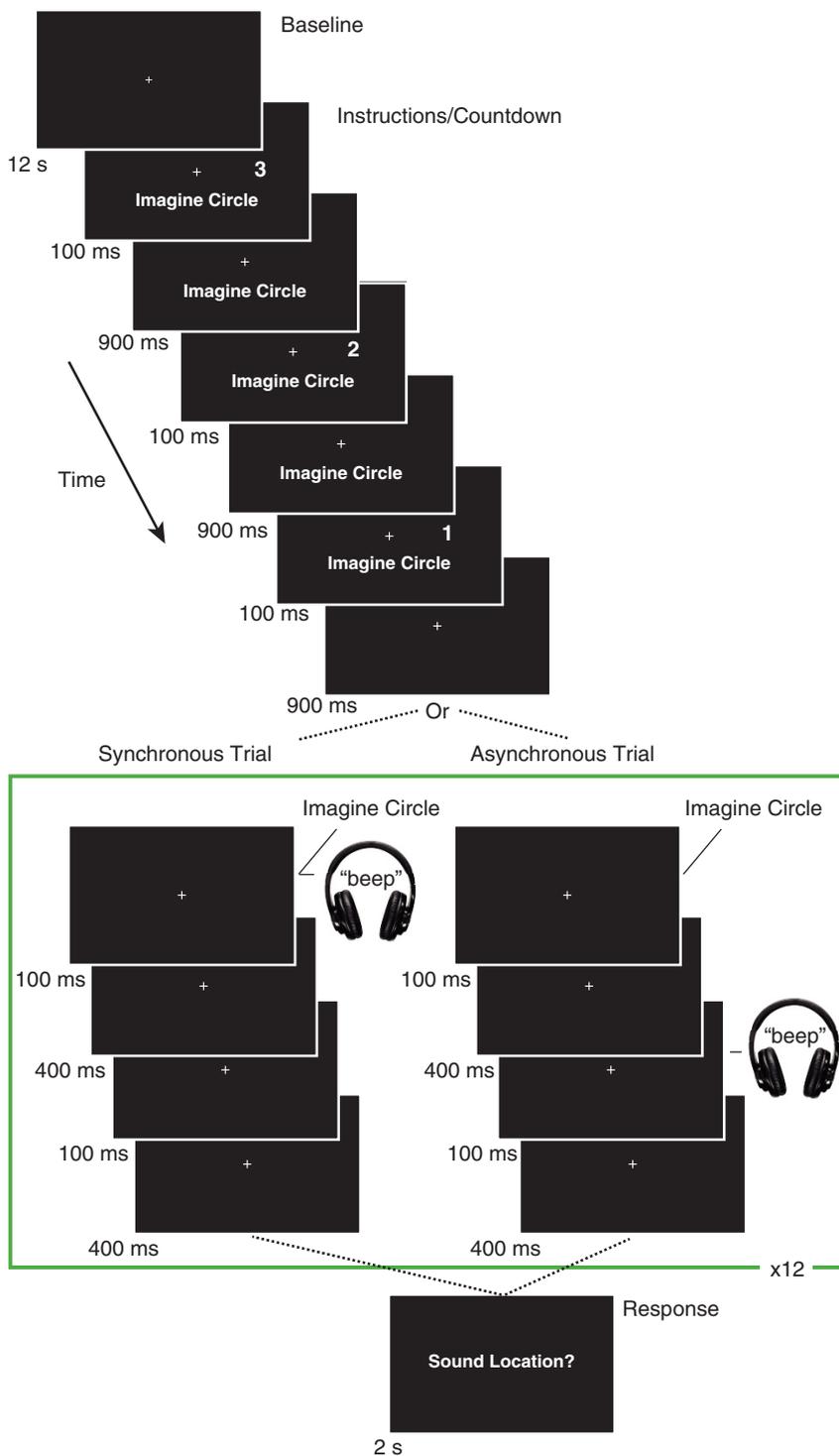
The authors declare no competing financial interests.

This article is freely available online through the *JNeurosci* Author Open Choice option.

Correspondence should be addressed to Christopher C. Berger at the above address. E-mail: christopher.c.berger@ki.se.

DOI:10.1523/JNEUROSCI.0943-14.2014

Copyright © 2014 the authors 0270-6474/14/3313684-09\$15.00/0



**Figure 1.** Task design. Each trial was preceded by 12 s of fixation, followed by instructions and a countdown. The instructions informed the participants that they should imagine the circle on that trial, and the countdown cued the participants to the timing and location (left or right; right in the above example trial) of the to-be-imagined visual stimulus. Following the countdown, the participants imagined the brief appearance (100 ms) of the visual stimulus once per second for 12 s while a brief auditory (100 ms) stimulus was presented in synchrony or asynchronously (i.e., 500 ms following the onset of the first imagined visual stimulus). Participants indicated whether they perceived the sound to come from the left, center, or right of fixation after each 12 s trial.

Based on previous research on multisensory perception, we hypothesized that neural activity in the multisensory temporal association cortex constitutes the key mechanism underlying the integration of convergent signals from imagined visual and real

auditory stimuli, and therefore expected that the strength of the imagery-induced ventriloquist illusion would be reflected in increases in the blood oxygenation level-dependent (BOLD) contrast signal in this area. Furthermore, we tested the prediction that the imagery-induced ventriloquist illusion is associated with increases in effective connectivity between the temporal association cortex and the auditory cortex.

### Materials and Methods

**Participants.** Twenty-two healthy, right-handed participants (age,  $29.09 \pm 5.56$  years; 10 females) participated in the experiment. Two additional participants were excluded because they were unable to maintain fixation throughout the experiment due to excessive sleepiness. All participants were recruited from the student population in the Stockholm area, were healthy, reported no history of psychiatric illness or neurologic disorder, and had no problems with hearing or vision (or had corrected to normal vision). All participants provided written informed consent before the start of the experiment, which was approved by the Regional Ethical Review Board of Stockholm.

**Design and procedures.** The general experimental procedures were explained to the participants before entering the scanner. Once inside the scanner, the details of each condition and the timing of all of the events for each trial were explained. For the imagery conditions, participants were given instructions about where, when, and how they should imagine the appearance of the visual stimulus (i.e., a white circle; radius, 20 mm) while looking at the visual display. First, participants were shown the object that they should imagine. The object was then taken away, and participants were asked to imagine it as vividly as possible. Instructions continued once the participant felt confident and comfortable with imagining the stimulus. Before data acquisition, participants performed a practice trial of each of the possible stimulus combinations during the MRI calibration scans. This allowed the participant to practice imagining the visual stimulus vividly and with the correct timing with the distraction of the scanner noise present before the actual experiment.

