Behavioral/Cognitive

A Trade-Off between Somatosensory and Auditory Related Brain Activity during Object Naming But Not Reading

Mohamed L. Seghier,[1] Thomas M.H. Hope,[1] Susan Prejawa,[1] 'Ōiwi Parker Jones,[1,2] Melanie Vitkovitch,[3] and Cathy J. Price[1]

[1]Wellcome Trust Centre for Neuroimaging, Institute of Neurology, UCL, London WC1N 3BG, United Kingdom, [2]Wolfson College, University of Oxford, Oxford OX2 6UD, United Kingdom, and [3]School of Psychology, University of East London, Water Lane, London E15 4LZ, United Kingdom

The parietal operculum, particularly the cytoarchitectonic area OP1 of the secondary somatosensory area (SII), is involved in somato-sensory feedback. Using fMRI with 58 human subjects, we investigated task-dependent differences in SII/OP1 activity during three familiar speech production tasks: object naming, reading and repeatedly saying "1-2-3." Bilateral SII/OP1 was significantly suppressed (relative to rest) during object naming, to a lesser extent when repeatedly saying "1-2-3" and not at all during reading. These results cannot be explained by task difficulty but the contrasting difference between naming and reading illustrates how the demands on somatosensory activity change with task, even when motor output (i.e., production of object names) is matched. To investigate what determined SII/OP1 deactivation during object naming, we searched the whole brain for areas where activity increased as that in SII/OP1 decreased. This across subject covariance analysis revealed a region in the right superior temporal sulcus (STS) that lies within the auditory cortex, and is activated by auditory feedback during speech production. The tradeoff between activity in SII/OP1 and STS was not observed during reading, which showed significantly more activation than naming in both SII/OP1 and STS bilaterally. These findings suggest that, although object naming is more error prone than reading, subjects can afford to rely more or less on somatosensory or auditory feedback during naming. In contrast, fast and efficient error-free reading places more consistent demands on both types of feedback, perhaps because of the potential for increased competition between lexical and sublexical codes at the articulatory level.

Key words: functional MRI; naming; parietal operculum; reading; somatosensory cortex; speech production

Introduction

Speech production is supported by a complex multilevel neural system which includes somatosensory and auditory processing of the spoken word during and after articulation. Activation related to somatosensory processing during articulation is observed in the lateral posterior part of the secondary somatosensory cortex (Guenther et al., 2006) known as secondary somatosensory area (SII)/OP1 (Eickhoff et al., 2006a), where responses increase when articulations are longer (Baciu et al., 2002; Shuster and Lemieux, 2005), slower (Binder et al., 2005), or less familiar (Wilson et al., 2009). Interestingly, SII/OP1 responses are suppressed during

some speech production tasks (Borden, 1979; Keller, 1987; Bookheimer et al., 1995; Dhanjal et al., 2008). This has been interpreted as an efficient adaptation process (Eliades and Wang, 2008) because it maximizes the somatosensory response when speech production becomes error prone, for example, in noisy environments, during perturbed feedback, or when producing novel or less familiar articulations (Golfinopoulos et al., 2011; Simmonds et al., 2011). The mismatch between the experienced somatosensory response and that expected from the intended speech output can then be used for on-line error correction.

Our study pursued an answer to the following question: if SII/OP1 activation increases in error-prone contexts, why is it greater for reading object names than naming the same objects in pictures (Bookheimer et al., 1995), given that reading is faster and less error-prone than object naming? One possibility is that more somatosensory activity for reading might reflect: (1) greater articulatory and acoustic variability in the presence of phonemic competition (McMillan et al., 2009; McMillan and Corley, 2010), particularly when written words have conflicting lexical and sub-lexical motor commands (Hennessey and Kirsner, 1999); (2) more attention to somatosensory responses to ensure that any acoustic variability in the output is minimized; or (3) suppression of somatosensory processing during object naming because, in the absence of conflicting sublexical influences, somatosensory activity can be predicted at the point of lexical retrieval, and is not required to fine tune the acoustic output. To summarize, we are

proposing that precision in the speech output is primarily controlled at an articulatory/somatosensory stage during reading, but an earlier lexical retrieval stage during object naming. Reduced somatosensory activity during more semantic tasks would explain why SII/OP1 activity is less for narrative speech than counting (Dhanjal et al., 2008) or syllable repetition (Brownsett and Wise, 2010).

To investigate which brain areas might be influencing somatosensory activation, we searched the whole brain for regions where activation increased or decreased, across subjects, in proportion to that in the secondary somatosensory cortex. We anticipated that the identified areas might be: (1) left frontal, middle temporal, or posterior parietal areas involved in the semantic mediation of speech production (Hope et al., 2014); (2) primary somatosensory regions of the postcentral gyrus where tactile and proprioceptive representations of articulation are likely to be coded (see discussion by Guenther et al., 2006); and/or (3) the auditory cortices involved in auditory feedback (Zheng et al., 2013), given the tradeoff between somatosensory and auditory feedback that has been observed in a behavioral study when both auditory and somatosensory feedback were altered (Lametti et al., 2012).

Materials and Methods

This study was approved by the London Queen Square Research Ethics Committee. All subjects gave written informed consent to participate in this study.

Experiment 1

Subjects. Fifty-eight right-handed native English speakers participated (33 females, 25 males, aged 29.6 ± 19 years). They had normal or corrected-to-normal vision, and had no history of neurological or psychiatric disorders.

Experimental design. All stimuli were derived from a set of 192 objects with three to six letter common names with regular spelling to sound relationships: 33 objects had three letter names (cat, bus, hat), 65 had four letter names (ship, bell, frog, hand), 58 had five letter names (teeth, camel, snake), and 36 had six letter names (spider, dagger, button). During two separate scanning sessions, subjects were asked to: (1) name pictures of familiar objects, (2) read aloud their written names, and (3) repeatedly say "1-2-3" in response to meaningless pictures of unfamiliar symbols or unfamiliar non-objects. During each of the two sessions, there were four different blocks of each condition, with 12 stimuli per block presented as triads (3 stimuli at a time) every 4.5 s (i.e., total duration per block = 18 s). To be able to characterize both activations and deactivations in the parietal operculum, six blocks of fixation (14.4 s per block) were also included. For the reading and object naming conditions, triads of stimuli were constructed where there was no obvious semantic relationship between the three different items in the triad (e.g., slide, axe, cup). Accuracy of vocal responses during all conditions was recorded with a MRI-compatible microphone. To minimize artifacts from head motion and airflow caused by the mouth opening and closing, subjects were instructed to produce vocal responses with minimal mouth movement. A sound cancellation system allowed us to identify the accuracy of vocal response. Stimulus presentation was via a video projector, a front-projection screen, and a system of mirrors fastened to a head coil. Additional details about the paradigm and stimuli can be found in our previous work (cf. Josse et al., 2008; Seghier et al., 2010).

