



Published in final edited form as:

Neuroimage. 2007 May 15; 36(1): 131–143.

Human Brain Activation during Phonation and Exhalation: Common Volitional Control for Two Upper Airway Functions

Torrey M. J. Loucks, Ph.D.¹, Christopher J. Poletto, Ph.D.¹, Kristina Simonyan, M.D., Ph.D.¹, Catherine L. Reynolds, B.Sc.¹, and Christy L. Ludlow, Ph.D.¹

1 Laryngeal & Speech Section, Medical Neurology Branch, National Institute of Neurological Disorders & Stroke, National Institutes of Health, Bethesda, MD, USA

Abstract

Phonation is defined as a laryngeal motor behavior used for speech production, which involves a highly specialized coordination of laryngeal and respiratory neuromuscular control. During speech, brief periods of vocal fold vibration for vowels are interspersed by voiced and unvoiced consonants, glottal stops and glottal fricatives (/h/). It remains unknown whether laryngeal/respiratory coordination of phonation for speech relies on separate neural systems from respiratory control or whether a common system controls both behaviors. To identify the central control system for human phonation, we used event-related fMRI to contrast brain activity during phonation with activity during prolonged exhalation in healthy adults. Both whole-brain analyses and region of interest comparisons were conducted. Production of syllables containing glottal stops and vowels was accompanied by activity in left sensorimotor, bilateral temporoparietal and medial motor areas. Prolonged exhalation similarly involved activity in left sensorimotor and temporoparietal areas but not medial motor areas. Significant differences between phonation and exhalation were found primarily in the bilateral auditory cortices with whole brain analysis. The ROI analysis similarly indicated task differences in the auditory cortex with differences also detected in the inferolateral motor cortex and dentate nucleus of the cerebellum. A second experiment confirmed that activity in the auditory cortex only occurred during phonation for speech and did not depend upon sound production. Overall, a similar central neural system was identified for both speech phonation and voluntary exhalation that primarily differed in auditory monitoring.

Keywords

phonation; exhalation; speech; fMRI

Phonation for human speech is a highly specialized manner of laryngeal sensorimotor control that extends beyond generating the sound source of speech vowels and semi-vowels (e.g. /y/ and /r/). Laryngeal gestures for voice onset and offset distinguish voiced and unvoiced consonant pairs through precise timing of voice onset (Raphael et al., 2006). Brief interruptions in phonation by glottal stops (/ʔ/) mark word and syllable boundaries through hyper-adduction of the vocal folds to offset phonation. Each of these vocal fold gestures, in addition to rapid pitch changes for intonation require precisely controlled variations in intrinsic laryngeal activity (Poletto et al., 2004). To understand the neural control of phonation for speech, laryngeal control must be investigated independent of the oral articulators (lips, tongue, jaw). Several neuroimaging studies of phonation for speech included oral articulatory movements

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

with phonation, which can limit the examination of laryngeal control independently from oral articulatory function (Huang et al., 2001;Schulz et al., 2005). We expect that the neural control over repetitive laryngeal gestures for speech phonation will show a greater hemodynamic response (HDR) response in the left relative to right inferolateral sensorimotor areas (Wildgruber et al., 1996;Riecker et al., 2000;Jeffries et al., 2003). On the other hand, the neural control for vocalizations that are not specific to speech, such as whimper or prolonged vocalization, will show a more bilateral distribution (Perry et al., 1999;Ozdemir et al., 2006).

Phonation for speech also involves volitional control of respiration as subglottal pressure throughout exhalation must be controlled to initiate and maintain vocal fold vibration (Davis et al., 1996;Finnegan et al., 2000). To better understand the central neural control of laryngeal gestures for speech, the contribution of volitional control over respiration must also be examined. Previous neuroimaging findings of volitional respiratory control, however, have been inconsistent. In one positron emission tomography (PET) study (Ramsay et al., 1993), brain activity for volitional exhalation, involved similar cortical and subcortical regions as those reported for voiced speech (Riecker et al., 2000;Turkeltaub et al., 2002;Bohland et al., 2006). During exhalation, activity increased in the left inferolateral frontal gyrus (IFG) where the laryngeal motor cortex is thought to be located (Ramsay et al., 1993). On the other hand, functional MRI (fMRI) studies of volitional exhalation found increases in more dorsolateral sensorimotor regions, potentially corresponding to diaphragm and chest wall control, but not in the laryngeal motor regions (Evans et al., 1999;McKay et al., 2003). The neural correlates for voluntary exhalation need to be distinguished to understand the neural control of phonation for speech.

The two aims of the current study were to identify the central organization of phonation for speech and to contrast phonation for speech with volitional exhalation using event-related fMRI. We hypothesized that brain activity for glottal stops plus vowels produced as syllables involves a left predominant sensorimotor system as suggested by the results of Reicker et al (2000) and Jeffries et al (2003) for simple speech tasks. We also expected that this left predominant system for laryngeal syllable production would differ from upper airway control for prolonged exhalation, which may involve a more bilateral and more dorsolateral sensorimotor system.

Phonation and exhalation differ in auditory feedback because phonation produces an overt sound heard by the speaker while exhalation is quiet. To test whether auditory feedback affects the HDR during phonation for speech, we manipulated auditory feedback during phonation and whimper (a non-speech task with similar levels of auditory feedback). If no differences were found between phonation for speech and whimper then it would be the auditory feedback which could explain a greater auditory response during phonation than during exhalation. On the other hand, if there were greater responses in the auditory area during phonation when compared to whimper, it might be because the participants attend to a greater degree to their own voice during a phonatory task and than during the non-speech task of whimper. Once masking noise was applied we expected that phonation and whimper should not differ in their auditory responses because the participants could not hear their own productions. Therefore, we predicted that there would be differences in responses in auditory areas between the phonated and whimper productions in the normal feedback condition but not during the masked condition.

Materials and Methods

Experiment 1: Identifying and contrasting of the neural control of phonation for speech and voluntary exhalation using event-related fMRI

Subjects—Twelve adults between 23–69 years participated in the study (mean = 35 years, s.d. = 12 years, 7 females). All were right-handed, native English speakers. None had a history of neurological, respiratory, speech, hearing or voice disorders. Each had normal laryngeal structure and function on laryngeal nasoendoscopy by a board certified otolaryngologist. All provided written informed consent before participating in the study, which was approved by the Internal Review Board of the National Institute of Neurological Disorders and Stroke.

Phonatory and Respiration Tasks—The two phonation tasks involved: 1) repetitions of the syllable beginning with a glottal stop requiring vocal fold hyperadduction /ʔ/ and followed by a vowel requiring precise partial opening of the vocal folds to allow vibration /i/ (similar to the “ee” in “see”). This syllable was produced reiteratively (ʔiʔiʔi) at a rate similar to normal speech production (between 3–5 repetitions per second); and, 2) The onset and offset of vocal fold vibration for the same vowel /i/ which was prolonged for up to 4.5 seconds. The phonation tasks were chosen to minimize the linguistic and oral requirements of the task and focus on voice onset and offset using laryngeal gestures involved in speech, i.e., glottal stops and vowels.

The exhalation task involved voluntarily prolonged oral exhalation produced in quiet after a rapid inhalation. The subjects were trained to prolong the exhalation for the same duration as the phonatory tasks but to maintain quiet exhalation not producing a sigh or a whisper. The exhalation productions were monitored throughout the study to assure that participants did not produce a sigh or whisper. A quiet rest condition was used as the control condition.

On each phonation or exhalation trial, the subject first heard an auditory example of the target for 1.5 s presented at a comfortable loudness level through MRI compatible headphones (Avotec, Stuart, FL, USA). The exhalation example was amplified to the same intensity as the phonation example to serve as a clearly audible task cue, although subjects were trained to produce quiet exhalation. The acoustic presentation was followed at 3.0 seconds by the onset of a visual cue (green arrow) to produce the target. The subjects were trained to begin phonating or exhaling after the arrow onset and to stop immediately following offset of the arrow 4.5 s later or 7.5 s into the trial. A single whole brain volumetric scan was then acquired over the next 3 seconds (0.3 s silent period + 2.7 s TR) resulting in a total trial duration of 10.5 s (Figure 1). The subjects were not provided with any instructions for the quiet rest condition, except that on certain trials there would not be an acoustic cue, although the visual arrow was presented. A fixation cue in the form of a cross was presented throughout the experiment except when the arrow was shown. Phonation was recorded through a tubing-microphone system, placed on the subject’s chest approximately six inches from the mouth (Avotec, Stuart, FL, USA). An MRI compatible bellows system placed around the subject’s abdomen recorded changes in respiration.