In total, there were seven possible stimulus combinations. The possible stimulus combinations were AV<sub>i</sub> synchronous left; AV<sub>i</sub> synchronous right; AV<sub>i</sub> asynchronous left; AV<sub>i</sub> asynchronous right; V<sub>i</sub> only left; V<sub>i</sub> only right; A<sub>i</sub> only, where A stands for auditory stimulus, V<sub>i</sub> stands for imagined visual stimulus, A<sub>i</sub> stands for imagined auditory stimulus, and left and right denote the location of the imagined visual stimulus. Each stimulus combination was presented in a pseudorandom order and repeated twice per run (except A<sub>i</sub> only, which was presented four times per session), resulting in six repetitions of each stimulus combination throughout the entire experiment (stimulus combinations from the left and the right were ultimately combined to examine the effects of AV<sub>i</sub> synchrony over AV<sub>i</sub> asynchrony, resulting in 12

was presented four times per session), resulting in six repetitions of each stimulus combination throughout the entire experiment (stimulus combinations from the left and the right were ultimately combined to examine the effects of AV<sub>i</sub> synchrony over AV<sub>i</sub> asynchrony, resulting in 12

repetitions per condition throughout the experiment). Only AV<sub>i</sub> synchronous and AV<sub>i</sub> asynchronous conditions were used for the main analyses. These conditions were chosen for our main analysis in light of several previous neuroimaging studies, which found robust multisensory integration effects by manipulating the temporal correspondence between audiovisual stimuli (Calvert et al., 2000; Bischoff et al., 2007; Noeselt et al., 2007; Lewis and Noppeney, 2010; Marchant et al., 2012). Furthermore, comparing synchronous and asynchronous trials enabled us to obtain a behavioral estimate of the ventriloquist effect that could be compared and related to the observed brain activity under conditions controlling for nonspecific activity related to the mere presence (vs absence) of the stimuli.

The visual display was presented via MR-compatible LCD video goggles (NordicNeuroLab) and auditory stimuli were transmitted in mono mode via stereo MR-compatible headphones. The auditory stimulus consisted of a brief 100 ms “beep” tone (mixed 3000/4000 Hz sinusoidal tone) that was presented at a comfortable volume still audible over the scanner noise. Presentation of the auditory stimulus in mono mode to stereo headphones resulted in the perception that the tone was spatially centered. During the experiment, the participants indicated whether they heard the sound come from the left, center, or right, by pressing one of three designated buttons on a hand-held MR-compatible fiber optic response pad (Current Designs). Stimulus presentation was controlled using PsychoPy software (Peirce, 2007, 2008) on a 13-inch MacBook computer.

In the main experiment, the participant first saw a fixation cross (12 s), which served as our baseline control condition. Next, instructions appeared (3 s) just below fixation informing the participant what he or she should imagine on that trial. For the AV<sub>i</sub> synchrony and AV<sub>i</sub> asynchrony trials, the instructions “Imagine circle” informed the participants that they should imagine the visual stimulus on that trial. The words “same” or “different” were also included at the end of this instruction to inform the participant that a sound would be presented synchronously or asynchronously, respectively, with the imagined visual stimulus; this allowed participants to maintain their timing and obviated the possibility that participants would begin to synchronize their imagery with the salient auditory stimulus over the course of the 12 s asynchronous trials. Simultaneous with the instruction screen, a countdown from 3 appeared 20° to the left or the right of fixation. Each number in the countdown appeared for 100 ms on the screen every 900 ms. The participants were instructed at the outset of the experiment to imagine the first appearance of the circle as vividly as possible at the next beat in the countdown (i.e., 900 ms after the disappearance of the “1” in the countdown) in the exact location of the countdown while maintaining fixation. Thus, the countdown instructed the participants when they should imagine the first appearance of the circle as well as where they should imagine it. Importantly, the participants’ task was the same in both the synchrony and asynchrony trials (i.e., the timing of their imagery was the same in every trial; only the timing of the auditory stimulus was manipulated). Moreover, participants were explicitly instructed (and practiced before the start of the experiment) to imagine the visual stimulus at the same pace (i.e., one per second for 12 s) on every trial and not to rely on the timing of the sound for their timing of the imagined visual stimulus. Thus, auditory stimuli were as task irrelevant as could be made possible in such a paradigm. To ensure participants maintained fixation, eye movements were monitored online via an MR-compatible EyeTracking camera (NordicNeuroLab) and ViewPoint EyeTracker systems software (Arrington Research) throughout the experiment.

**Behavioral data analysis.** To determine whether visual imagery of a spatially incongruent circle led to the ventriloquist effect during synchronous trials, the left, right, and center responses were coded as  $-1$ ,  $1$ , and  $0$ , respectively, and averaged within each condition. In this way, an average response of  $0$  would reflect no localization bias;  $1$  would reflect an extreme bias to the right; and  $-1$  would reflect an extreme bias to the left. A ventriloquism index—an unbiased estimate of the effect of visual stimuli (or visual imagery) over auditory stimuli—was then calculated by subtracting each participant’s average bias in the auditory-only condition (from functional localizer blocks; see below) from their averages in all other conditions. The benefit of such an estimate is that it controls for

false positives (i.e., instances in which the participant indicates that the auditory stimulus came from the same direction as the imagined visual stimulus because of baseline response/perceptual biases), which could confound our results. Ventriloquism indices from conditions with visual stimuli imagined (or perceived during functional localizer blocks; see below for details) on the left and the right were collapsed to increase statistical power by reverse scoring the indices from the left conditions and averaging them with those from the right conditions. Behavioral data were analyzed using R (R Development Core Team, 2010). The paired differences (between synchrony and asynchrony) were first plotted (i.e., density and quantile–quantile plots) and assessed for normality. As the paired differences followed a normal (i.e., Gaussian) distribution, a repeated-measures *t* test (two-tailed) was then used to assess statistical significance between the synchronous and asynchronous conditions (this was also used for the functional localizer behavioral data).

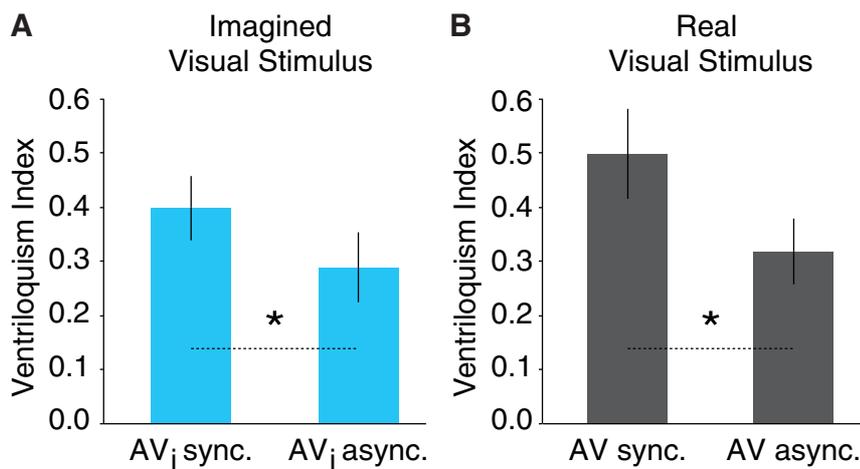
**fMRI data acquisition.** Participants were scanned using a 3T General Electric 750, MR scanner equipped with an eight-channel head coil to acquire gradient-echo T2\*-weighted echo-planar images with BOLD contrast as an index of local increases in synaptic activity (Logothetis et al., 2001; Magri et al., 2012). A functional image volume comprised 49 continuous slices that were 3 mm in thickness to ensure that the whole brain was within the field of view [FOV;  $96 \times 96$  matrix;  $3.0 \times 3.0$  mm; echo time (TE) = 30 ms]. One functional image volume was collected every 2.5 s [repetition time (TR) = 2500 ms] in an ascending, interleaved protocol. Thus, at the conclusion of the three experimental runs, 1028 image volumes were acquired for each participant. A high-resolution structural image was also acquired for each participant at the end of the experiment (3D MPRAGE sequence; voxel size:  $1 \times 1 \times 1$  mm; FOV,  $230.4 \times 230.4$  mm; 170 slices; TR = 6656 ms; TE = 2.93 ms; flip angle, 11°).