MRI acquisition. Experiments were performed on a 1.5T Siemens system (Siemens Medical Systems). Functional imaging consisted of an EPI GRE sequence (TR/TE/Flip = 3600 ms/50 ms/90°, FOV = 192 mm, matrix = 64 × 64, 40 axial slices, 2 mm thick with 1 mm gap). Functional scanning was always preceded by 14.4 s of dummy scans to ensure tissue steady-state magnetization. Anatomical T1-weighted images were acquired using a three-dimensional modified driven equilibrium Fourier transform sequence (TR/TE/TI = 12.24 ms/3.56 ms/530 ms, matrix = 256 × 224, 176 sagittal slices with a final resolution of 1 mm^3).

fMRI data analysis. Data processing and statistical analyses were performed with the Statistical Parametric Mapping SPM5 software package (Wellcome Trust Centre for Neuroimaging, London UK; http://www.fil.ion.ucl.ac.uk/spm/). All functional volumes were spatially realigned, unwarped, normalized to MNI space using the unified normalization-segmentation procedure of SPM5, and smoothed with an isotropic 6 mm FWHM Gaussian kernel, with resulting voxel size of 2 × 2 × 2 mm^3. Time-series from each voxel were high-pass filtered (1/128 Hz cutoff) to remove low-frequency noise and signal drift. The preprocessed functional volumes of each subject were then submitted to a fixed-effects analysis, using the general linear model at each voxel. Each stimulus onset was modeled as an event in condition-specific "stick-functions" with a duration of 4.32 s per trial and a stimulus onset interval of 4.5 s. The resulting stimulus functions were convolved with a canonical hemodynamic response function, which provided regressors for the linear model. For each subject, we generated a contrast image that summarizes the activation level during each condition relative to fixation.

Parcellation of the parietal operculum. The SII region is mainly located on the upper bank and parietal operculum of the Sylvian sulcus, and is defined as the ventral part of a large sensory association cortex of the parietal lobe (Caselli, 1993). Several subdivisions (or areas) have previously been identified in SII (Disbrow et al., 2000; Eickhoff et al., 2007; Hinkley et al., 2007). Here we defined these areas using the probabilistic cytoarchitectonic maps that are available in the MNI space within the Anatomy Toolbox in SPM8 (v1.7; Eickhoff et al., 2005). Specifically, four parietal opercular areas in each hemisphere, labeled OP1, OP2, OP3, and OP4, are included in the toolbox (Eickhoff et al., 2007). Areas OP1, OP3, and OP4, but not OP2, are considered to be part of human SII (see discussion by Eickhoff et al., 2007) although other work has shown significant somatosensorial responses in OP2 as well (Burton et al., 2008).

OP1 is located at the lateral-posterior part of the parietal operculum and is equivalent to area SII of nonhuman primates. OP4 is located at the lateral-anterior parietal operculum and immediately ventral to the primary somatosensory region and is considered to be the homolog of the parietal-ventral area of nonhuman primates. OP3 is located at mesial-anterior parietal operculum and is considered to be equivalent to the ventral somatosensory area of nonhuman primates. OP2 is located at the mesial-posterior part of the parietal operculum and is considered to be the homolog of the parieto-insular vestibular cortex of nonhuman primates (Eickhoff et al., 2006b). Figure 1 illustrates the location of SII in respect to other anatomical landmarks and its different subdivisions.

It is worth noting that previous studies have used functional localizers to identify relevant SII voxels, for example using voluntary jaw/tongue movement tasks (Dhanjal et al., 2008; Simmonds et al., 2011; Geranmayeh et al., 2012). Here we investigated signal change in the whole SII region, as defined anatomically above. Although all our analyses were performed across the whole brain, we were particularly interested in speech production related signal changes in subdivisions OP1, OP3, and OP4 (Fig. 1, voxels shown in red).

Movement artifacts and data quality. It has been shown that speech production can cause artifacts during fMRI data acquisition (Yetkin et al., 1996; Birn et al., 1998). We have therefore optimized our procedures to minimize their impact. This included instructing all subjects to move their mouth as little as possible while speaking, the use of short block durations of 18 s (Soltysik and Hyde, 2006), unwarping during data preprocessing to correct artifacts caused by the interaction between head motion and geometric distortion (Andersson et al., 2001), and excluding any session with any motion parameter more than the size of 1 voxel (3 mm). Under these methodological conditions, we were able to identify robust activations during speech production; for a discussion see Heim et al. (2006).

We also conducted a systematic examination of motion parameters during each condition/task. This *ad hoc* analysis aimed to test whether the experimental conditions (reading, naming, and saying "1-2-3") differed systematically in their susceptibility to head motion artifacts. During realignment, each EPI image was transformed via a six-parameter rigid-body transformation, three translation and three rotation parameters, so as to be as similar as possible to the first image in the time series. Using these parameters after detrending, we computed an interscan displace-
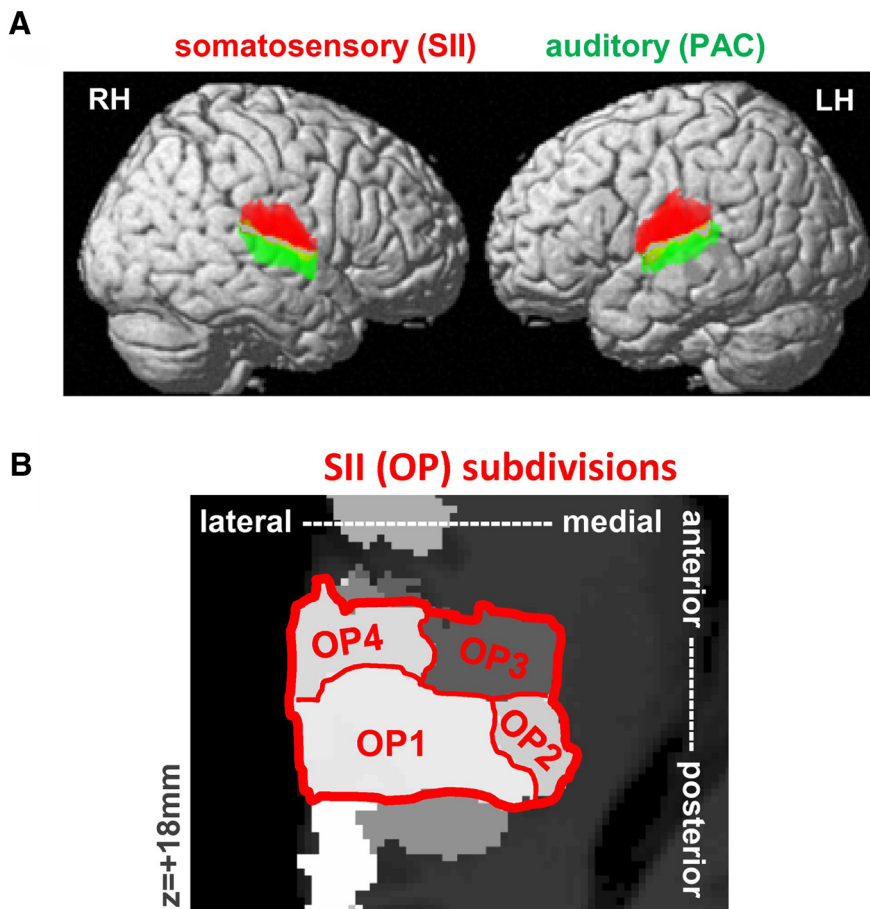
A

somatosensory (SII)   auditory (PAC)

B

SII (OP) subdivisions

Figure 1. *A*, rendering of the secondary somatosensory region (SII in red) and the primary auditory cortex (PAC, green) as defined from the Anatomy toolbox. *B*, illustrates the four opercular areas OP1–OP4 as defined from the cytoarchitectonic maps.

significant deactivation in the somatosensory cortex SII was extracted in each individual subject for each condition (naming, reading and repeatedly saying "1-2-3"). These effects were then entered as regressors in a repeat of the second-level analysis described above. This allowed us to search the whole brain for voxels where activation varied in proportion to that in the seed region (in SII), and to determine how such a relationship depended on the type of task. The aim was to identify regions that were functionally related to the seed region; for a similar rationale, see Seghier et al. (2008).