Each experiment consisted of 45 repetitive syllable trials, 45 continuous phonation trials, 90 exhalation trials and 120 rest trials distributed evenly over 5 functional scanning runs giving 60 trials per run. Two initial scans at the beginning of each run, to reach homogenous magnetization, were not included in the analyses. The stimuli were randomized so each subject received a different task order. Prior to the fMRI scanning session, each subject was trained to produce continuous and repeated syllable phonations and prolonged exhalation for 4.5 seconds.

Functional Image Acquisition—To minimize head movements during syllable that could produce artifacts in the blood oxygenation level dependent (BOLD) signal, a vacuum pillow cradled the subject’s head. Additionally, an elastic head band placed over the forehead provided

tactile feedback and a slight resistance to shifts in head position. A sparse sampling approach was used; the scanner gradients were turned off during stimulus presentation and production so the task could be performed with minimal background noise (Birn et al., 1999; Hall et al., 1999). A single echo-planar imaging volume (EPI) of the whole brain was acquired starting 4.8 seconds after the onset of the production phase with the following parameters: TR- 2.7 s, TE-30 ms, Flip Angle- 90°, FOV- 240 mm, 23 sagittal slices, slice thickness - 6 mm (no gap). Because the subjects took between 0.5–1.0 seconds to initiate the task following the visual cue, the acquisition from 4.0 to 6.7 seconds occurred during the predicted peak of the HDR for speech based on previous studies (Birn et al., 1999; Huang et al., 2001; Birn et al., 2004; Langers et al., 2005). The HDR could possibly have also been affected by the auditory example stimulus because scanning began 5.8 s after auditory example offset (Birn et al., 2004; Langers et al., 2005), but this would have affected all phonatory and exhalation trials to the same degree as the auditory examples were all presented at the same sound intensity level. The possible effect of the auditory stimulus was addressed later in experiment 2.

A high-resolution T1-weighted MPRAGE scan for anatomical localization was collected after the experiment: TE-3.0 ms, TI- 450 ms, Flip Angle- 12°, BW- 31.25 mm, FOV- 240 mm, slices- 128, slice thickness 1.1 mm (no spacing).

Functional Image Analysis—All image processing and analyses were conducted with Analysis of Functional Neuroimages (AFNI) software (Cox, 1996). The initial image processing involved a motion correction algorithm for three translation and three rotation parameters that registered each functional volume from the five functional runs to the fourth volume of the first functional run. The motion estimate of the algorithm indicated most subjects only moved their head slightly during the experiment, typically ≈ 1 mm. If movement exceeded 1 mm, then the movement parameters were modeled as separate coefficients for head movement and included in the multiple regression analysis as noted below (only conducted for two subjects).

The HDR signal time-course of each voxel from each run (60 volumes) was normalized to percent change by dividing the HDR amplitude at each time point by the mean amplitude of all the trials from the same run and multiplying by 100. Each functional image was then spatially blurred with an 8 mm FWHM Gaussian filter followed by a concatenation of the five functional runs into a single time-course. The amplitude coefficients for the phonation and exhalation factors for each subject were estimated using multiple linear regression. The statistical maps and anatomic volume were transformed into the Talaraich & Tournoux coordinate system using a 12 point affine transformation.

The first group comparison was a conjunction analysis to determine whether the syllable repetition and extended vowel phonation conditions differed or whether the conditions could be combined for further analyses. A second related comparison was a voxel-wise paired t-test between the amplitude coefficients (% change) for the repetitive syllable and continuous phonation to assess task differences parametrically. If the tasks did not differ then they would be combined to compare with the exhalation task. To correct for multiple voxel-wise comparisons, a cluster-size thresholding approach was used to determine the spatial cluster size that would be expected by chance. The AlphaSim module within AFNI provided a probabilistic distribution of voxel-wise cluster sizes under the null hypothesis based on 1000 Monte Carlo simulations. The AlphaSim output indicated that clusters smaller than 6 contiguous voxels in the original coordinates (equivalent volume = 506 mm³) should be rejected at a corrected voxel-wise p value of $\alpha \leq 0.01$.

A voxel-wise, paired t-test was then used to compare whether the combined means of the repeated syllable and prolonged vowel trials differed significantly from zero. The same cluster

threshold of 6 contiguous voxels at the corrected alpha level ($p \leq 0.01$) described previously was used as the statistical threshold. The amplitude coefficient maps (% change) for the two combined phonatory tasks and the exhalation task were then compared directly using a voxel-wise paired t-test at the same threshold of $p < 0.01$ for 6 contiguous voxels. The resultant statistics from each test were normalized to Z scores. A conjunction analysis was then used to assess overlapping and distinct responses in the different task conditions.

Region of Interest (ROI) Analyses—An ROI analysis compared brain responses to the combined repetition and vowel prolongation tasks versus prolonged exhalation in anatomical regions previously related to voice and speech production in the literature (Riecker et al., 2000;Jurgens, 2002;Turkeltaub et al., 2002;Schulz et al., 2005;Bohland et al., 2006;Tremblay et al., 2006). The ROIs were obtained from the neuroanatomical atlas plug-in available with the Medical Imaging Processing, Analysis & Visualization (MIPAV) software (Computer and Information Technology, National Institutes of Health, Bethesda, MD). The ROIs in Talaraich and Tournaux space were developed by the Laboratory of Medical Image Computing (John Hopkins University, Baltimore, MD). MIPAV was used to register each pre-defined ROI to the normalized anatomical scan of the individual subjects. When certain ROI's encompassed skull/meningeal tissue or white matter in different subjects due to individual variations in brain anatomy, slight manual adjustments were applied to position the ROI within adjacent grey matter, but no changes in the shape or volume of the ROIs occurred. Individual ROI templates were converted to a binary mask dataset using MIPAV and segmented into right and left sides. To restrict the subsequent ROI analyses of percent volume and signal change to the anatomically defined ROI's, the mask datasets of the ROIs from each subject were applied to their own functional volumes from the phonation and exhalation results (without spatial blurring).

Certain ROI's were segmented or combined to form or distinguish regions. Only the lateral portion of the Brodmann Area (BA) 6 ROI was used to test for activation in the precentral gyrus, an area thought to encompass the laryngeal motor cortex based on sub-human primate studies (Simonyan et al., 2002;Simonyan et al., 2003). The primary motor and sensory ROI's were combined into a single ROI to create a primary sensorimotor region because functional activation is typically continuous across this region. However, the primary sensorimotor ROI was then divided at the level of $z=30$ to provide a ventral ROI thought to encompass the orofacial sensorimotor region and a dorsal ROI for the thoracic / abdomen region. The final set of 11 ROIs included the ventrolateral precentral gyrus (BA 6), inferior primary sensorimotor cortex (MI/SI-inf, BA 2-4), superior primary sensorimotor cortex (MI/SI-sup, BA 2-4), Broca's area (BA 44), anterior cingulate cortex (ACC, BA 24 & 32), insula (BA 13), superior temporal gyrus (STG, BA 22), dentate nucleus of the cerebellum, putamen, ventral lateral (VLn) and ventral posterior medial (VPm) nuclei of the thalamus. A suitable ROI was not available from the digital atlas for the supplementary motor area (SMA) because the only available ROI covered the mediodorsal BA 6, which was excessively large.

Within an ROI, the voxels that were active for a task (voiced syllables or prolonged exhalation) in each subject were identified using a statistical threshold ($p < 0.001$) to detect voxels with positive values (task > rest condition) that were less than 6 mm apart. The 6 mm criterion was chosen to extend beyond the in-plane resolution (3.75 mm * 3.75 mm) of the original coordinates while also encompassing the original slice thickness (6 mm). The volume of active voxels was converted to a percentage of the total ROI volume. A three way repeated ANOVA was computed to test for differences in the percent volume activation across three factors: Task (phonation versus exhalation); ROI (region of interest); and Laterality (right versus left hemisphere).

To compare the mean amplitude of activation in ROIs, we rank-ordered the voxels within each ROI starting with the most positive values and then selected the highest twenty percent of voxels in each ROI to calculate a cumulative mean amplitude. This approach ensured that the number of voxels for a particular ROI contributing to the mean amplitude was the same for each subject in each task. Selection of the 20% criterion was based on a pilot study conducting multiple simulations using cumulative mean values from 0–100% from different ROI's. Criteria below 15% only tended to capture the most intense activation, which likely did not represent the range of significant activity while criteria greater than 30% typically approached 0. The 20% criterion generated mean signal values that were around the asymptote between intense means of low criterions and the negligible mean signal values above 30%. A three way repeated ANOVA examined the factors: ROI, Task and Laterality. The overall significance level of $p \leq 0.025$ ($0.05/2$) that was corrected for the two ROI ANOVAs (Extent and Amplitude) was used to test for significant effects.