**fMRI data analysis.** The fMRI data were analyzed using the Statistical Parametric Mapping software package, version 8 (SPM8; <http://www.fil.ion.ucl.ac.uk/spm>; Wellcome Department of Cognitive Neurology). The functional images were realigned to correct for head movements and coregistered with each participant’s high-resolution structural scan. The anatomical image was then segmented into white matter, gray matter, and CSF partitions and normalized to the Montréal Neurological Institute (MNI) standard brain. The same transformation was then applied to all functional volumes, which were resliced to a  $2.0 \times 2.0 \times 2.0$  mm voxel size. The functional images were then spatially smoothed with an 8 mm full-width-at-half-maximum isotropic Gaussian kernel.

A linear regression model [general linear model (GLM)] was fitted to each participant’s data (first-level analysis) with regressors defined for each of the stimulus combinations described above. We also defined a condition of no interest corresponding to the 12 s baseline condition, the 3 s countdown and instructions, and the 2 s response. Each condition was modeled with a boxcar function and convolved with the standard SPM8 hemodynamic response function. Linear contrasts were defined within the GLM. The resulting contrast images from each subject were then entered into a random effects group analysis (second level). One-sample *t* tests were then used (21 degrees of freedom) to assess statistical significance.

To relate the imagery-induced ventriloquist illusion to BOLD activity, values corresponding to the strength of the ventriloquist illusion (i.e., the difference between the ventriloquism indices in the AV<sub>i</sub> synchrony and AV<sub>i</sub> asynchrony conditions) for each subject were entered as a covariate alongside the AV<sub>i</sub> synchrony–AV<sub>i</sub> asynchrony contrast images in a multiple linear regression model that was then estimated for the entire brain. Thus, the effect of the ventriloquism-strength covariate revealed all voxels displaying a significant positive relationship between the synchrony manipulation and strength of the imagery-induced ventriloquist effect.

In the main and the multiple linear regression analyses, we only report peaks of activation (unless otherwise stated) corresponding to  $p \leq 0.05$ , after correcting for multiple comparisons [familywise error (FWE) correction] within functionally defined regions of interest (fROIs). fROIs in the left superior temporal sulcus and left parietal cortex were identified in an orthogonal contrast involving real audiovisual stimuli from a functional localizer task (i.e., AV synchrony vs AV asynchrony; see below for details). Peaks of activation outside these fROIs were corrected for mul-



**Figure 2.** Imagery-induced ventriloquism. **A**, Behavioral results obtained in the scanner revealed a stronger ventriloquist effect when auditory stimuli were presented synchronously with imagined visual stimuli (AV<sub>i</sub> sync.) compared with asynchronously (AV<sub>i</sub> async.). **B**, The same effect was found for real visual stimuli presented synchronously with an auditory stimulus (AV sync.) compared with asynchronously (AV async.) during functional localizer scans. Error bars denote  $\pm$ SEM; asterisks between bars indicate significance ( $*p < 0.05$ ).

multiple comparisons based on the number of comparisons in the whole brain. Thus, we performed a whole-brain analysis that made use of fROIs to constrain the correction for multiple comparisons in a priori-specified regions of the brain. Because we had no a priori hypotheses concerning the functional significance of deactivations, i.e., less activity ( $p_{\text{uncorrected}} < 0.05$ ) for synchronous audiovisual stimuli compared with the resting baseline, such patterns of activation are not reported or discussed. However, in the interest of transparency and to make these data available for future research, the deactivations are still displayed in the figures (the activity visible in the parietal cortex, as observed in Fig. 3A, represents one area displaying such a pattern).

Effective connectivity changes between the left superior temporal sulcus (L. STS) and remote brain areas were assessed in a psychophysiological interaction (PPI) analysis (Friston et al., 1997) by defining a seed region for each participant centered on the peak voxel found within an 8 mm sphere centered on the group peak for the contrast AV<sub>i</sub> synchrony greater than AV<sub>i</sub> asynchrony. The seed region's time series was computed as the first eigenvariate of all voxels within a 4-mm-radius sphere centered on each participant's peak voxel. For each participant, regressors corresponding to the time series of the seed region (i.e., the physiological variable), the conditions of interest (i.e., the psychological variable), and their product (i.e., the PPI) were created and entered into a GLM estimated for each participant. Contrast estimates for the PPI regressor were analyzed in a random effects group analysis using a one-sample  $t$  test.

To relate the subject-to-subject variability in the strength of the imagery-induced ventriloquist effect to the effective connectivity to the L. STS, values corresponding to the strength of the ventriloquist effect for each subject (as described in the multiple-regression analysis above) were entered as a covariate alongside the PPI contrast images in a multiple linear regression model that was then estimated for the entire brain. Thus, the effect of the covariate revealed all voxels displaying a significant positive relationship between the PPI estimate and strength of the imagery-induced ventriloquist effect.

All reported peaks (unless otherwise stated) from the main PPI and PPI multiple regression analysis were FWE corrected for multiple comparisons within fROIs identified in an orthogonal PPI analysis conducted on scans from a functional localizer task (see below for more details from functional localizer scans).

**Functional localizer.** The corrections for multiple comparisons in all analyses were made within fROIs that had been identified by functional localizer scans that were interleaved throughout the experiment between imagery blocks. The possible stimulus combinations for the multisensory functional localizer blocks were as follows: AV-synchronous left; AV-

synchronous right; AV-asynchronous left; AV-asynchronous right; V-only left; V-only right; A only, where A stands for auditory stimulus, V stands for visual stimulus, and left and right denote the location of the presented visual stimulus. The task, timing, and number of stimulus presentations during the perceptual multisensory localizer were exactly the same as those used in the main experiment except that instead of imagining a visual stimulus, the participants actually saw the visual stimulus appear. These perceptual localizer scans were included in the same runs as the main experiment to minimize unspecific time or context differences, but importantly, the localizer scans and the main scans were completely orthogonal and thus statistically independent. For consistency, the pretrial instructions and countdown were also included in these runs [although the instruction now informed the participants that they would see a circle appear on that trial (e.g., "See Circle")].