In-scanner behavioral data. Although we were able to monitor and record in-scanner accuracy both on line (during scanning) and off line (checking the recordings), the recording apparatus was not set up to measure speech latencies (RTs). It was started manually by the operator and did not indicate the exact start of each scan or the onset of the stimulus. We can therefore only measure time from the onset of one vocal response to another. This is not informative because the interstimulus interval was kept constant. Thus, faster response times from stimulus onset to speech onset were cancelled out by more time from speech onset to next stimulus. Last but not least, the stimuli were grouped into triads (3 stimuli per trial/ event) and thus speech onset for each individual stimulus could not be extracted. To examine whether and how the effects identified in the first experiment (above) were modulated by in-scanner response times, we included data from a second Experiment (see Experiment 2: validation in an independent sample with in-scanner RTs, below) that also allowed us to replicate the results of the main experiment, in a new cohort of participants.

ment (ISD) as an average path length (D'Esposito et al., 1999) for each condition, with ISD computed separately for translation and rotation parameters (Yoo et al., 2005). An ISD value reflects the average 3D Euclidean distance (in mm) between two consecutive EPI volumes. For each subject, ISD values were averaged over the two sessions for each task (naming, reading, saying "1-2-3" and fixation). As illustrated in Figure 2, ISD values were comparable across tasks, although slightly higher for naming. The ISD values were strongly correlated between tasks (e.g., $r = 0.87$ between naming and reading), which means that when subjects moved a lot during one condition they also tended to move a lot during the other conditions. Perhaps more importantly, the amount of head motion during each task (Fig. 2) did not reflect or mirror the task-dependent activation in SII regions (see Results, below).

Second-level condition-dependent analyses. The contrast images of each of our three speech production conditions (i.e., naming, reading, and repeatedly saying "1-2-3") relative to fixation, from each of the 58 subjects' first-level analyses, were entered into a second-level ANOVA analysis (i.e., random-effects analysis). From this second level analysis, we generated statistical parametric maps of the t statistic at each voxel SPM{t} for condition-dependent activation differences with the expectation that we would observe deactivation for speaking relative to fixation. Statistical comparisons are reported at a threshold of $p < 0.05$ FWE-corrected for height across the whole brain. We investigated the influence of age on these effects because previous work has shown that speech production may differ between old and young adults (Sörös et al., 2011; for review, see Mortensen et al., 2006).

Second-level covariance analysis. The aim of this analysis was to search over the whole brain for regions where activation increased across our 58 subjects in proportion to deactivation in the somatosensory cortex. Practically, the effect size (parameter estimate) in the voxel showing the most

Experiment 2: validation in an independent sample with in-scanner RTs

This experiment aimed to test:(1) whether brain responses in bilateral SII varied with speech latencies (Do faster subjects suppress SII more?), and (2) whether the strength of the covariance between SII and other brain regions interacted with RTs (Is covariance strongest in subjects with faster latencies?). The participants were 55 healthy subjects (age: 43 ± 17 years, 23 males, 32 females) in whom in-scanner RTs were measured during overt object naming. The subjects in this new sample were older than our original 58 subjects ($t = 4.0$, $p < 0.001$, df = 111). They were visually presented with pictures of two objects at a time. In the first scanning session, they decided whether the two objects were semantically related (e.g., pirate–boat) or unrelated (barrel–deer). In the second scanning session, the task was to overtly name both objects that were always semantically unrelated (e.g., pirate–deer). Only the data from the second scanning session are considered in the current paper. Data from both sessions have already been reported in a recent report (Sanjuán et al., 2015). There were four blocks of stimuli in each scanning session. Each block presented five different pairs of objects (at a rate of one pair every 5 s and stimulus duration of 1.5 s). Each block of object stimuli was followed by a block of fixation. In the second (naming) session of interest, half the objects had been seen in the previous (semantic decision) session but in a different pair. Repeated and novel items were presented in different blocks so that the effect of novelty/repetition could be investigated on object naming activation. The first and fourth blocks presented pictures of objects that had been seen in the previous (semantic decision) session. The second and third blocks presented pictures of objects that had not been seen in the previous session.

Stimulus onsets and the participant's spoken responses were recorded via a noise-cancelling MRI microphone (FOMRI IIITM, Optoacoustics). To compute the reaction times, we used an adaptive window moving average filter that was tailored to each subject. The optimal window length (i.e., the width that maximally smoothed the audio stream) was based on a portion of the audio file collected during rest. After smoothing the whole audio recording, we defined the onset of speech as a rise in the absolute amplitude of the smoothed audio stream beyond 1.5 SD from the mean.

Functional images were acquired from a Siemens 3T scanner, with a 12-channel head coil, using a gradient-echo EPI sequence (TR/TE = 3080 ms/30 ms; flip angle = 90°; matrix size = 64 × 64; FOV = 192 × 192 mm, slice thickness = 2 mm, interslice gap = 1 mm). Each functional run consisted of 66 scans, including five "dummy scans" to allow for T1 equilibration effects. All data analyses were conducted using standard routines, as in the first experiment, using SPM12. First level, subject specific, fixed-effect analyses modeled each stimulus onset as a single event. For the naming session, there were four regressors of interest, which corresponded to the correct responses in each of the four stimulus blocks. Three extra regressors modeled the instructions, incorrect or "other" (self-corrected, delayed, or no-response) trials. Contrast images for the four regressors of interest (correct trials > fixation averaged across block) were then entered into the second level analysis, which focused on replicating the effects observed in the first experiment. We focused on whether the BOLD response in bilateral SII was deactivated. Then we repeated the same analysis, using the BOLD signal in the deactivated SII area as a covariate, specifically checking whether decreases in SII activity covaried with increased activity in the same region(s) as the first experiment. Having replicated the effects of the first experiment, we could then investigate whether they correlated with in-scanner response times.