Results

The syllabic (glottal stop plus vowel) and prolonged vowel conditions elicited highly similar patterns of bilateral activity in cortical and sub-cortical regions that encompassed the orofacial sensorimotor cortex, insula, STG, SMA, anterior cingulate & cingulate gyrus, thalamus, putamen and cerebellum. A conjunction analysis was also conducted to identify overlapping and unique responses at a threshold of $p < 0.01$, which indicated a completely overlapping pattern of responses in the cerebral regions mentioned previously. A voxel-wise group comparison between the amplitude coefficients of the two conditions indicated the only task difference was a cluster of activation in the left midline SMA (Table 1) where syllabic repetition exceeded prolonged vowel production at the predetermined p value of $p < 0.01$ for a cluster of 6 voxels (Figure 2a). The amplitude of the HDR for the prolonged vowel never exceeded the amplitude of the response to the repeated syllable. The single cluster of differential activation, although theoretically important, was the only difference so the syllable repetition and extended vowel conditions were combined for all subsequent analyses and figures.

The combined phonatory task activation for the group using the criterion of $p < 0.01$ for 6 contiguous voxels showed a widespread bilateral pattern of brain activation. To discriminate separate spatial clusters for this descriptive analysis, we applied a more stringent criterion Z value of 4.0 for clusters $\geq 506 \text{ mm}^3$ in cortical and sub-cortical regions where the phonatory task activation was significantly greater than zero (Table 1). This stringent threshold indicated the most intense clusters of activation associated with the combined phonatory tasks. The most prominent cluster was in the left lateral cortex extending from the inferior frontal gyrus, through the post-central gyrus to the superior temporal gyrus encompassing Brodmann areas 1–4, 6, 22, 44 (Figure 2b, $z = 17$). Three activation clusters were detected in the right hemisphere, which included the right cerebellum (Figure 2b, $z = -17$), the right supramarginal gyrus (BA 40; Figure 2b, $z = 34$) and the right inferior lingual gyrus (Table 1). Bilateral activation was evident in the lateral pre & post central gyrus (BA 3, 4/6) in a region superior to the left ventrolateral cluster (Figure 2b, $z = 34$). Mid-line cortical activation was evident in a cluster that encompassed the SMA (BA 6, $z = 51$) and extended into the ACC (Table 1). Prominent subcortical activation was found in ventral and medial nuclei of the right thalamus and in the right caudate nucleus (Table 1). Activation in the PAG was not detected by the voxel-wise group analysis.

As with the phonatory tasks, voluntary exhalation elicited a widespread pattern of activation. Using the same criteria of a Z score of 4.0 (for 6 contiguous voxels), a set of cortical and sub-cortical regions similar to the phonatory tasks was identified (Figure 2b), although the volume of activation was decreased and fewer regions were detected (Table 1). Regions of activation in the left hemisphere included a large inferolateral cluster (BA 2-4, 6, 22, 44; Figure 2b, $z =$

17) and a more superior cluster (BA 3–4, 6; Figure 2b, $z = 34$), both encompassing primary sensorimotor regions. Two clusters of right hemisphere activation included the supramarginal (BA 40; Figure 2b, $z = 17$) and occipital gyri (BA 19). Subcortical activation involved large clusters of activation in the right cerebellum (Figure 2b, $z = -17$) and the right thalamus (Table 1).

The conjunction analysis of phonatory and exhalation tasks at the same statistical threshold ($Z \geq 4.0$) indicated a highly overlapping pattern of responses in left primary sensorimotor regions, left frontal operculum, bilateral insula, bilateral thalamus and right BA 40 (Figure 3a). The phonatory responses were more widespread throughout the left STG (BA 22/41/42), medial motor areas (SMA and ACC) and the right SMG, STG and lateral precentral cortex (BA 6). Almost no unique responses were detected for exhalation, which were encompassed by the phonatory task responses.

The HDR amplitude for the phonatory task exceeded that for exhalation at the corrected alpha level ($p < 0.01$ for 6 contiguous voxels) in only three regions (Figure 3b): the right and left superior temporal gyrus (STG) and in the right insula (Table 1). In the left STG, large clusters were found in BA 22 and more posteriorly in BA 40/42. A single large cluster was detected in the right STG that extended into the right insula. The HDR amplitude for exhalation was significant in numerous brain regions at the statistical threshold ($p < 0.01$ for 6 contiguous voxels), but did not exceed the amplitude of the HDR to phonation in any region.

ROI Comparison of Percent Region Volume

Increases in the percent of voxels in an ROI that were active ($p < 0.001$) occurred during both the phonatory and exhalation tasks in each subject for all of the 11 ROIs with few exceptions, suggesting that the brain areas captured by these ROIs were active for both tasks. The repeated measures ANOVA for the percent of voxels in an ROI which met the criterion of $p < 0.001$ indicated significant main effects for Task ($F(1,11) = 10.43$, $p = 0.008$) and Side ($F(1,11) = 6.94$, $p = 0.02$) and a three-way interaction between ROI, Task and Side ($F(1,11) = 2.928$, $p = 0.013$ – Huynh-Feldt corrected p value). We chose to evaluate the three-way interaction by investigating Task and Side effects and interactions within each ROI. Greater percent volumes occurred during the phonatory tasks compared to exhalation in 4 ROIs (Figure 4a–d): STG ($F(1,11) = 35.23$, $p < 0.001$), MI/SI-inferior ($F(1,11) = 35.23$, $p < 0.001$), dentate nucleus ($F(1,11) = 35.23$, $p < 0.001$) and VLn ($F(1,11) = 35.23$, $p < 0.001$). The percent volume was significantly greater on the left relative to the right for the MI/SI-inferior ROI ($F(1,11) = 7.94$, $p = 0.017$) (Figure 4e). Significant Task by Side interactions (Figure 5) were found for BA 44 ($F(1,11) = 7.50$, $p = 0.020$) and the insula ($F(1,11) = 10.71$, $p < 0.01$) where the response was greater on the right for phonatory tasks and equal during exhalation, while it was greater on the left than on the right for exhalation in the insula.

ROI Comparison of Response Amplitude

Repeated measures ANOVA for mean percent change indicated significant main effects for ROI ($F(10,110) = 6.29$, $p < 0.001$ – Huynh-Feldt corrected p value) and an ROI by Task interaction ($F(1,11) = 2.94$, $p = 0.015$ – Huynh-Feldt corrected p value). To focus on differences within ROIs, we used paired t -tests to compare the tasks – $p \leq 0.025$. The results indicated that percent change was significantly greater during the phonatory tasks compared to exhalation (Figure 6a, b) in BA 22 ($t(11) = 3.56$, $p < 0.01$) and the dentate nucleus ($t(11) = 2.68$, $p = 0.02$).

Experiment 2: A test of auditory feedback manipulations on the hemodynamic response for laryngeal tasks

Experiment 2 was aimed at determining whether the basis for the auditory response differences between the phonatory task and exhalation in experiment 1 was due to the presence of auditory

feedback during the phonatory task or because of auditory monitoring differences during speech and non-speech tasks (Belin et al., 2002). In experiment 2 we determined if auditory region response differences occurred: 1) between phonated and whispered syllable repetition tasks with auditory feedback differences; and 2) between phonated syllable repetition and a non-speech vocalization task (whimper), when both had the same auditory feedback of vocalization. We also conducted these comparisons when auditory feedback was masked by acoustic feedback during the tasks.

Subjects

Twelve adults between 22–69 years participated in the study (mean – 33 years, s.d. – 11 years, 6 females). Three of the subjects also participated in experiment 1. All were right-handed, native English speakers who met the same inclusion criteria as experiment 1. All provided written informed consent before participating in the study, which was approved by the Internal Review Board of the National Institute of Neurological Disorders and Stroke.