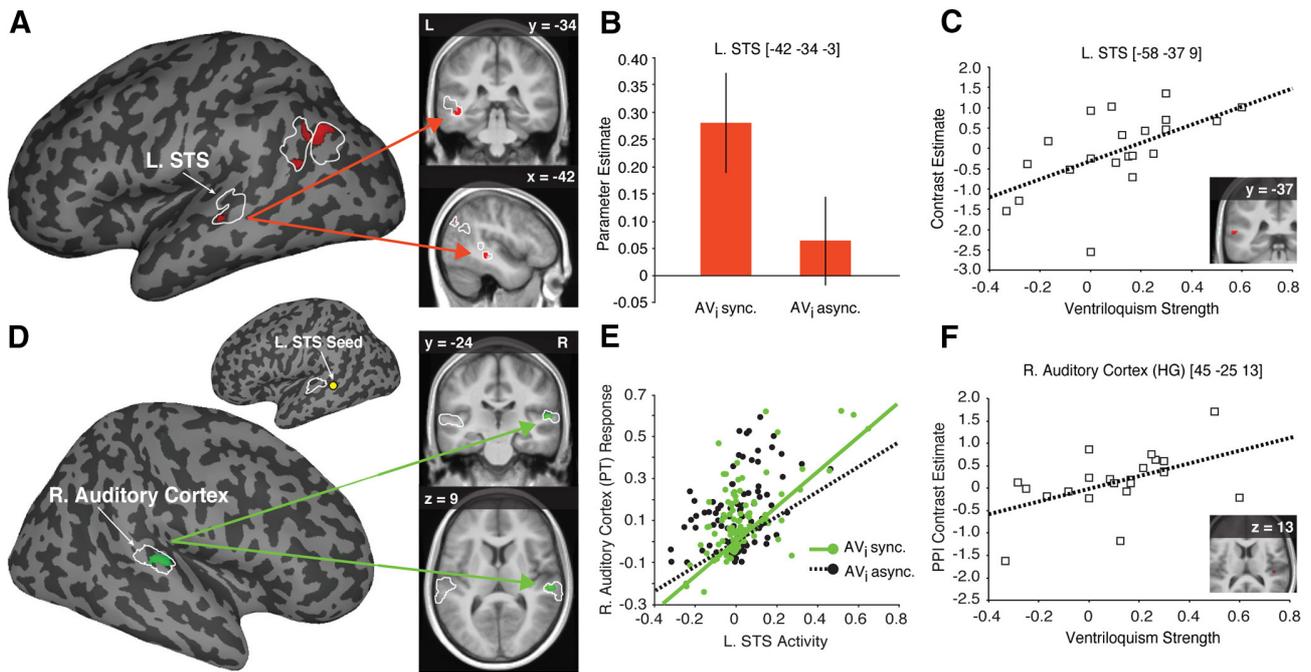
To identify multisensory areas sensitive to audiovisual synchrony, the AV synchrony trials were contrasted with the AV asynchrony trials. The resulting regions identified ( $p_{\text{uncorrected}} < 0.05$ ) were the L. STS and the left inferior parietal lobe (see Fig. 3A). MarsBar (<http://marsbar.sourceforge.net>; Brett et al., 2002) was used to create and export fROIs that were then used to correct for multiple comparisons in the main and multiple-regression analyses described above.

A PPI analysis was conducted on functional localizer data to assess effective connectivity changes between the L. STS and remote brain areas, particularly the auditory cortex, when real visual stimuli were presented synchronously with auditory stimuli. Clusters of activation were identified ( $p_{\text{uncorrected}} < 0.005$ ) in the left and right auditory cortices. MarsBar was used to export these clusters of activation to be used as fROIs in the main PPI and PPI multiple-regression analyses described above (see Fig. 3D). Significant peaks of activation from the functional localizer PPI analysis are reported in Figure 4 and were FWE-corrected for multiple comparisons within a sphere (radius, 6 mm) centered on the peak coordinates of the left ( $-50, -30, 11$ ;  $t_{(21)} = 9.52$ ,  $p_{\text{uncorrected}} < 0.001$ ) and right ( $59, -23, 9$ ;  $t_{(21)} = 11.01$ ,  $p_{\text{uncorrected}} < 0.001$ ) planum temporale (PT), and left ( $-45, -25, 9$ ;  $t_{(21)} = 6.90$ ,  $p_{\text{uncorrected}} < 0.001$ ) and right ( $44, -21, 9$ ;  $t_{(21)} = 4.70$ ,  $p_{\text{uncorrected}} < 0.001$ ) Heschl's gyrus (HG) portions of the auditory cortex from the orthogonal (i.e., statistically independent) A only-baseline contrast (see Fig. 4).

## Results

An analysis of the behavioral data obtained from the scanner revealed that imagining a spatially disparate visual stimulus in synchrony with an auditory stimulus (vs asynchronously) led to a significant translocation of the auditory stimuli ( $t_{(21)} = 2.15$ ,  $p = 0.043$ ,  $d = 0.38$ ; Fig. 2A). This effect mirrored the comparison of equivalent conditions from functional localizer scans,  $t_{(21)} = 2.91$ ,  $p = 0.008$ ,  $d = 0.53$  (Fig. 2B), consistent with previous behavioral evidence of visual imagery-induced ventriloquism (Berger and Ehrsson, 2013).

The primary analysis of the fMRI data focused on identifying the neural correlates of the imagery-induced ventriloquist effect by testing whether there were differences in activity within multisensory areas when the participants imagined a spatially incongruent visual stimulus in synchrony (vs asynchrony) with an auditory stimulus. We found that audiovisual synchrony of imagined visual stimuli was associated with a significant increase in activity in the L. STS compared with asynchrony ( $-42, -34, -3$  [ $x, y, \text{ and } z$  coordinates in MNI standard space];  $t_{(21)} = 3.89$ ,



**Figure 3.** Neural basis of imagery-induced ventriloquism. **A**, Activity associated with audiovisual synchrony (vs asynchrony) for imagined visual stimuli within the functionally defined multi-sensory regions of interest (fROIs outlined in white) overlaid on a representative inflated cortical surface (left); coronal and sagittal sections displaying the peak activation in the L. STS is overlaid on the average normalized anatomical image from our participants (right). The activity differences observed in the parietal cortex were the result of deactivations, i.e., less activity for synchronous auditory and imagined visual stimuli compared with the resting baseline (see Results). **B**, Bar plot shows the parameter estimates from the significant peak of activation in the L. STS; error bars denote  $\pm$ SEM. **C**, *Post hoc* multiple-regression analysis demonstrating that the activity in the L. STS in the AV<sub>i</sub> synchronous (AV<sub>i</sub> sync.; vs AV<sub>i</sub> asynchronous (AV<sub>i</sub> async.)) condition could be predicted by the strength (i.e., the difference of AV<sub>i</sub> synchrony and AV<sub>i</sub> asynchrony ventriloquism indices) of the imagery-induced ventriloquist effect. **D**, Significant enhanced connectivity between the right auditory cortex and the L. STS seed region for the AV<sub>i</sub> synchronous (vs AV<sub>i</sub> asynchronous) condition overlaid on a representative inflated cortical surface (bottom left). A yellow circle marks the approximate location of the L. STS seed on an inflated left hemisphere cortical surface (top left; fROIs outlined in white). Coronal and axial sections displaying the peak connectivity to the right auditory cortex are overlaid on the average anatomical image from our participants (right). **E**, Plot of the PPI for one representative subject showing a steeper regression slope relating L. STS activity to the response magnitude of the right (R.) auditory cortex during the AV<sub>i</sub> synchrony (AV<sub>i</sub> sync., green) compared with the AV<sub>i</sub> asynchrony condition (AV<sub>i</sub> async., black). **F**, *Post hoc* multiple-regression analysis demonstrating that effective connectivity from L. STS to the right (R.) auditory cortex in the AV<sub>i</sub> synchrony (vs AV<sub>i</sub> asynchrony) condition could be predicted by the strength of the imagery-induced ventriloquist effect. Activation maps are displayed at  $p_{\text{uncorrected}} < 0.01$  for display purposes only.

**Table 1. Anatomical regions, MNI coordinates, and statistics from all analyses**

Analysis; anatomical region	MNI <i>x, y, z</i> (mm)	Peak <i>t</i>	Peak <i>z</i>	<i>p</i> (FWE corrected)
Main analysis (AV <sub>i</sub> synchrony greater than AV <sub>i</sub> asynchrony); L. STS	−42, −34, −3	3.89	3.34	0.013
Multiple-regression analysis with ventriloquism covariate; L. STS	−58, −37, 9	3.18	2.83	0.052
PPI analysis; right PT (auditory cortex)	57, −24, 9	3.44	3.03	0.021
PPI analysis; left HG (auditory cortex)	−42, −24, 0	2.59	2.39	*0.008
PPI multiple-regression analysis with ventriloquism covariate; right HG (auditory cortex)	45, −25, 13	2.65	2.42	*0.008
PPI multiple-regression analysis with ventriloquism covariate; left PT (auditory cortex)	−54, −18, 6	3.21	2.85	0.054

\*Uncorrected for multiple comparisons based on a priori hypotheses.