Results

Experiment 1: in-scanner performance

In the first fMRI experiment, a response was considered correct when all three stimuli in a triad were read/named correctly. Over our 58 healthy controls, accuracy across sessions was on average 91 ± 9% for object naming, 99 ± 1% for word reading, and 100% for saying "1-2-3," with subjects making more errors during naming than reading (Wilcoxon signed rank test: $p < 0.001$). Increased errors for object naming could be a consequence of: (1) timing constraints because the interstimulus interview was not sufficient for 3 different naming responses, (2) failures to select a lexical phonological representation related to the picture, or (3) in-accurate pronunciation. A detailed analysis of performance during object naming revealed that 83% of all naming errors corresponded to trials where subjects failed to name the third stimulus (36%) or the third response was uncertain or incorrect (47%). In all these cases, the other two stimuli were correctly named. This suggests that the majority of the errors during object naming were the consequence of time constraints. Indeed, it takes longer to retrieve phonological representations from the semantic content of pictures than to read aloud written words
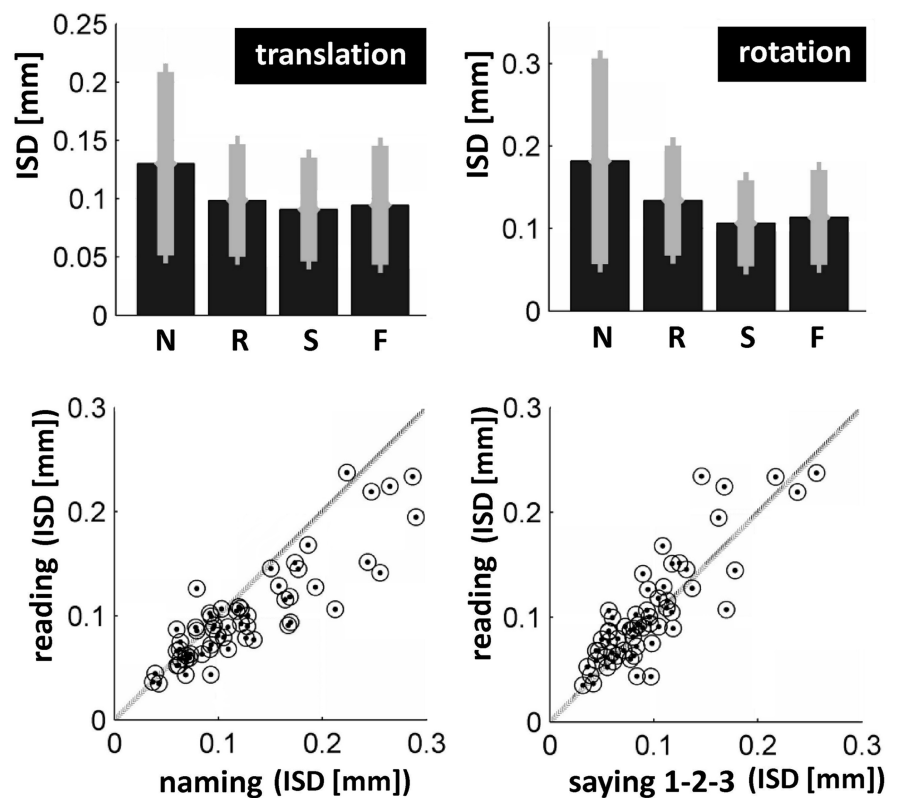

Figure 2. Top, Bar graphs of the mean ISD (in mm) over subjects for each task when using detrended translation and rotation parameters that are derived from the six-parameter rigid-body transformation during realignment. Error bars represent the SD. N, Naming; R, reading; S, saying "1-2-3"; F = fixation. Bottom, A scatter plot of ISD for reading against naming (left) and reading against saying "1-2-3" (right). This is based on the translation parameters; a similar profile for rotation parameters was also observed (data not shown here). The ISD values were strongly correlated between tasks ($r > 0.8, p < 0.001$). Each dot represents the ISD value of one subject. The solid gray line represents the diagonal "y = x."

that are facilitated by nonsemantic links between orthography and phonology.

Experiment 1: deactivations confined to bilateral SII/OP1

As predicted, speech production decreased activation in bilateral SII relative to fixation (Fig. 3). This was observed in a large lateral cluster in the inferior parietal cortex (cytoarchitectonic areas PF/PFop) and the postcentral gyrus extending ventrally to area OP1 at $x = -60, y = -28; z = 24$ and $x = 60, y = -26, z = 30$, in left and right hemispheres, respectively (Fig. 3B, coronal slices). These coordinates correspond to those that Guenther et al. (2006) associated with the detection and processing of mismatches between expected and actual somatosensory processing (cf. Guenther et al. (2006), their Table 1 at MNI: $x = -62, y = -28, z = 32; x = 66, y = -24, z = 35$). The extension of the effects beyond the cytoarchitectonic maps (Eickhoff et al., 2006a) in the dorsal and posterior direction is consistent with that reported in (Dhanjal et al., 2008).

No other deactivations were seen in other parietal opercular areas, including OP3 and OP4 even at lower thresholds. Thus, somatosensory deactivation relative to fixation was limited to lateral SII/OP1 foci, bilaterally. Figure 3C illustrates the task effects at these SII/OP1 foci, demonstrating that deactivation was significant during naming (Z-score = 4.6 and 5.4, for left and right OP1, respectively) and for repeatedly saying "1-2-3" (Z-score = 4.8 and 3.2, for left and right OP1, respectively) but not during reading aloud. The difference between (de)activation during naming and reading (i.e., "reading > naming") at the same