Experimental Tasks

The experiment involved 2 separate sessions in which the subjects produced speech-like phonatory and whispered syllable repetition and non-speech vocalization (whimper) tasks (separate trials) during one session with normal auditory feedback and during another session with auditory masking during production. The repeated syllable used for both the phonatory and whisper tasks was the same as experiment 1 (Gallena et al., 2001) produced reiteratively (?i?i?) at a rate similar to normal speech production (3–5 repetitions per second) – except the whisper task did not include vocal fold vibration. The non-speech vocalization task mimicked a short whimper sound and was repeated 3–5 times per second. The whispered / phonated syllable repetition and whimper each require active control over expiratory airflow to maintain production, with the primary difference being that the syllabic tasks involved speech like laryngeal gestures and the other a non-speech vocalization. The phonated syllable and whimper were produced at a similar speech sound intensity while whispered speech was much quieter. The acoustic examples for the phonatory and whispered syllable repetition and whimper tasks were amplified to the same intensity level to serve as clearly audible task cues. The subjects were further trained to prolong the tasks until the offset of the visual production cue. Each of these three tasks produced either vocal fold vibration or whispering involved the same oral posture so the oral articulatory and linguistic constraints were minimized; instead, the tasks focused on laryngeal movement control for either phonation, whisper or non-speech vocalization. Quiet rest was used as the control condition.

The time course of the stimulus presentation and response were the same as the first experiment (Figure 1). The auditory presentation format, visual arrow cue, response requirements and presentation/recording set-up were identical to experiment 1. For each subject, there were 36 syllable phonation trials, 36 whisper trials and 36 non-speech vocalization trials. The experiment involved other stimuli that are not reported here. The stimuli were randomized so each subject received a different task order.

In the normal feedback condition, the subjects' heard their own phonatory, vocal or whispered productions. The sound intensity level through air conduction was likely reduced because the subjects were wearing earplugs for hearing protection, but they could still hear their own productions via bone conduction except in the whispered condition. In the masking condition, the same reversed speech was presented binaurally at 80 dB SPL to mask both air and bone conduction of the subjects' own productions during all three conditions, phonated syllables, whispered syllables and whimper. We verified for each subject that the reversed speech noise was subjectively louder than their own production. The speech noise began at the same time as the visual production cue and continued for the entire production phase (4.5 sec). The

reversed speech provided effective masking of phonatory productions because the reversed speech masking noise and the utterance have similar long-term spectra (Katz et al., 2002).

The same event-related sparse-sampling design as described for experiment 1 was used for experiment 2, except that a software upgrade allowed us to shorten the TR from 2.7 seconds to 2.0 seconds and acquire 35 sagittal slices instead of 23. Voxel dimensions were 3.75 * 3.75 mm in-plane as per experiment 1 but the slice thickness was decreased from 6 mm to 4 mm, while the other EPI parameters remained unchanged. A whole-head anatomical MPRage volume was acquired for each subject (parameters identical to experiment 1). The functional image analysis used the same methods as experiment 1, except that 372 EPI volumes were collected over 6 runs.

The purpose of experiment 2 was to determine whether the auditory area HDR differences in experiment 1 were due to louder auditory feedback during the phonatory task that was not present during the quiet exhalation or whether the auditory findings could have resulted from differences in auditory processing of the speech-like aspects of the phonatory task versus exhalation (Belin et al., 2002). Therefore, in experiment 2, we contrasted the phonatory and whispered productions of the same syllables during normal feedback to determine if auditory area differences in the BOLD signal were due to auditory feedback differences. We then compared the phonatory syllable repetitions with non-speech whimper to determine if auditory area responses differed in speech-like and non-speech tasks with similar levels of auditory feedback. Finally we conducted both comparisons (phonated versus whispered syllables and phonated syllables versus whisper) during masking to identify if any auditory area differences occurred. Four voxel-wise planned comparisons (paired t-tests) were used to identify voxels where the group mean amplitude coefficients for the laryngeal tasks (versus rest) differed for - 1) phonatory syllables vs. whispered syllables (normal feedback); 2) phonatory syllables vs. whimper (normal feedback); 3) phonatory syllables vs. whispered syllables (masked feedback); 4) phonatory syllables vs. whimper (masked feedback). A cluster-size thresholding approach based on the same approach described for experiment 1 was used to correct for multiple voxel-wise comparisons (AlphaSim module within AFNI). A threshold of $p < 0.01$ was indicated as the nominal voxel-wise detection level for clusters of 9 or more contiguous voxels (equivalent volume – 506.25 mm³). Voxel clusters smaller than 9 contiguous voxels were rejected under the null hypothesis.

Results

In the normal feedback condition, no differences occurred for phonated syllable repetition versus whispered syllable repetition in the STG (BA 22/41) on either side (Figure 7a) while the repetitive syllable phonated task was associated with a significantly greater HDR in the left STG (BA 22/41) compared to whimper (Figure 7b). No other significant task differences were detected in the STG or temporal lobe between either the phonated syllable repetition versus whispered syllable repetition or the repetitive syllable phonated task versus whimper in the masking feedback conditions (Figure 7c-d).

Discussion

Volitional Laryngeal Control System

The central control of laryngeal gestures for producing glottal stops and vowels in syllables involved a widely distributed system encompassing the lateral sensory, motor and pre-motor regions in the left hemisphere; bilateral dorsolateral sensorimotor regions; right temporoparietal, cerebellar, and thalamic areas and the medial SMA and ACC. The left predominant responses and medial responses for phonatory tasks encompassed similar regions to those mapped by Penfield and Roberts (1959) where voice could be elicited or interrupted

with electrical stimulation. This functional system for phonation is similar to that found in recent neuroimaging studies of phonation (Haslinger et al., 2005; Ozdemir et al., 2006) and other studies involving simple speech production tasks (Murphy et al., 1997; Wise et al., 1999; Riecker et al., 2000; Huang et al., 2001).

A similar pattern of HDR activity was found for exhalation, which involved left ventrolateral and dorsolateral primary sensorimotor areas, and right sided responses in temporoparietal, cerebellar and thalamic regions. The group data suggest that relative to exhalation, phonation further involved responses in the SMA and ACC and more extensive left sensorimotor responses, but that the difference was quantitative rather than qualitative.

The major difference between the phonatory and exhalation response patterns was the greater response in the superior temporal gyrus in experiment 1. In experiment 2 we compared phonated and whispered syllable productions to determine if the finding in experiment 1 was due to a difference in sound due to voice feedback. Because no differences were found between phonated and whispered productions, it is unlikely the auditory area response differences in experiment 1 were due to either auditory feedback or differences in the auditory stimuli. On the other hand, greater auditory area responses occurred when comparing the phonated syllables with whimper, a non-speech task. These results suggest that the auditory area response differences in experiment 1 might have been found because participants were monitoring their own voice productions for phonatory tasks and not for exhalation because similar differences in auditory area responses were found between phonated syllables and whimper. Further, these differences were eliminated with auditory masking, suggesting that the ability to hear one's own voice during phonatory tasks is important for this auditory area response. These results are in agreement with behavioral and neuroimaging evidence of the importance of voice feedback during voice production. Rapid pitch shifts in voice feedback have been shown to produce rapid changes in voice production during extended vowels, and vowel pitch glides (Burnett et al., 1998; Burnett et al., 2002) and Belin and colleagues found greater auditory HDR responses to speech sounds than non-speech sounds (Belin et al., 2002).

The results point to a common volitional sensorimotor system for the production of laryngeal gestures for speech and voluntary breath control. For speech-like phonatory changes in voice onset and offset, precise timing of vocal fold opening and closing is required to distinguish segments (using glottal stops) and voiced versus voiceless consonants (Lofqvist et al., 1984). The opening and closing positions of the vocal folds to produce onsets and offsets are voluntarily controlled, even though vocal fold vibration is mechanically induced by air flow during exhalation (Titze, 1994). Voluntary exhalation and phonation both require volitional control over the chest wall and laryngeal valving to prolong the breath stream. The findings show that the brain regions involved in volitional control of exhalation also contribute to rapid changes in voice onset and offset for speech. The similar response patterns in the dorsolateral primary sensorimotor cortex controlling chest wall activity for thoracic, abdominal and diaphragm contractions were expected for voice and exhalation, but other similarities were not expected. The common left predominant pattern for phonation and exhalation is similar to a report of activity in the precentral gyrus during exhalation using PET (Ramsay et al., 1993), but these findings differ from fMRI studies of inhalation and exhalation (Evans et al., 1999; McKay et al., 2003), which did not find left predominant, inferolateral sensorimotor activation. The precentral gyrus region that was active for both tasks is close to the anatomical region previously identified from which electrical stimulation elicits vocal fold adduction in rhesus monkeys (Simonyan et al., 2002; Simonyan et al., 2003). Although previous studies have found that this region is not essential for innate phonations in subhuman primates (Kirzinger et al., 1982), this region may be part of for the volitional laryngeal control network for human speech.