$p_{\text{FWE-corrected}} < 0.05$  (Fig. 3A,B; Table 1, main analysis). The two peaks of activity observed in the parietal cortex (Fig. 3A) were the result of significant differences in deactivations, i.e., less activity in the AV<sub>i</sub> synchrony condition compared with the resting state baseline [−42, −72, 45:  $M_{\text{AV}_i \text{ sync}} = 0.01 \pm 1.49$ ,  $M_{\text{AV}_i \text{ async}} = -1.03 \pm 2.05$ ; −39, −75, 42:  $M_{\text{AV}_i \text{ sync}} = -0.18 \pm 1.31$ ,  $M_{\text{AV}_i \text{ async}} = -1.00 \pm 1.64$  (MNI coordinates and means  $\pm$  SDs of parameter estimates from peaks of activation in the AV<sub>i</sub> synchrony and AV<sub>i</sub> asynchrony conditions)]. Because we did not have an a priori hypothesis regarding deactivations in this region, this activity was assumed to be unrelated to the multisensory percept under investigation. There were no peaks of activation outside the fROIs that survived the correction for multiple comparisons at the whole-brain level, and no statistically trending peaks ( $p_{\text{uncorrected}} < 0.001$ ) were observed outside the fROIs in other areas related to

audiovisual processing, such as the primary sensory cortices, pre-frontal cortex, basal ganglia, or superior colliculus.

Next we examined whether any synchrony-specific BOLD activity in the brain could be predicted by the strength of the imagery-induced ventriloquist illusion for each subject. Thus, we examined whether the unbiased estimate of the strength of the ventriloquist illusion, based on the difference between the ventriloquism indices in the AV<sub>i</sub> synchrony and AV<sub>i</sub> asynchrony conditions calculated from the participants' responses, was linearly related to the strength of the BOLD response in the AV<sub>i</sub> synchrony condition compared with the AV<sub>i</sub> asynchrony condition in an additional whole-brain multiple-regression analysis. This analysis revealed that participants whose auditory perception was biased most in synchronous (vs asynchronous) trials also showed the strongest activity in the L. STS (−58, −37, 9;  $t_{(21)} = 3.18$ ,

$p_{\text{FWE-corrected}} = 0.052$ ; see Fig. 3C; Table 1, multiple-regression analysis with ventriloquism covariate). These two findings link the imagery-induced ventriloquist effect to activity in the L. STS.

In light of previous findings demonstrating increased effective connectivity between the STS and primary visual and auditory areas during audiovisual synchrony (Noesselt et al., 2007; Marchant et al., 2012), we also conducted a separate PPI analysis (Friston et al., 1997), in which we tested whether imagining a visual stimulus in synchrony with an auditory stimulus was associated with increased effective connectivity between the L. STS and primary visual and/or auditory areas. A significant increase in effective connectivity was observed between the L. STS and the right auditory cortex (PT; 57, -24, 9;  $t_{(21)} = 3.44$ ,  $p_{\text{FWE-corrected}} = 0.021$ ) when participants imagined a visual stimulus in synchrony with a real auditory stimulus compared with imagining a visual stimulus in asynchrony with a real auditory stimulus [a *post hoc* analysis, conducted for descriptive purposes, also revealed increased connectivity to the left auditory cortex at a lower statistical threshold (HG; -42, -24, 0;  $t_{(21)} = 2.59$ ,  $p_{\text{uncorrected}} = 0.008$ ), but did not survive correction for multiple comparisons (see Fig. 3D,E; Table 1, PPI analysis).

Finally, we also examined whether any synchrony-specific increase in connectivity could be predicted by the strength of the imagery-induced ventriloquist effect in an additional whole-brain multiple-regression analysis, and found that participants whose auditory perception was most biased when auditory stimuli were presented in synchrony with their imagination of a circle also showed the strongest effective connectivity between L. STS and the left auditory cortex [HG; -54, -18, 6;  $t_{(21)} = 2.85$ ,  $p_{\text{FWE-corrected}} = 0.054$ ; a *post hoc* analysis, conducted purely for descriptive purposes, also revealed a positive relationship between the right auditory cortex and the strength of the ventriloquist effect (HG; 45, -25, 13;  $t_{(21)} = 2.65$ ,  $p_{\text{uncorrected}} = 0.008$ ) but did not survive correction for multiple comparisons; see Fig. 3F; Table 1, PPI multiple-regression analysis with ventriloquism covariate]. Thus, the imagery-induced ventriloquist effect is associated with a strong functional interplay between the auditory cortex and the L. STS.

## Discussion

We have demonstrated that the illusory translocation of auditory stimuli toward the location of an imagined visual stimulus—the imagery-induced ventriloquist effect—is associated with increased activity in the L. STS and with increased effective connectivity between the L. STS and the auditory cortex. Moreover, we found that the strength of this illusion is related to the degree of increased activity in the L. STS and to the degree of increased effective connectivity between the L. STS the auditory cortex. These findings are in line with those obtained using the standard ventriloquist effect (using real stimuli) observed in the present study and in previous neuroimaging studies (Bischoff et al., 2007; Bonath et al., 2007). Together, these results suggest that the fusion of imagery and real sensory signals is mediated by the same integrative mechanisms in the association cortex and primary sensory cortex as those that mediate the fusion of real sensory stimuli.

The L. STS has previously been implicated as a key site for the integration of audiovisual stimuli (Beauchamp et al., 2004a, 2004b) and in studies on the perceptual effects of audiovisual integration (Bushara et al., 2003; Bischoff et al., 2007; Stevenson and James, 2009; Werner and Noppeney, 2010b; Marchant et al., 2012). Anatomically, the STS is situated between the visual and auditory cortices, with direct connections from both, making it

an ideal candidate for the integration of convergent auditory and visual stimuli (Seltzer and Pandya, 1994; Lyon and Kaas, 2002; Kaas and Collins, 2004; Wallace et al., 2004). Moreover, electrophysiological recordings in nonhuman primates have demonstrated that this region contains cells that have the capacity to integrate auditory and visual signals at the single-neuron level (Bruce et al., 1981; Schroeder and Foxe, 2002; Dahl et al., 2009; Perrodin et al., 2014); and neuroimaging studies on humans have also implicated the STS in the integration of a wide range of audiovisual stimuli (Noesselt et al., 2007; Stein and Stanford, 2008; Marchant et al., 2012), including one neuroimaging study linking the ventriloquist illusion to increased activity in the L. STS (Bischoff et al., 2007). In the present study, we found that STS activity was greater when the participants imagined the visual stimuli in synchrony, compared with asynchrony, with the auditory stimuli, and that the degree of this BOLD effect was correlated with the behaviorally indexed imagery-induced ventriloquist effect across participants. Our findings suggest that neuronal signals produced by imagined visual stimuli are combined with signals generated by real auditory stimuli in the STS, thereby facilitating the creation of a coherent audiovisual representation of a single external event.