A

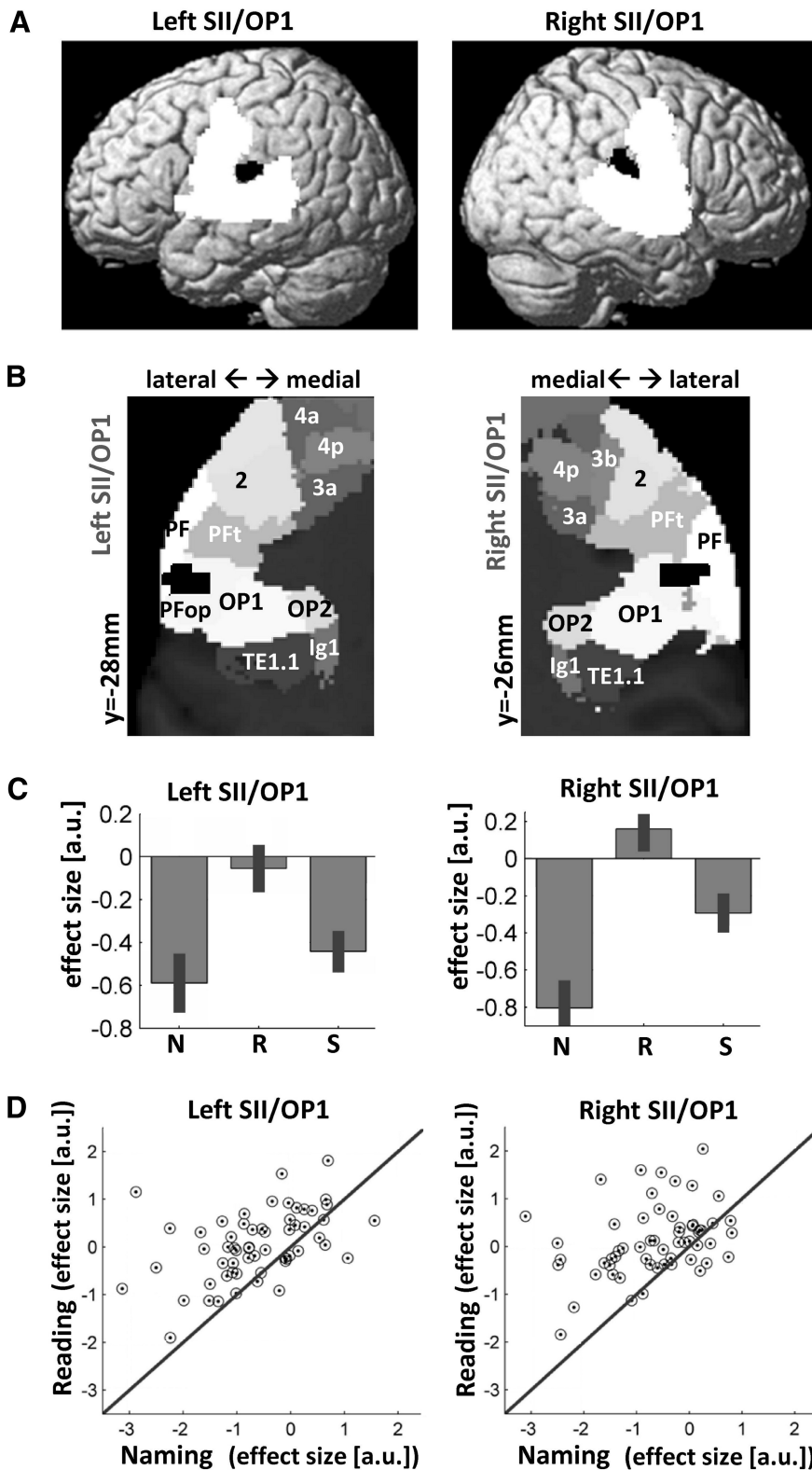**Left SII/OP1**   **Right SII/OP1**

B

C

D

Figure 3. *A*, A 3D-rendering illustrating the location of the suppressed responses in left and right SII/OP1 (black clusters) in relation to the activated production system (in white, defined during the simple production task of saying "1-2-3."). *B*, Projection of the peaks (at *p* < 0.05 FWE-corrected) of the deactivated clusters (in black) on the cytoarchitectonic maps. The abbreviations of the different areas are defined from the Anatomy toolbox. *C*, The bar graphs show the effect size at each peak in left and right SII/OP1 during the three production tasks. N, Naming; R, reading; S, saying "1-2-3." *D*, A scatter plot showing the signal in SII/OP1 being consistently more suppressed during naming than reading in the majority of subjects (i.e., each point represents 1 subject, and the majority subjects falls above the diagonal line shown in black). a.u., Arbitrary unit.

OP1 coordinates was also highly significant: Z-score = 4.7 and 5.6, in left and right OP1 respectively, *p* < 0.05 FWE-corrected). This task-specific effect was highly consistent at the individual subject level (Fig. 3*D*) and was not significantly influenced by age (*p* > 0.05 uncorrected).

Experiment 1: strong negative covariance between SII/OP1 and the right auditory cortex

The second-level covariance analysis used both left and right SII/OP1 of Figure 3 as seed regions for naming and reading tasks. The only region to show significant across-subject negative covariance with left OP1 was located in the right auditory cortex in the superior temporal sulcus (STS) at [*x* = 56, *y* = −18, *z* = −8]; see Figure 4*A*. The significant negative covariance with left OP1 that was observed during object naming (Z-score = 5.1, *p* < 0.05 FWE-corrected for multiple comparisons across the whole brain) was not observed during either reading or repeatedly saying "1-2-3" (*p* > 0.05 uncorrected). The effect of task (object naming versus reading) on the relationship between left OP1 and auditory cortex was significant at *p* < 0.001 uncorrected (Z-score = 3.5). When the seed region was located in right OP1, there was no significant relationship with activity in either left or right STS (*p* > 0.05 uncorrected). Figure 4 illustrates the task-dependent activation in left and right STS, as well as the relationship between activation in each SII/OP1 region and each STS region during naming and reading.

Experiment 2: validation in an independent sample with in-scanner RTs

In-scanner RTs did not vary across blocks of the naming session (*p* > 0.05 uncorrected). Area SII (−48, −25, 23; Z-score = 5.4) was most deactivated in the context of the first block of novel items (Block 2) and this effect survived significance at *p* < 0.05 FWE-corrected across the whole brain. Negative covariance between SII and right auditory cortex was also strongest for Block 2 with a local peak at (51, −22, −10) where activity increased as that in left SII decreased (Z-score = 2.84), and this cluster was also highly activated during object naming (Z-score = 4.7).

To investigate how these effects were related to RTs, we correlated mean correct naming RTs for Block 2 only (in-scanner accuracy = 94%, in-scanner naming RTs = 1.29 ± 0.2 s) with activation in left SII and right STS. We also compared the strength of covariance between these regions for the 28 subjects with faster RTs (range, 0.95–1.28 s)

and the 27 subjects with slower RTs (range, 1.31–1.66 s) using a two-tailed Z-test performed on the Fisher's Z-transformed correlations. The effect of RTs was not significant in any of these analyses, though we note that there was a trend ($p = 0.14$, two tailed) for covariance to be stronger across participants with faster RTs.

Discussion

This study sought to explain why more somatosensory activation has been reported for tasks that are less error prone when the prevailing theories of somatosensory function predict more activation for more error prone tasks (Borden, 1979; Keller, 1987; Ventura et al., 2009; Golfinopoulos et al., 2010; Behroozmand and Larson, 2011; Chang et al., 2013). First, we replicated prior evidence (Bookheimer et al., 1995) that somatosensory activity in SII/OP1 is higher for reading (less error prone) than object naming (more error prone). We also found that SII/OP1 activity for repeatedly saying "1-2-3" was less than reading but more than naming. This cannot simply be explained by the demands on semantic processing, which are stronger for naming and reading than saying "1-2-3." Nor can it be explained by "automaticity," which is greatest for saying "1-2-3" and least for naming.

To investigate the brain regions, and their processing functions, that might explain SII/OP1 activation, we searched the whole brain for areas where activation increased or decreased proportionately with somatosensory activity during object naming. This allowed us to explicitly test whether SII/OP1 activity was: (1) inversely related to areas involved in the semantic mediation of speech production, consistent with SII/OP1 suppression being greatest when semantic mediation was high; (2) positively related to SI of the postcentral gyrus (Guenther et al., 2006); and/or (3) inversely related to activity in the auditory cortices as would be expected if there was a tradeoff between somatosensory and auditory feedback (Lametti et al., 2012).