The lack of difference between the phonatory tasks and exhalation in other than the auditory regions, however, may be due to the limited types of phonatory tasks sampled in this study. The phonatory tasks used included sustained vowels and repetitive syllable tasks, which do not probe the full range of laryngeal control for connected speech such as for pauses between phrases, and suprasegmental changes in pitch, loudness and rate for speech intonation. Further fMRI studies manipulating these vocal parameters during phonation may elicit more distinctive HDR activity patterns relative to voluntary exhalation.

Lateral Predominance of Sensorimotor Control for Laryngeal Gestures

The voxel-wise analyses indicated a left hemisphere predominance of laryngeal sensorimotor control for phonation and exhalation, which correspond to neuroimaging studies involving speech tasks with phonation where the sensorimotor activation was left-lateralized (Riecker et al., 2000;Jeffries et al., 2003;Brown et al., 2006). On the other hand, a more bilaterally symmetrical distribution of sensorimotor activity has been found in other studies using tasks involving phonation (Murphy et al., 1997;Wise et al., 1999;Haslinger et al., 2005;Ozdemir et al., 2006). Similar discrepancies have occurred in the study of singing, where some reports point to right-side predominance of sensorimotor control of vocalization (Riecker et al., 2000;Jeffries et al., 2003), while others report a symmetrical pattern (Perry et al., 1999;Brown et al., 2004;Ozdemir et al., 2006). The same has occurred for exhalation where bilaterally symmetrical sensorimotor activation was reported in orofacial and more dorsal regions for prolonged exhalation (Colebatch et al., 1991;Ramsay et al., 1993;Murphy et al., 1997), while two studies, along with our results suggest left sided predominance in the laryngeal motor cortex (Ramsay et al., 1993) and an inferolateral sensory area (Evans et al., 1999). The complex variations in the tasks used by different studies involve different motor/cognitive response patterns that could account for the differences in hemispheric responses.

The task by hemisphere interactions in BA 44 and the insula indicated that the response volume for phonation was greater in the right hemisphere compared to the left in these regions. The predominance of right side responses in these regions during syllabic phonation is not consistent with the view that speech activity in these regions is left predominant. One possible explanation is that the phonatory control aspects for voice onset and offset are independent of speech and language processing. Right sided predominance has been shown both for voice perception (Belin et al., 2002), as well as studies of widely varying singing tasks (Riecker et al., 2000;Brown et al., 2004;Brown et al., 2006;Ozdemir et al., 2006). This leaves the possibility open that subjects performed the tasks as a singing-like or tonal task involving right sided mechanisms for pitch processing. However, the right hemisphere IFG/insula activity could also be motor planning activity for non-speech vocalizations (Gunji et al., 2003). In any case, the interaction is difficult to explain without further study. The left sided predominance of the insula response during prolonged exhalation conforms to a previous report of left inferolateral cortical activation during volitional exhalation (Ramsay et al., 1993). A response in BA 44 has not been found consistently during simple repetitive speech tasks (Murphy et al., 1997;Wise et al., 1999), vowel phonation or humming (Ozdemir et al., 2006) or speech exhalation (Murphy et al., 1997). Studies that involved singing of a single pitch or overlearned songs also have not reported BA 44 responses predominantly on the left or right (Perry et al., 1999;Riecker et al., 2000;Jeffries et al., 2003;Ozdemir et al., 2006). One study of simple vowel phonation reported bilaterally symmetric activation in BA 44 for (Haslinger et al., 2005) while another study involving monotonic vocalization reported right BA 44 activation (Brown et al., 2004). As stated previously, further experimentation involving a broader range of phonatory tasks is necessary to resolve the contribution of inferofrontal areas.

Components of the Phonatory System

Despite the similarities in the lateral brain regions, the SMA and ACC responses in the volumetric analysis were only significant for phonation and not for exhalation. Others have found SMA responses during phonation (Colebatch et al., 1991), exhalation (Colebatch et al., 1991) and speech (Murphy et al., 1997; Wise et al., 1999; Riecker et al., 2005) along with the ACC (Barrett et al., 2004; Schaufelberger et al., 2005; Schulz et al., 2005). The greater involvement of the SMA during the repetitive syllable task in contrast with prolonged vowels suggests that SMA activity may be greater during complex movements such as syllable production for speech. The responses of the SMA reported here are consistent with an important role for the SMA in phonation for speech and conforms to theoretical positions that the SMA is involved in highly programmed movements (Goldberg, 1985).

The ACC responses found here and in other studies of speech phonation (Barrett et al., 2004; Schulz et al., 2005) suggest that the ACC may have a role in the human phonation system for speech. A significant difference between phonation and exhalation was not detected, however, which suggests that a BOLD response did occur in the ACC for exhalation (Colebatch et al., 1991; Corfield et al., 1995; Murphy et al., 1997), but it did not reach the statistical threshold. The lack of PAG responses contrasts with a PET study that reported ACC-PAG co-activation during voiced speech but not during whispered speech (Schulz et al., 2005). Schulz et al., (2005) sampled narrative descriptions of personal experiences, which likely involved more complex speech and emotional vocal expression that may have activated the ACC-PAG. The syllabic phonatory tasks used here did not have emotional content so the ACC-PAG system may not have been involved. However, to properly test a role for the PAG in phonation using fMRI, a cardiac/respiratory gating procedure would be required.

Phonation was associated with a significantly higher response in the dentate nucleus compared to exhalation. A task effect within the dentate has significance because the cerebellum is thought to have a particular role in error correction and coordination during speech production (Ackermann et al., 2000; Riecker et al., 2000; Wildgruber et al., 2001; Guenther et al., 2006). Similar preferential responses in the cerebellum during phonation relative to whispering were found during narrative speech (Schulz et al., 2005). Other areas of significance for the control of phonation suggested by our results include the left inferolateral sensorimotor cortex a region previously identified for voice and speech production (Penfield et al., 1959; Riecker et al., 2000) and which may subserve laryngeal motor control for speech in humans.

Conclusions

Generalizing the current findings may be limited by the nature of the tasks used. The phonation task involved vowels and glottal stops dependent upon precise control of vocal fold closing and opening at the larynx. We used this task to avoid the rapid orofacial movements of consonant production and the linguistic formulation of meaningful words. However, simple syllable production may not engage the same brain regions that are involved in articulated speech. In separate studies, however, the brain activation associated with either complex speech phonation (Schulz et al., 2005) or simple vowel production (Haslinger et al., 2005) involved similar bilateral regions as identified in the current study.

Voluntary exhalation may itself be seen as a 'component' in phonation because it drives vocal fold vibration. It remains possible that subjects controlled exhalation in a similar fashion to exhalation for phonation, which led to minimal task differences. A broad range of non-speech tasks involving volitional laryngeal control such as whimper studied here, may could involve similar bilateral sensory and motor regions (Dresel et al., 2005) to those found here for phonatory control. In future studies, speech and other laryngeal motor control tasks should be compared with syllabic phonation.

This study contrasted phonation for vowel and syllable production with exhalation in an effort to identify the human phonation system. A previous neuroimaging study compared phonation and exhalation and used the subtraction paradigm may have missed the central control aspects of voice for speech by assuming that the components of different tasks are additive (Murphy et al., 1997). Instead, the two tasks appear to share highly related central control that would not be observed when using subtraction (Sidtis et al., 1999). By contrasting the percent change coefficient for phonatory tasks with voluntary exhalation we identified the regions of the human brain system involved in voice control for syllabic phonation. Furthermore, the HDR was estimated from single events, which has greater specificity than the longer duration events used in the PET studies (Murphy et al., 1997; Schulz et al., 2005). These methodological differences may explain why we found left hemisphere laryngeal motor control predominant for phonation that was not found in previous studies. In conclusion, our results suggest that the laryngeal gestures for vowel and syllable production and controlled exhalation involve left hemisphere mechanisms similar to speech articulation.

Acknowledgements

The study was supported by the Division of Intramural Research of the National Institute of Neurological Disorders and Stroke, National Institutes of Health. The authors wish to thank Pamela Kearney, M.D. and Keith G. Saxon, M.D. for conducting the nasolaryngoscopy examinations. We also thank Chen Gang, Ph.D. and Ziad Saad, Ph.D. for assistance with image processing and statistical analyses.