In addition to the STS, previous work has also implicated the primary auditory and visual cortices in multisensory interactions during the processing of synchronous audiovisual stimuli (Driver and Noesselt, 2008; Kayser et al., 2010; Werner and Noppeney, 2010a). Interestingly, we did not observe any significant activity in either the auditory or visual cortices even at lower statistical thresholds ( $p_{\text{uncorrected}} < 0.05$ ) for synchronous (vs asynchronous) audiovisual stimuli in the functional localizer data, nor in the main analyses when comparing synchronously and asynchronously imagined visual and real auditory stimuli. However, the results of our effective connectivity analyses, which showed that the imagery-induced ventriloquist illusion was associated with an enhanced effective connectivity between the L. STS and the auditory cortex, are in line with previous studies implicating the involvement of the auditory cortex in multisensory processing (Bonath et al., 2007; Driver and Noesselt, 2008; Werner and Noppeney, 2010a).

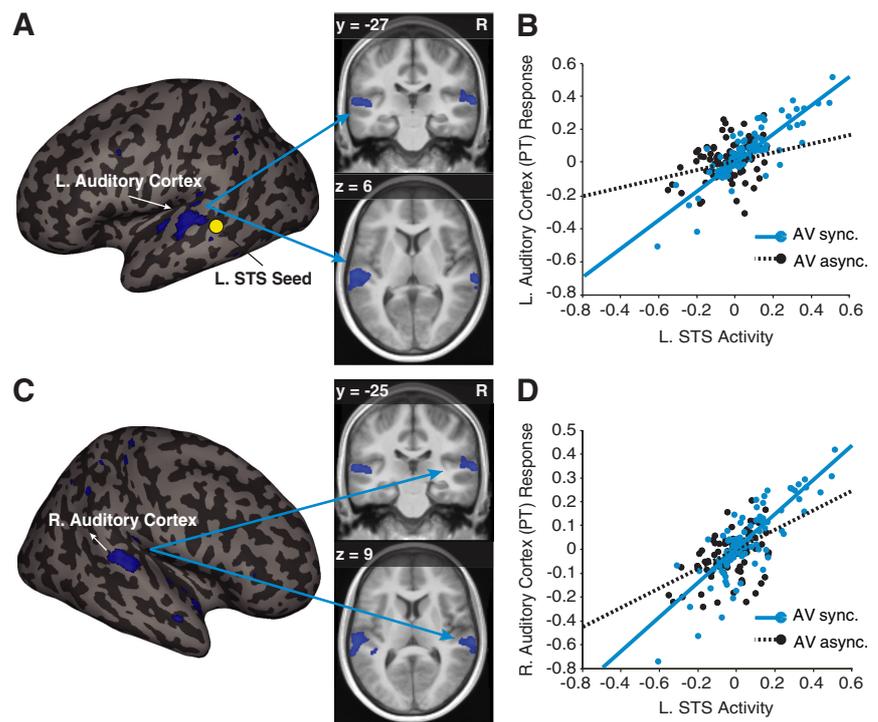
The involvement of the auditory cortex in perceptual multisensory interactions has been attributed to inputs from higher-order areas in the association cortex or from other sensory areas via long-range anatomical connections (Ghazanfar et al., 2005, 2008; Driver and Noesselt, 2008; Marchant et al., 2012). The posterior portions of the auditory cortex, observed in our connectivity analysis, have also been implicated in the spatial localization of auditory stimuli (Tian et al., 2001; Bonath et al., 2007; Lomber and Malhotra, 2008; Ahveninen et al., 2013), and in one previous neuroimaging study of the ventriloquist effect (Bonath et al., 2007). Our findings from the effective connectivity analyses are in line with these observations. Our interpretation is that this increase in effective connectivity to the auditory cortex reflects an important mechanism by which visual stimuli (real or imagined) lead to changes in the processing of auditory stimuli in external space, such that endogenously and exogenously induced multisensory perception is mediated by the association cortex, and by the information exchange between the association cortex and early “modality-specific” cortex.

The present results provide new insight regarding top-down effects in multisensory integration. While there has been a great deal of research examining the effects of attention, expectation, and prior knowledge on multisensory integration (Engel et al., 2001, 2012; Talsma et al., 2010), our results sug-

gest that coherent multisensory representations of external objects are not only modulated by top-down processing, but can indeed be formed from signals that are partly real and partly the product of our explicit mental images. That is, signals from imagined stimuli are capable of perceptually fusing with real stimuli by engaging the same integrative mechanisms as real cross-modal sensory stimuli. This finding suggests that imagery can substitute sensation in multisensory perception rather than just modulate sensory processing related to external stimuli, as in the case of attention.

It is important to note that many experiments on the ventriloquist effect have successfully demonstrated that the ventriloquist illusion reflects a genuine perceptual phenomenon that cannot merely be explained by cognitive bias or postperceptual decisions (Bertelson and Aschersleben, 1998; Bertelson et al., 2000, 2006; Vroomen et al., 2001; Alais and Burr, 2004). In a recent behavioral experiment using a psychophysical staircase procedure, we were able to demonstrate that the imagery-induced version of the ventriloquist illusion is also indicative of a genuine perceptual phenomenon (Berger and Ehrsson, 2013). Therefore, we are confident that the imaging results reported here reflect the genuine perceptual translocation of the auditory stimulus toward the imagined visual stimulus. Such an interpretation is in agreement with the results presented here, in which the strength of the illusion is reflected in the strength of activity in the STS and connectivity to the auditory cortex rather than activity in or connectivity to prefrontal regions previously implicated in perceptual decisions (Noppeney et al., 2010).

Although the present data suggest that the STS and increased connectivity between the STS and auditory cortex play an important role in integrating imagined visual stimuli with the auditory stimuli that we perceive in the external world, future research may serve to further investigate the specific mechanisms associated with other features of this multisensory integrative process. For instance, we were able to relate the unbiased estimate of the strength of the ventriloquist effect for each subject to the strength of the BOLD response in the AV<sub>i</sub> synchrony condition compared with the AV<sub>i</sub> asynchrony condition; however, future research may be able to make use of the trial-to-trial variability in the illusory percept to further our understanding of the relationship between the consciously reported percepts and the basic multisensory integration mechanisms. Further, while we demonstrated here that the integration of mental imagery and perception relies on at least partially overlapping neural mechanisms, we hope that our findings will provide the basis for future investigations into the mechanisms by which the brain distinguishes between imagery and perception. Such an investigation may be useful in understanding circumstances in which one fails to distinguish between sensory



**Figure 4.** Effective connectivity during perceptual functional localizer scans. **A**, Significant enhanced connectivity between the left auditory cortex ( $-51, -27, 6$ ;  $t_{(21)} = 3.74$ ,  $p_{FWE-corrected} < 0.05$ ) and the L. STS seed region for the AV synchrony (vs AV asynchrony) condition overlaid on a representative inflated cortical surface (left) and in a coronal (top right) and axial (bottom right) section of the average anatomical image from our participants. A yellow circle marks the approximate location of the L. STS seed. **B**, Plot of the PPI for one representative subject showing a steeper regression slope relating L. STS activity with the response magnitude of the left (L.) auditory cortex during the AV synchrony (AV sync., blue) compared with the AV asynchrony condition (AV async., black). **C**, Significant ( $63, -25, 9$ ;  $t_{(21)} = 3.65$ ,  $p_{FWE-corrected} < 0.05$ ) enhanced connectivity between the right (R.) auditory cortex and the L. STS seed region for the AV synchrony (vs AV asynchrony) condition overlaid on a representative inflated cortical surface (left) and in the coronal (top right) and axial (bottom right) section of the average anatomical image from our participants. **D**, Plot of the PPI for one representative subject showing a steeper regression slope relating L. STS activity with the response magnitude of the right (R.) auditory cortex during the AV synchrony (AV sync., blue) compared with the AV asynchrony condition (AV async., black). Activation maps are displayed at  $p_{uncorrected} < 0.005$  for display purposes.

stimuli generated in one's mind and sensory stimuli perceived in the external world, such as hallucinations.