This covariance analysis revealed one highly significant effect in the right STS, where activation was higher during object naming when activation in left SII/OP1 was lower. This relationship was not observed during reading or saying "1-2-3." The same right STS region (+54, −18, −6) has been associated with auditory feedback during speech production (Zheng et al., 2010) in an experiment that found increased activation at these coordinates when auditory processing was perturbed by introducing conflict between the expected auditory input (i.e., the speech production response) and the actual auditory input (noise). Our covariance analysis therefore provided evidence that SII/OP1 activity during object naming is
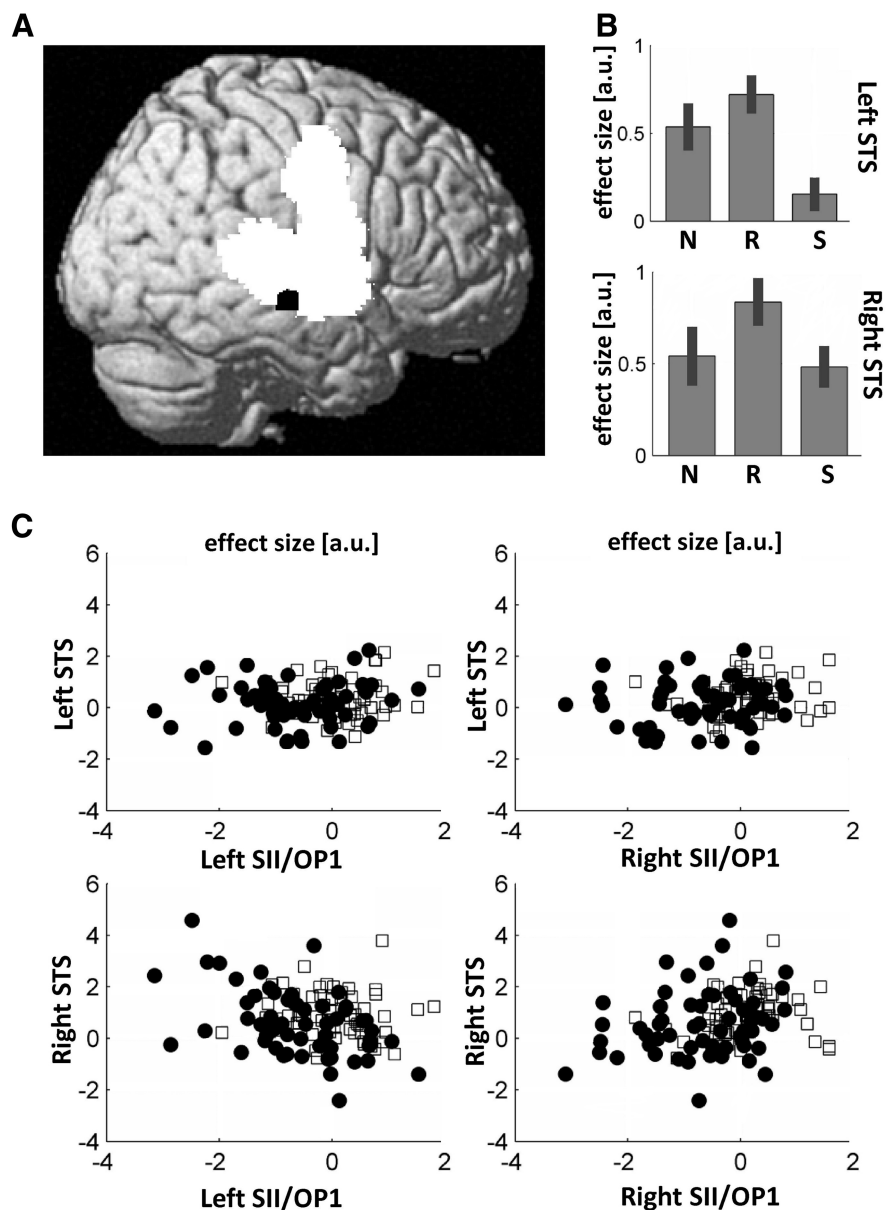
Figure 4. *A*, Illustrates the location of the right STS cluster (black voxels) where activation increased in proportion to deactivation in SII/OP1. The white surround shows activation related to speech production (saying "1-2-3"). *B*, The bar graphs illustrate the responses profile in both left and right STS during the three production tasks. N, Naming; R, reading; S, saying "1-2-3." As shown, activation is higher for reading than naming in both right and left STS (Z-score = 2.3 and 1.8, respectively). *C*, Illustrates the scatter plots of the signal in left and right SII/OP1 against the signal in left and right STS during naming (black dots) and reading (open squares). Each point in the scatter plots represents one subject (i.e., the scatter plots are for illustration purposes only). A significant relationship across subjects was only observed between left SII/OP1 and right STS during naming. There is no clear relationship between the signal in left SII/OP1 and left STS or between right SII/OP1 and both left and right STS, regardless of task. a.u., Arbitrary unit.

more suppressed when auditory responses are high (Lametti et al., 2012). In contrast, we did not find evidence that SII/OP1 suppression was related to activity in any other areas including those previously associated with semantic processing or motor output.

The tradeoff between auditory and somatosensory feedback during speech production was previously demonstrated in behavioral experiments that altered somatosensory and auditory feedback while subjects repeated a simple speech utterance (Feng et al., 2011; Lametti et al., 2012). Interestingly, the amount of compensation for each perturbation depended on the preferen-

tial reliance that individuals showed for either somatosensory, auditory feedback or both during speech production (Lametti et al., 2012, their Fig. 8). Those who were more reliant on auditory feedback failed to adapt to the somatosensory disturbances, while those who were more reliant on somatosensory feedback ignored the auditory perturbation. Lametti et al.'s (2012) behavioral study is therefore consistent with what we observe here: subjects with low activation in a somatosensory area (SPII/OP1) had high activation in an auditory processing area (right STS), whereas those with high SPII/OP1 activation showed low activation in right auditory cortex. The three key differences between studies are that we: (1) provide the neuroanatomical substrates for the previously reported behavioral effect, (2) demonstrate that this compensatory relationship is task-dependent because it is observed during object naming but not during reading or saying "1-2-3," and (3) observe a tradeoff between somatosensory and auditory activity even when feedback was not experimentally distorted as by Lametti et al. (2012).

As we did not find any evidence, in Experiment 2, that naming response times were faster or slower with higher or lower SII/OP1 activity or with stronger covariance between left SII/OP1 and right STS, we have no indication of which type of feedback strategy might be more efficient (i.e., somatosensory, auditory, or both). Plausibly the strategy used depends on individual differences, with no noticeable effect on processing speed. Alternatively, (1) between-subject variance in mean naming responses may have been insufficiently precise to detect differences in speech performance at the individual stimulus level, or (2) speech performance may have been at ceiling for all our healthy native-speaking participants. We therefore plan, in future, to explore how the somatosensory–auditory processing tradeoff during naming is influenced by between group differences in speech proficiency; e.g., those speaking in a first or second language or those who have developmental delay or acquired damage to the speech production network. We could also investigate how somatosensory and auditory processing (in SII and right STS) tradeoff within an individual subject during the naming task, and whether such a tradeoff influences response times in that individual. The neural mechanisms supporting such a tradeoff could be investigated with effective connectivity analyses that measure how the different regions influence each other and how these inter-regional interactions are modulated by response times or task.

With respect to the neuroanatomical substrates for the tradeoff in somatosensory and auditory activation, we observed a striking lateralized and cross-hemisphere relationship. Namely, activity in left SII/OP1 was inversely correlated with activity in right STS, while there was an absence of such a relationship between left SII/OP1 with left STS, right SII/OP1 with right STS or right SII/OP1 with left STS (Fig. 4C). Although further investigations are required to understand such lateralized effects, we speculate that the right hemisphere (STS and SII/OP1) might be more involved when speech is predictable whereas the left hemisphere is more involved when speech is unfamiliar or needs monitoring. In the auditory cortices, evidence for this conjecture comes from a recent study (Ylinen et al., 2014) that found activity in the right but not left auditory cortex to be suppressed when the same auditory stimulus is repeated during covert speech rehearsal. In contrast, activity in the left auditory cortex increased when there was a mismatch between the auditory stimulus and covert speech. In SII, evidence comes from findings that left SII at $x = -68$, $y = -26$, $z = 30$ is more important for processing unfamiliar pseudowords than familiar words (Roux et al., 2012), whereas right

SII ($x = 64$, $y = -28$, $z = 34$) is more activated (less suppressed) for words than pseudowords (Mechelli et al., 2003). Thus in both studies, activation in the left hemisphere was increasing with unexpected/unfamiliar feedback, whereas activation in the right hemisphere was increasing with predictable feedback.