References

- Ackermann H, Hertrich I. The contribution of the cerebellum to speech processing. *J. Neuroling* 2000;13 (2-3):95-116.
- Barrett J, Pike GB, Paus T. The role of the anterior cingulate cortex in pitch variation during sad affect. *Eur. J. Neurosci* 2004;19 (2):458-464.
- Belin P, Zatorre RJ, Ahad P. Human temporal-lobe response to vocal sounds. *Cog. Brain Res* 2002;13 (1):17-26.
- Birn RM, Bandettini PA, Cox RW, Shaker R. Event-related fMRI of tasks involving brief motion. *Hum. Brain Mapp* 1999;7 (2):106-114.
- Birn RM, Cox RW, Bandettini PA. Experimental designs and processing strategies for fMRI studies involving overt verbal responses. *Neuroimage* 2004;23 (3):1046-58. [PubMed: 15528105]
- Bohland JW, Guenther FH. An fMRI investigation of syllable sequence production. *Neuroimage* 2006;32 (2):821-41. [PubMed: 16730195]
- Brown S, Martinez MJ, Hodges DA, Fox PT, Parsons LM. The song system of the human brain. *Cog. Brain Res* 2004;20 (3):363-375.
- Brown S, Martinez MJ, Parsons LM. Music and language side by side in the brain: a PET study of the generation of melodies and sentences. *Eur. J. Neurosci* 2006;23 (10):2791-2803.
- Burnett TA, Freedland MB, Larson CR, Hain TC. Voice F0 responses to manipulations in pitch feedback. *Journal Of The Acoustical Society Of America* 1998;103 (6):3153-3161. [PubMed: 9637026]
- Burnett TA, Larson CR. Early pitch-shift response is active in both steady and dynamic voice pitch control. *Journal Of The Acoustical Society Of America* 2002;112 (3):1058-1063. [PubMed: 12243154]
- Colebatch JG, Adams L, Murphy K, Martin AJ, Lammertsma AA, Tochondanguy HJ, Clark JC, Friston KJ, Guz A. Regional Cerebral Blood-Flow During Volitional Breathing In Man. *J Physiol-London* 1991;443:91-103. [PubMed: 1822545]
- Corfield DR, Fink GR, Ramsay SC, Murphy K, Harty HR, Watson JDA, Adams L, Frackowiak RSJ, Guz A. Evidence For Limbic System Activation During Co2-Stimulated Breathing In Man. *J Physiol-London* 1995;488 (1):77-84. [PubMed: 8568667]
- Cox RW. AFNI: Software for analysis and visualization of functional magnetic resonance neuroimages. *Computers And Biomedical Research* 1996;29 (3):162-173. [PubMed: 8812068]
- Davis PJ, Zhang SP, Winkworth A, Bandler R. Neural control of vocalization: respiratory and emotional influences. *J Voice* 1996;10 (1):23-38. [PubMed: 8653176]

- Dresel C, Castrop F, Haslinger B, Wohlschlaeger AM, Hennenlotter A, Ceballos-Baumann AO. The functional neuroanatomy of coordinated orofacial movements: Sparse sampling fMRI of whistling. *Neuroimage* 2005;28 (3):588–597. [PubMed: 16084116]
- Evans KC, Shea SA, Saykin AJ. Functional MRI localisation of central nervous system regions associated with volitional inspiration in humans. *J Physiol-London* 1999;520 (2):383–392. [PubMed: 10523407]
- Finnegan EM, Luschei ES, Hoffman HT. Modulations in respiratory and laryngeal activity associated with changes in vocal intensity during speech. *Journal of Speech Language and Hearing Research* 2000;43 (4):934–50.
- Gallena S, Smith PJ, Zeffiro T, Ludlow CL. Effects of levodopa on laryngeal muscle activity for voice onset and offset in Parkinson disease. *Journal of Speech Language and Hearing Research* 2001;44 (6):1284–99.
- Goldberg G. Supplementary motor area structure and function: Review and hypotheses. *Behavioral And Brain Sciences* 1985;8:567–616.
- Guenther FH, Ghosh SS, Tourville JA. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain and language* 2006;96 (3):280–301. [PubMed: 16040108]
- Gunji A, Kakigi R, Hoshiyama M. Cortical activities relating to modulation of sound frequency: how to vocalize? *Cog. Brain Res* 2003;17 (2):495–506.
- Hall DA, Haggard MP, Akeroyd MA, Palmer AR, Summerfield AQ, Elliott MR, Gurney EM, Bowtell RW. "Sparse" temporal sampling in auditory fMRI. *Hum. Brain Mapp* 1999;7 (3):213–223.
- Haslinger B, Erhard P, Dresel C, Castrop F, Roettinger M, Ceballos-Baumann AO. "Silent event-related" fMRI reveals reduced sensorimotor activation in laryngeal dystonia. *Neurology* 2005;65 (10):1562–9. [PubMed: 16301482]
- Huang J, Carr TH, Cao Y. Comparing cortical activations for silent and overt speech using event-related fMRI. *Hum. Brain Mapp* 2001;15 (1):39–53.
- Jeffries KJ, Fritz JB, Braun AR. Words in melody: an H(2)15O PET study of brain activation during singing and speaking. *Neuroreport* 2003;14 (5):749–54. [PubMed: 12692476]
- Jurgens U. Neural pathways underlying vocal control. *Neuroscience and Biobehavioral Reviews* 2002;26 (2):235–58. [PubMed: 11856561]
- Katz, J.; Lezynski, J. Clinical Masking. In: Katz, j, editor. *Handbook of Clinical Audiology*. Philadelphia: Lippincott Williams & Wilkins; 2002. p. 124-141.
- Kirzinger A, Jurgens U. Cortical lesion effects and vocalization in the squirrel monkey. *Brain Res* 1982;233 (2):299–315. [PubMed: 7059813]
- Langers DRM, Van Dijk P, Backes WH. Interactions between hemodynamic responses to scanner acoustic noise and auditory stimuli in functional magnetic resonance imaging. *Magn. Reson. Med* 2005;53 (1):49–60.
- Lofqvist A, McGarr NS, Honda K. Laryngeal Muscles And Articulatory Control. *Journal Of The Acoustical Society Of America* 1984;76 (3):951–954. [PubMed: 6491054]
- McKay LC, Evans KC, Frackowiak RS, Corfield DR. Neural correlates of voluntary breathing in humans. *J. App. Physiol* 2003;95 (3):1170–8.
- Murphy, K.; Corfield, DR.; Guz, A.; Fink, GR.; Wise, RJ.; Harrison, J.; Adams, L. *Journal of applied physiology*. 83. Bethesda, Md: 1997. 1985. Cerebral areas associated with motor control of speech in humans; p. 1438-47.
- Ozdemir E, Norton A, Schlaug G. Shared and distinct neural correlates of singing and speaking. *Neuroimage* 2006;33 (2):628–635. [PubMed: 16956772]
- Penfield, W.; Roberts, L. *Speech and Brain Mechanisms*. Princeton: Princeton University Press; 1959.
- Perry DW, Zatorre RJ, Petrides M, Alivisatos B, Meyer E, Evans AC. Localization of cerebral activity during simple singing. *Neuroreport* 1999;10 (18):3979–3984. [PubMed: 10716244]
- Poletto CJ, Verdun LP, Strominger R, Ludlow CL. Correspondence between laryngeal vocal fold movement and muscle activity during speech and nonspeech gestures. *J. App. Physiol* 2004;97 (3): 858–66.
- Ramsay SC, Adams L, Murphy K, Corfield DR, Grootoonek S, Bailey DL, Frackowiak RSJ, Guz A. Regional Cerebral Blood-Flow During Volitional Expiration In Man - A Comparison With Volitional Inspiration. *J Physiol-London* 1993;461:85–101. [PubMed: 8350282]