The results we described here advance our understanding of the functional and neuroanatomical similarities between imagery and perception. Numerous imaging studies have compared activation when imagining or perceiving a sensory stimulus, and have described a remarkable degree of neuroanatomical overlap of the activation patterns in sensory cortices (Farah, 1984, 1989; O'Craven and Kanwisher, 2000; Kosslyn et al., 2001; Ehrsson et al., 2003; Ganis et al., 2004; Oh et al., 2013). Recent studies using brain-decoding techniques have also shown that the fine-grained patterns of activity in sensory areas when imagining a stimulus are similar to those when perceiving it (Thirion et al., 2006; Stokes et al., 2009; Cichy et al., 2012; Horikawa et al., 2013). Our results, however, go beyond these observations by showing that endogenously generated sensory signals are not only capable of activating areas responsible for perceiving sensory stimuli, but are in fact of sufficient quality and signal strength as to fully integrate with exogenous sensory stimuli from a different sensory modality to form coherent multisensory representations of external events. To the best of our knowledge, this study is the first to image such a behaviorally relevant interaction between imagery and perception. These findings provide renewed support for perceptually based theories of imagery.

## References

- Ahveninen J, Huang S, Nummenmaa A, Belliveau JW, Hung AY, Jääskeläinen IP, Rauschecker JP, Rossi S, Tiitinen H, Raij T (2013) Evidence for distinct human auditory cortex regions for sound location versus identity processing. *Nat Commun* 4:2585. [CrossRef Medline](#)
- Alais D, Burr D (2004) The ventriloquist effect results from near-optimal bimodal integration. *Curr Biol* 14:257–262. [CrossRef Medline](#)
- Anema HA, de Haan AM, Gebuis T, Dijkerman HC (2012) Thinking about touch facilitates tactile but not auditory processing. *Exp Brain Res* 218:373–380. [CrossRef Medline](#)
- Beauchamp MS, Lee KE, Argall BD, Martin A (2004a) Integration of auditory and visual information about objects in superior temporal sulcus. *Neuron* 41:809–823. [CrossRef Medline](#)
- Beauchamp MS, Argall BD, Bodurka J, Duyn JH, Martin A (2004b) Unraveling multisensory integration: patchy organization within human STS multisensory cortex. *Nat Neurosci* 7:1190–1192. [CrossRef Medline](#)
- Beauchamp MS, Nath AR, Pasalar S (2010) fMRI-guided transcranial magnetic stimulation reveals that the superior temporal sulcus is a cortical locus of the McGurk effect. *J Neurosci* 30:2414–2417. [CrossRef Medline](#)
- Berger CC, Ehrsson HH (2013) Mental imagery changes multisensory perception. *Curr Biol* 23:1367–1372. [CrossRef Medline](#)
- Bertelson P, Aschersleben G (1998) Automatic visual bias of perceived auditory location. *Psychon Bull Rev* 5:482–489. [CrossRef](#)
- Bertelson P, Vroomen J, de Gelder B (2000) The ventriloquist effect does not depend on the direction of deliberate visual attention. *Percept Psychophys* 62:321–332. [CrossRef Medline](#)
- Bertelson P, Frissen I, Vroomen J, de Gelder B (2006) The aftereffects of ventriloquism: patterns of spatial generalization. *Percept Psychophys* 68:428–436. [CrossRef Medline](#)
- Bischoff M, Walter B, Blecker CR, Morgen K, Vaitl D, Sammer G (2007) Utilizing the ventriloquism-effect to investigate audio-visual binding. *Neuropsychologia* 45:578–586. [CrossRef Medline](#)
- Bonath B, Noesselt T, Martinez A, Mishra J, Schwiecker K, Heinze HJ, Hillyard SA (2007) Neural basis of the ventriloquist illusion. *Curr Biol* 17:1697–1703. [CrossRef Medline](#)
- Brett M, Anton J, Valabregue R, Poline J (2002) Region of interest analysis using an SPM toolbox [abstract]. 8th Int Conf Funct Mapp Hum Brain (OHBM), June 2–6; Sendai, Japan. *Neuroimage* 16:497.
- Bruce C, Desimone R, Gross CG (1981) Visual properties of neurons in a polysensory area in superior temporal sulcus of the macaque. *J Neurophysiol* 46:369–384. [Medline](#)
- Bunzeck N, Wuestenberg T, Lutz K, Heinze HJ, Jancke L (2005) Scanning silence: mental imagery of complex sounds. *Neuroimage* 26:1119–1127. [CrossRef Medline](#)
- Bushara KO, Hanakawa T, Immisch I, Toma K, Kansaku K, Hallett M (2003) Neural correlates of cross-modal binding. *Nat Neurosci* 6:190–195. [CrossRef Medline](#)
- Calvert GA, Campbell R, Brammer MJ (2000) Evidence from functional magnetic resonance imaging of crossmodal binding in the human heteromodal cortex. *Curr Biol* 10:649–657. [CrossRef Medline](#)
- Cichy RM, Heinze J, Haynes JD (2012) Imagery and perception share cortical representations of content and location. *Cereb Cortex* 22:372–380. [CrossRef Medline](#)
- Dahl CD, Logothetis NK, Kayser C (2009) Spatial organization of multisensory responses in temporal association cortex. *J Neurosci* 29:11924–11932. [CrossRef Medline](#)
- Driver J, Noesselt T (2008) Multisensory interplay reveals crossmodal influences on “sensory-specific” brain regions, neural responses, and judgments. *Neuron* 57:11–23. [CrossRef Medline](#)
- Ehrsson HH, Geyer S, Naito E (2003) Imagery of voluntary movement of fingers, toes, and tongue activates corresponding body-part-specific motor representations. *J Neurophysiol* 90:3304–3316. [CrossRef Medline](#)
- Ehrsson HH, Spence C, Passingham RE (2004) That’s my hand! Activity in premotor cortex reflects feeling of ownership of a limb. *Science* 305:875–877. [CrossRef Medline](#)
- Engel AK, Fries P, Singer W (2001) Dynamic predictions: oscillations and synchrony in top-down processing. *Nat Rev Neurosci* 2:704–716. [CrossRef Medline](#)
- Engel AK, Senkowski D, Schneider TR (2012) Multisensory integration through neural coherence. In: *The neural bases of multisensory processes* (Murray M, Wallace M, eds). Boca Raton, FL: CRC.
- Farah MJ (1984) The neurological basis of mental imagery: a componential analysis. *Cognition* 18:245–272. [CrossRef Medline](#)
- Farah MJ (1989) The neural basis of mental imagery. *Trends Neurosci* 12:395–399. [Medline](#)
- Friston KJ, Buechel C, Fink GR, Morris J, Rolls E, Dolan RJ (1997) Psychophysiological and modulatory interactions in neuroimaging. *Neuroimage* 6:218–229. [CrossRef Medline](#)
- Ganis G, Thompson WL, Kosslyn SM (2004) Brain areas underlying visual mental imagery and visual perception: an fMRI study. *Brain Res Cogn Brain Res* 20:226–241. [CrossRef Medline](#)
- Gentile G, Guterstam A, Brozzoli C, Ehrsson HH (2013) Disintegration of multisensory signals from the real hand reduces default limb self-attribution: an FMRI study. *J Neurosci* 33:13350–13366. [CrossRef Medline](#)
- Ghazanfar AA, Maier JX, Hoffman KL, Logothetis NK (2005) Multisensory integration of dynamic faces and voices in rhesus monkey auditory cortex. *J Neurosci* 25:5004–5012. [CrossRef Medline](#)
- Ghazanfar AA, Chandrasekaran C, Logothetis NK (2008) Interactions between the superior temporal sulcus and auditory cortex mediate dynamic face/voice integration in rhesus monkeys. *J Neurosci* 28:4457–4469. [CrossRef Medline](#)
- Horikawa T, Tamaki M, Miyawaki Y, Kamitani Y (2013) Neural decoding of visual imagery during sleep. *Science* 340:639–642. [CrossRef Medline](#)
- Howard IP, Templeton WB (1966) Human spatial orientation. London: Wiley.
- Kaas J, Collins C (2004) The resurrection of multisensory cortex in primates: connection patterns that integrate modalities. In: *The handbook of multisensory processes* (Calvert G, Spence C, Stein B, eds), pp 285–294. Cambridge, MA: Bradford.
- Kamitani Y, Tong F (2005) Decoding the visual and subjective contents of the human brain. *Nat Neurosci* 8:679–685. [CrossRef Medline](#)
- Kayser C, Logothetis NK, Panzeri S (2010) Visual enhancement of the information representation in auditory cortex. *Curr Biol* 20:19–24. [CrossRef Medline](#)
- Kosslyn SM, Ganis G, Thompson WL (2001) Neural foundations of imagery. *Nat Rev Neurosci* 2:635–642. [CrossRef Medline](#)
- Lewis R, Noppeney U (2010) Audiovisual synchrony improves motion discrimination via enhanced connectivity between early visual and auditory areas. *J Neurosci* 30:12329–12339. [CrossRef Medline](#)
- Logothetis NK, Pauls J, Augath M, Trinath T, Oeltermann A (2001) Neurophysiological investigation of the basis of the fMRI signal. *Nature* 412:150–157. [CrossRef Medline](#)
- Lomber SG, Malhotra S (2008) Double dissociation of “what” and “where” processing in auditory cortex. *Nat Neurosci* 11:609–616. [CrossRef Medline](#)
- Lyon DC, Kaas JH (2002) Evidence for a modified V3 with dorsal and ventral halves in macaque monkeys. *Neuron* 33:453–461. [CrossRef Medline](#)
- Macaluso E, Driver J (2005) Multisensory spatial interactions: a window onto functional integration in the human brain. *Trends Neurosci* 28:264–271. [CrossRef Medline](#)
- Magri C, Schridde U, Murayama Y, Panzeri S, Logothetis NK (2012) The amplitude and timing of the BOLD signal reflects the relationship between local field potential power at different frequencies. *J Neurosci* 32:1395–1407. [CrossRef Medline](#)
- Marchant JL, Ruff CC, Driver J (2012) Audiovisual synchrony enhances BOLD responses in a brain network including multisensory STS while also enhancing target-detection performance for both modalities. *Hum Brain Mapp* 33:1212–1224. [CrossRef Medline](#)
- Noesselt T, Rieger JW, Schoenfeld MA, Kanowski M, Hinrichs H, Heinze HJ, Driver J (2007) Audiovisual temporal correspondence modulates human multisensory superior temporal sulcus plus primary sensory cortices. *J Neurosci* 27:11431–11441. [CrossRef Medline](#)
- Noppeney U, Ostwald D, Werner S (2010) Perceptual decisions formed by accumulation of audiovisual evidence in prefrontal cortex. *J Neurosci* 30:7434–7446. [CrossRef Medline](#)
- O’Craven KM, Kanwisher N (2000) Mental imagery of faces and places activates corresponding stimulus-specific brain regions. *J Cogn Neurosci* 12:1013–1023. [CrossRef Medline](#)
- Oh J, Kwon JH, Yang PS, Jeong J (2013) Auditory imagery modulates frequency-specific areas in the human auditory cortex. *J Cogn Neurosci* 25:175–187. [CrossRef Medline](#)
- Pearce JW (2007) PsychoPy—psychophysics software in Python. *J Neurosci Methods* 162:8–13. [CrossRef Medline](#)