Putting these results together with our own findings supports the following hypothesis: when left hemisphere somatosensory cortex increases its response to aberrations in speech output, the auditory signal becomes less relevant which decreases the response in right STS. In contrast, when the left somatosensory cortex is less responsive, the auditory signal becomes more relevant which increases responses in right STS. This relationship between STS and SII/OP1 is anatomically plausible because it can be mediated by many heteromodal connections that exist between auditory regions and SII (Cappe and Barone, 2005; Hackett et al., 2007; Smiley and Falchier, 2009); however, further studies are required to determine how the relationship between SII/OP1 and STS is established. Finally, with respect to why the tradeoff between activity in somatosensory and auditory feedback areas was detected during naming but not reading, we note that both systems (bilateral SII/OP1 and bilateral STS) were more activated for reading than naming (Figs. 3C, Fig. 4B). This suggests that both systems were engaged for fast and efficient reading but naming involved variable contributions from one system or another. Specifically, reading may need to engage both feedback systems to resolve competition between conflicting phonemic information that can potentially arise at the articulation level from coactivation of lexical and sublexical orthographic to phonological mappings. By contrast, for object naming, somatosensory activity can be predicted at the point of lexical retrieval, probably because monitoring may primarily take place at the lexical level. The level of phonemic competition during reading and naming may also depend on the context (e.g., naming novel and repeated items in Experiment 2). Reading in Experiment 1 may also have been particularly susceptible to phonemic competition because it required the production of three unrelated words in rapid succession. By measuring acoustic and articulatory variation in speech responses (McMillan et al., 2009; McMillan and Corley, 2010), future studies can provide insight into the different task- and context-dependent monitoring mechanisms.

In conclusion, our findings demonstrated that task-dependent responses in SII/OP1 may reflect the contribution of many different factors, the most significant of which is the relative contribution of both somatosensory and auditory feedback mechanisms. In particular, we show that both feedback systems are more engaged by reading than picture naming even when the motor outputs are controlled, and during picture naming different individuals show a preference for using one system over the other. Future work needs to investigate how both suppressed and nonsuppressed parietal opercular areas interact during speech production, how SII/OP1 interact with other subnetworks of the feedback system (Parker Jones et al., 2013; Zheng et al., 2013; Simmonds et al., 2014), and how this would vary with learning, expertise, and experimental context. For instance, previous developmental studies have suggested more reliance on sublexical reading and greater feedback mechanisms in children who were learning to read compared with skilled adult readers (Ratner et al., 1964; Borden, 1979; Awaida and Beech, 1995; Green et al., 2002; Shiller et al., 2010; MacDonald et al., 2012). The relative involvement of the somatosensory and auditory feedback systems may also show substantial changes when the speech production system is damaged in adults.

References

Andersson JL, Hutton C, Ashburner J, Turner R, Friston K (2001) Modeling geometric deformations in EPI time series. Neuroimage 13:903–919. CrossRef Medline

Awaida M, Beech JR (1995) Children's lexical and sublexical development while learning to read. J Exp Edu 63:97–113. CrossRef

Baciu M, Ans B, Carbonnel S, Valdois S, Juphard A, Pachot-Clouard M, Segebarth C (2002) Length effect during word and pseudo-word reading: an event-related fMRI study. Neurosci Res Commun 30:155–165. CrossRef

Behroozmand R, Larson CR (2011) Error-dependent modulation of speech-induced auditory suppression for pitch-shifted voice feedback. BMC Neurosci 12:54. CrossRef Medline

Binder JR, Medler DA, Desai R, Conant LL, Liebenthal E (2005) Some neurophysiological constraints on models of word naming. Neuroimage 27:677–693. CrossRef Medline

Birn RM, Bandettini PA, Cox RW, Jesmanowicz A, Shaker R (1998) Magnetic field changes in the human brain due to swallowing or speaking. Magn Reson Med 40:55–60. CrossRef Medline

Bookheimer SY, Zeffiro TA, Blaxton T, Gaillard W, Theodore W (1995) Regional cerebral blood flow during object naming and word reading. Hum Brain Mapp 3:93–106. CrossRef

Borden GJ (1979) An interpretation of research of feedback interruption in speech. Brain Lang 7:307–319. CrossRef Medline

Brownsett SL, Wise RJ (2010) The contribution of the parietal lobes to speaking and writing. Cereb Cortex 20:517–523. CrossRef Medline

Burton H, Sinclair RJ, Wingert JR, Dierker DL (2008) Multiple parietal operculum subdivisions in humans: tactile activation maps. Somatosens Mot Res 25:149–162. CrossRef Medline

Cappe C, Barone P (2005) Heteromodal connections supporting multisensory integration at low levels of cortical processing in the monkey. Eur J Neurosci 22:2886–2902. CrossRef Medline

Caselli RJ (1993) Ventrolateral and dorsomedial somatosensory association cortex damage produces distinct somesthetic syndromes in humans. Neurology 43:762–771. CrossRef Medline

Chang EF, Niziolek CA, Knight RT, Nagarajan SS, Houde JF (2013) Human cortical sensorimotor network underlying feedback control of vocal pitch. Proc Natl Acad Sci U S A 110:2653–2658. CrossRef Medline

D'Esposito M, Zarahn E, Aguirre GK, Rypma B (1999) The effect of normal aging on the coupling of neural activity to the bold hemodynamic response. Neuroimage 10:6–14. CrossRef Medline

Dhanjal NS, Handunnetthi L, Patel MC, Wise RJ (2008) Perceptual systems controlling speech production. J Neurosci 28:9969–9975. CrossRef Medline

Disbrow E, Roberts T, Krubitzer L (2000) Somatotopic organization of cortical fields in the lateral sulcus of Homo sapiens: evidence for SII and PV. J Comp Neurol 418:1–21. CrossRef Medline

Eickhoff SB, Stephan KE, Mohlberg H, Grefkes C, Fink GR, Amunts K, Zilles K (2005) A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. Neuroimage 25:1325–1335. CrossRef Medline

Eickhoff SB, Amunts K, Mohlberg H, Zilles K (2006a) The human parietal operculum: II. Stereotaxic maps and correlation with functional imaging results. Cereb Cortex 16:268–279. CrossRef Medline

Eickhoff SB, Weiss PH, Amunts K, Fink GR, Zilles K (2006b) Identifying human parieto-insular vestibular cortex using fMRI and cytoarchitectonic mapping. Hum Brain Mapp 27:611–621. CrossRef Medline

Eickhoff SB, Grefkes C, Zilles K, Fink GR (2007) The somatotopic organization of cytoarchitectonic areas on the human parietal operculum. Cereb Cortex 17:1800–1811. CrossRef Medline

Eliades SJ, Wang X (2008) Neural substrates of vocalization feedback monitoring in primate auditory cortex. Nature 453:1102–1106. CrossRef Medline

Feng Y, Gracco VL, Max L (2011) Integration of auditory and somatosensory error signals in the neural control of speech movements. J Neurophysiol 106:667–679. CrossRef Medline

Geranmayeh F, Brownsett SL, Leech R, Beckmann CF, Woodhead Z, Wise RJ (2012) The contribution of the inferior parietal cortex to spoken language production. Brain Lang 121:47–57. CrossRef Medline

Golfinopoulos E, Tourville JA, Guenther FH (2010) The integration of large-scale neural network modeling and functional brain imaging in speech motor control. Neuroimage 52:862–874. CrossRef Medline

Golfinopoulos E, Tourville JA, Bohland JW, Ghosh SS, Nieto-Castanon A, Guenther FH (2011) fMRI investigation of unexpected somatosensory feedback perturbation during speech. Neuroimage 55:1324–1338. CrossRef Medline