- Raphael, L.J.; Borden, G.J.; Harris, K.S. *Speech science primer: physiology, acoustics, and perception of speech*. Baltimore: Lippincott Williams & Wilkins; 2006.
- Riecker A, Ackermann H, Wildgruber D, Dogil G, Grodd W. Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *Neuroreport* 2000;11 (9):1997–2000. [PubMed: 10884059]
- Riecker A, Mathiak K, Wildgruber D, Erb M, Hertrich I, Grodd W, Ackermann H. fMRI reveals two distinct cerebral networks subserving speech motor control. *Neurology* 2005;64 (4):700–706. [PubMed: 15728295]
- Schaufelberger M, Senhorini MCT, Barreiros MA, Amaro E, Menezes PR, Scazufca M, Castro CC, Ayres AM, Murray RM, McGuire PK, Busatto GF. Frontal and anterior cingulate activation during overt verbal fluency in patients with first episode psychosis. *Revista Brasileira De Psiquiatria* 2005;27 (3): 228–232. [PubMed: 16224612]
- Schulz GM, Varga M, Jeffries K, Ludlow CL, Braun AR. Functional neuroanatomy of human vocalization: An (H2OPET)-O-15 study. *Cereb Cortex* 2005;15 (12):1835–1847. [PubMed: 15746003]
- Sidtis JJ, Strother SC, Anderson JR, Rottenberg DA. Are brain functions really additive? *Neuroimage* 1999;9 (5):490–496. [PubMed: 10329288]
- Simonyan K, Jurgens U. Cortico-cortical projections of the motorcortical larynx area in the rhesus monkey. *Brain Res* 2002;949 (1–2):23–31. [PubMed: 12213296]
- Simonyan K, Jurgens U. Efferent subcortical projections of the laryngeal motorcortex in the rhesus monkey. *Brain Res* 2003;974 (1–2):43–59. [PubMed: 12742623]
- Titze, IR. *Principles of Voice Production*. Englewood Cliffs, N.J.: Prentice Hall; 1994.
- Tremblay P, Gracco VL. Contribution of the frontal lobe to externally and internally specified verbal responses: fMRI evidence. *Neuroimage* 2006;33 (3):947–957. [PubMed: 16990015]
- Turkeltaub PE, Eden GF, Jones KM, Zeffiro TA. Meta-analysis of the functional neuroanatomy of single-word reading: method and validation. *Neuroimage* 2002;16 (3 Pt 1):765–80. [PubMed: 12169260]
- Wildgruber D, Ackermann H, Grodd W. Differential contributions of motor cortex, basal ganglia, and cerebellum to speech motor control: Effects of syllable repetition rate evaluated by fMRI. *Neuroimage* 2001;13 (1):101–109. [PubMed: 11133313]
- Wildgruber D, Ackermann H, Klose U, Kardatzki B, Grodd W. Functional lateralization of speech production at primary motor cortex: A fMRI study. *Neuroreport* 1996;7 (15–17):2791–2795. [PubMed: 8981469]
- Wise RJS, Greene J, Buchel C, Scott SK. Brain regions involved in articulation. *Lancet* 1999;353 (9158): 1057–1061. [PubMed: 10199354]

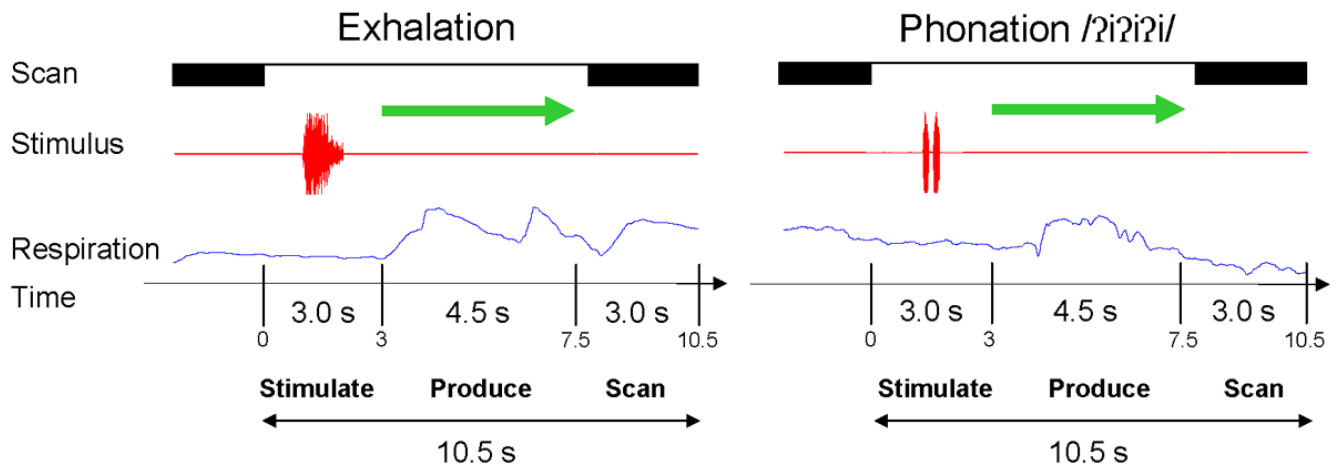


Figure 1.

A schematic representation of the event-related, clustered-volume design. The subject heard an auditory model of exhalation task or phonation followed by a visual cue to repeat the model. A single whole brain echo-planar volume was acquired following the removal of the visual cue.

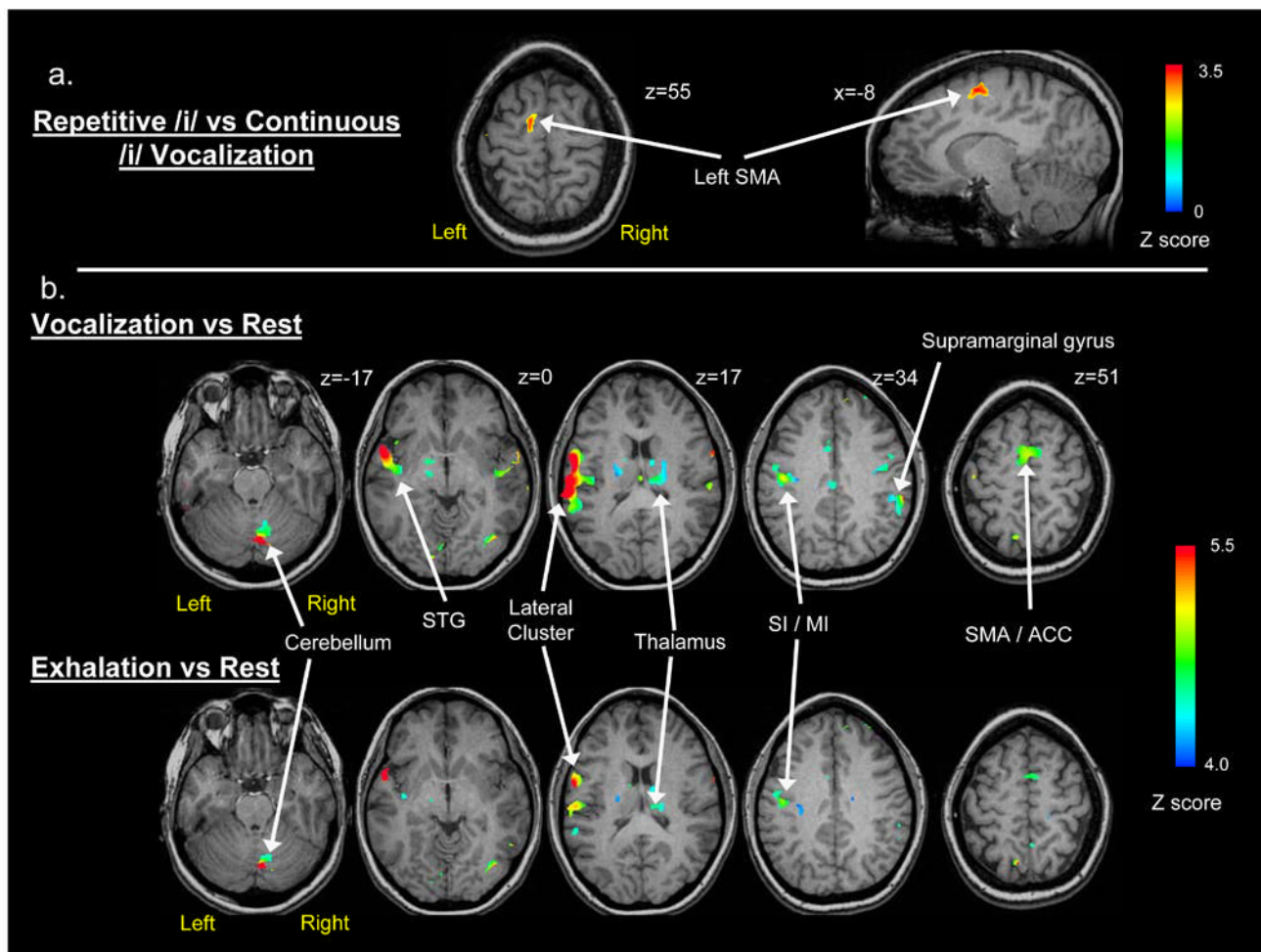


Figure 2. Brain imaging results from voxel-wise comparisons of the phonation and exhalation tasks based on average group data. The arrows indicate clusters of significant positive activation: a) The direct contrast between repetitive and continuous phonation ($p < .01$ for a cluster size of 506 mm^3); b) Axial brain images showing the statistical comparison ($Z \geq 4.0$ for a cluster size of 506 mm^3) between phonation and quiet respiration (top row) and the comparison between exhalation and quiet respiration (bottom row).