- Peirce JW (2008) Generating stimuli for neuroscience using PsychoPy. *Front Neuroinform* 2:10. [CrossRef Medline](#)
- Perrodin C, Kayser C, Logothetis NK, Petkov CI (2014) Auditory and visual modulation of temporal lobe neurons in voice-sensitive and association cortices. *J Neurosci* 34:2524–2537. [CrossRef Medline](#)
- R Development Core Team (2010) R: a language and environment for statistical computing. Vienna: R Foundation for Statistical Computing.
- Roth M, Decety J, Raybaudi M, Massarelli R, Delon-Martin C, Segebarth C, Morand S, Gemignani A, Décorps M, Jeannerod M (1996) Possible involvement of primary motor cortex in mentally simulated movement: a functional magnetic resonance imaging study. *Neuroreport* 7:1280–1284. [CrossRef Medline](#)
- Schroeder CE, Foxe JJ (2002) The timing and laminar profile of converging inputs to multisensory areas of the macaque neocortex. *Brain Res Cogn Brain Res* 14:187–198. [CrossRef Medline](#)
- Seltzer B, Pandya DN (1994) Parietal, temporal, and occipital projections to cortex of the superior temporal sulcus in the rhesus monkey: a retrograde tracer study. *J Comp Neurol* 343:445–463. [CrossRef Medline](#)
- Stein BE, Stanford TR (2008) Multisensory integration: current issues from the perspective of the single neuron. *Nat Rev Neurosci* 9:255–266. [CrossRef Medline](#)
- Stevenson RA, James TW (2009) Audiovisual integration in human superior temporal sulcus: Inverse effectiveness and the neural processing of speech and object recognition. *Neuroimage* 44:1210–1223. [CrossRef Medline](#)
- Stokes M, Thompson R, Cusack R, Duncan J (2009) Top-down activation of shape-specific population codes in visual cortex during mental imagery. *J Neurosci* 29:1565–1572. [CrossRef Medline](#)
- Talsma D, Senkowski D, Soto-Faraco S, Woldorff MG (2010) The multifaceted interplay between attention and multisensory integration. *Trends Cogn Sci* 14:400–410. [CrossRef Medline](#)
- Thirion B, Duchesnay E, Hubbard E, Dubois J, Poline JB, LeBihan D, Dehaene S (2006) Inverse retinotopy: inferring the visual content of images from brain activation patterns. *Neuroimage* 33:1104–1116. [CrossRef Medline](#)
- Tian B, Reser D, Durham A, Kustov A, Rauschecker JP (2001) Functional specialization in rhesus monkey auditory cortex. *Science* 292:290–293. [CrossRef Medline](#)
- Vroomen J, Bertelson P, de Gelder B (2001) The ventriloquist effect does not depend on the direction of automatic visual attention. *Percept Psychophys* 63:651–659. [CrossRef Medline](#)
- Wallace MT, Ramachandran R, Stein BE (2004) A revised view of sensory cortical parcellation. *Proc Natl Acad Sci U S A* 101:2167–2172. [CrossRef Medline](#)
- Werner S, Noppeney U (2010a) Distinct functional contributions of primary sensory and association areas to audiovisual integration in object categorization. *J Neurosci* 30:2662–2675. [CrossRef Medline](#)
- Werner S, Noppeney U (2010b) Superadditive responses in superior temporal sulcus predict audiovisual benefits in object categorization. *Cereb Cortex* 20:1829–1842. [CrossRef Medline](#)