Green JR, Moore CA, Reilly KJ (2002) The sequential development of jaw and lip control for speech. J Speech Lang Hear Res 45:66–79. CrossRef Medline

Guenther FH, Ghosh SS, Tourville JA (2006) Neural modeling and imaging of the cortical interactions underlying syllable production. Brain Lang 96:280–301. CrossRef Medline

Hackett TA, Smiley JF, Ulbert I, Karmos G, Lakatos P, de la Mothe LA, Schroeder CE (2007) Sources of somatosensory input to the caudal belt areas of auditory cortex. Perception 36:1419–1430. CrossRef Medline

Heim S, Amunts K, Mohlberg H, Wilms M, Friederici AD (2006) Head motion during overt language production in functional magnetic resonance imaging. Neuroreport 17:579–582. CrossRef Medline

Hennessey NW, Kirsner K (1999) The role of sub-lexical orthography in naming: a performance and acoustic analysis. Acta Psychol (Amst) 103:125–148. CrossRef Medline

Hinkley LB, Krubitzer LA, Nagarajan SS, Disbrow EA (2007) Sensorimotor integration in S2, PV, and parietal rostroventral areas of the human sylvian fissure. J Neurophysiol 97:1288–1297. CrossRef Medline

Hope TM, Prejawa S, Parker Jones O, Oberhuber M, Seghier ML, Green DW, Price CJ (2014) Dissecting the functional anatomy of auditory word repetition. Front Hum Neurosci 8:246. CrossRef Medline

Josse G, Seghier ML, Kherif F, Price CJ (2008) Explaining function with anatomy: language lateralization and corpus callosum size. J Neurosci 28:14132–14139. CrossRef Medline

Keller E (1987) The cortical representation of motor processes of speech. In: Motor and sensory language processes (Keller E, Gopnik M, eds), pp 125–162. Hillsdale, NJ: Lawrence Erlbaum.

Lametti DR, Nasir SM, Ostry DJ (2012) Sensory preference in speech production revealed by simultaneous alteration of auditory and somatosensory feedback. J Neurosci 32:9351–9358. CrossRef Medline

MacDonald EN, Johnson EK, Forsythe J, Plante P, Munhall KG (2012) Children's development of self-regulation in speech production. Curr Biol 22:113–117. CrossRef Medline

McMillan CT, Corley M (2010) Cascading influences on the production of speech: evidence from articulation. Cognition 117:243–260. CrossRef Medline

McMillan CT, Corley M, Lickley RJ (2009) Articulatory evidence for feedback and competition in speech production. Lang Cogn rocess 24:44–66. CrossRef

Mechelli A, Gorno-Tempini ML, Price CJ (2003) Neuroimaging studies of word and pseudoword reading: consistencies, inconsistencies, and limitations. J Cogn Neurosci 15:260–271. CrossRef Medline

Mortensen L, Meyer AS, Humphreys GW (2006) Age-related effects on speech production: a review. Lang Cogn Process 21:238–290. CrossRef

Parker Jones O, Seghier ML, Kawabata Duncan KJ, Leff AP, Green DW, Price CJ (2013) Auditory-motor interactions for the production of native and non-native speech. J Neurosci 33:2376–2387. CrossRef Medline

Ratner SC, Rawronski JJ, Rice FE (1964) The variable of concurrent action in the language of children: effects of delayed speech feedback. Psychol Rec 14:47–56.

Roux FE, Durand JB, Jucla M, Réhault E, Reddy M, Démonet JF (2012) Segregation of lexical and sub-lexical reading processes in the left perisylvian cortex. PLoS One 7:e50665. CrossRef Medline

Sanjuán A, Hope TMH, Parker Jones O, Prejawa S, Oberhuber M, Guerin J, Seghier ML, Green DW, Price CJ (2015) Dissociating the semantic function of two neighbouring subregion in the left lateral anterior temporal lobe. Neuropsychologia. Advance online publication. CrossRef Medline

Seghier ML, Lee HL, Schofield T, Ellis CL, Price CJ (2008) Inter-subject variability in the use of two different neuronal networks for reading aloud familiar words. Neuroimage 42:1226–1236. CrossRef Medline

Seghier ML, Fagan E, Price CJ (2010) Functional subdivisions in the left angular gyrus where the semantic system meets and diverges from the default network. J Neurosci 30:16809–16817. CrossRef Medline

Shiller DM, Gracco VL, Rvachew S (2010) Auditory-motor learning during speech production in 9–11-year-old children. PLoS One 5:e12975. CrossRef Medline

Shuster LI, Lemieux SK (2005) An fMRI investigation of covertly and

overtly produced mono- and multisyllabic words. Brain Lang 93:20–31. CrossRef Medline

Simmonds AJ, Wise RJ, Dhanjal NS, Leech R (2011) A comparison of sensory-motor activity during speech in first and second languages. J Neurophysiol 106:470– 478. CrossRef Medline

Simmonds AJ, Wise RJ, Collins C, Redjep O, Sharp DJ, Iverson P, Leech R (2014) Parallel systems in the control of speech. Hum Brain Mapp 35: 1930–1943. CrossRef Medline

Smiley JF, Falchier A (2009) Multisensory connections of monkey auditory cerebral cortex. Hear Res 258:37– 46. CrossRef Medline

Soltysik DA, Hyde JS (2006) Strategies for block-design fMRI experiments during task-related motion of structures of the oral cavity. Neuroimage 29:1260–1271. CrossRef Medline

Sörös P, Bose A, Sokoloff LG, Graham SJ, Stuss DT (2011) Age-related changes in the functional neuroanatomy of overt speech production. Neurobiol Aging 32:1505–1513. CrossRef Medline

Ventura MI, Nagarajan SS, Houde JF (2009) Speech target modulates speaking induced suppression in auditory cortex. BMC Neurosci 10:58. CrossRef Medline

Wilson SM, Isenberg AL, Hickok G (2009) Neural correlates of word pro-

duction stages delineated by parametric modulation of psycholinguistic variables. Hum Brain Mapp 30:3596–3608. CrossRef Medline

Yetkin FZ, Haughton VM, Cox RW, Hyde J, Birn RM, Wong EC, Prost R (1996) Effect of motion outside the field of view on functional MR. Am J Neuroradiol 17:1005–1009. Medline

Ylinen S, Nora A, Leminen A, Hakala T, Huotilainen M, Shtyrov Y, Mäkelä JP, Service E (2014) Two distinct auditory-motor circuits for monitoring speech production as revealed by content-specific suppression of auditory cortex. Cereb Cortex. Advance online publication. CrossRef Medline

Yoo SS, Choi BG, Juh R, Pae CU, Lee CU (2005) Head motion analysis during cognitive fMRI examination: application in patients with schizophrenia. Neurosci Res 53:84–90. CrossRef Medline

Zheng ZZ, Munhall KG, Johnsrude IS (2010) Functional overlap between regions involved in speech perception and in monitoring one's own voice during speech production. J Cogn Neurosci 22:1770–1781. CrossRef Medline

Zheng ZZ, Vicente-Grabovetsky A, MacDonald EN, Munhall KG, Cusack R, Johnsrude IS (2013) Multivoxel patterns reveal functionally differentiated networks underlying auditory feedback processing of speech. J Neurosci 33:4339– 4348. CrossRef Medline