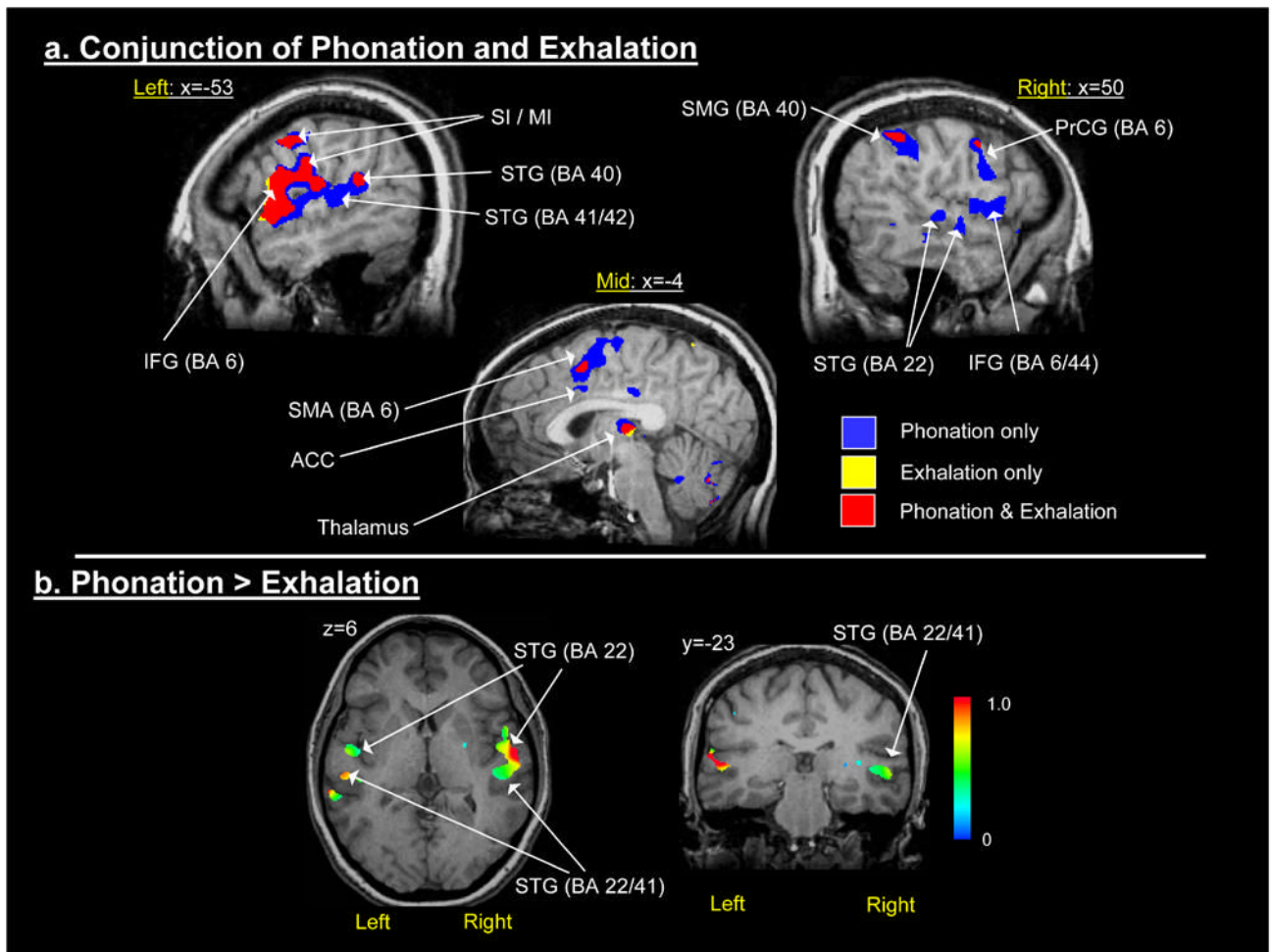


Figure 3. a) Conjunction map of the combined phonation condition and the exhalation condition ($Z \geq 4.0$); and, b) Paired t-test comparison of the amplitude coefficients from the combined phonation condition versus the exhalation condition.

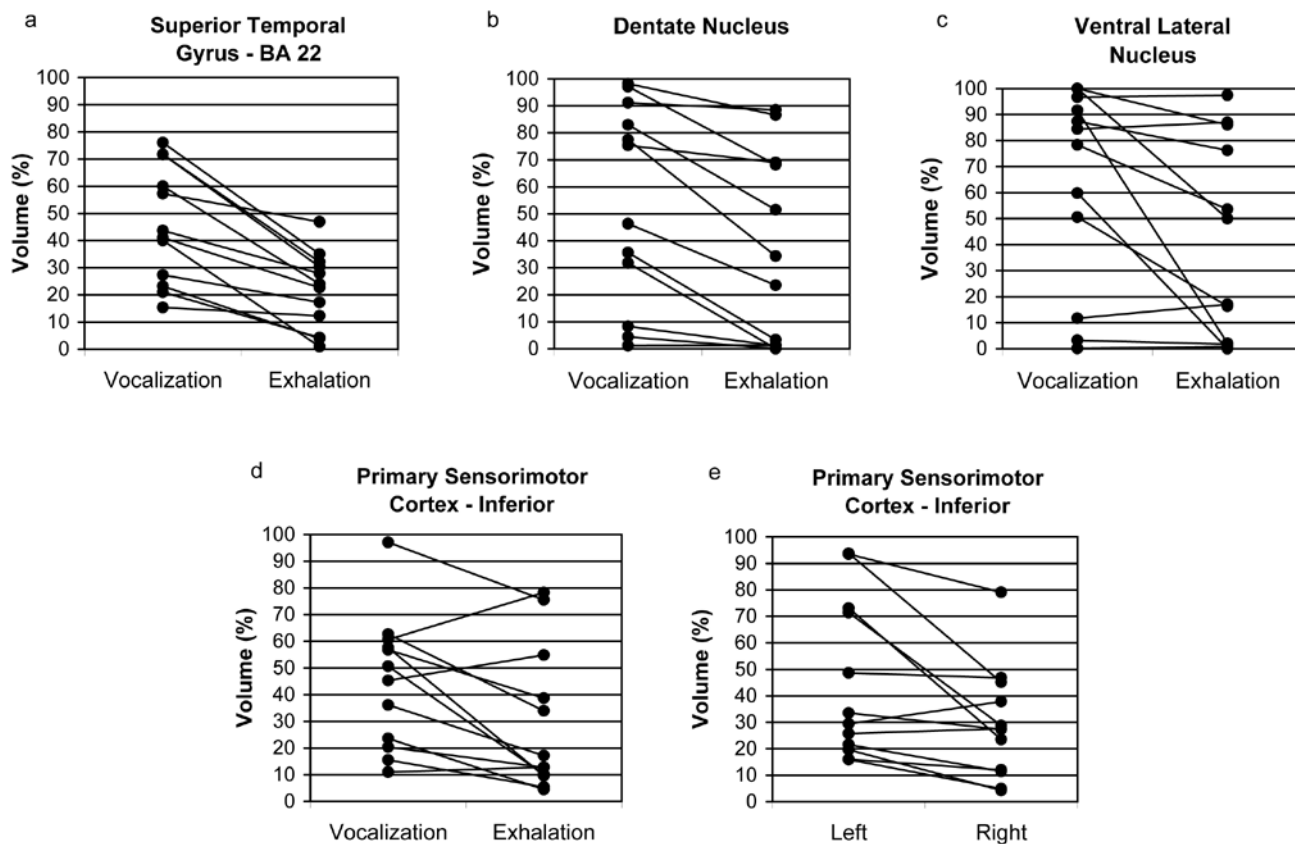


Figure 4. Functional activation extent (% activation volume) from each subject is compared for specific ROI's. Task differences are shown in figures 4a-d and a hemisphere difference is shown in figure 4e: a) superior temporal gyrus (BA 22); b) dentate nucleus of the cerebellum; c) ventral lateral nucleus of the thalamus (VLn); d-e) inferolateral primary sensorimotor cortex (MI / SI - Inferior).

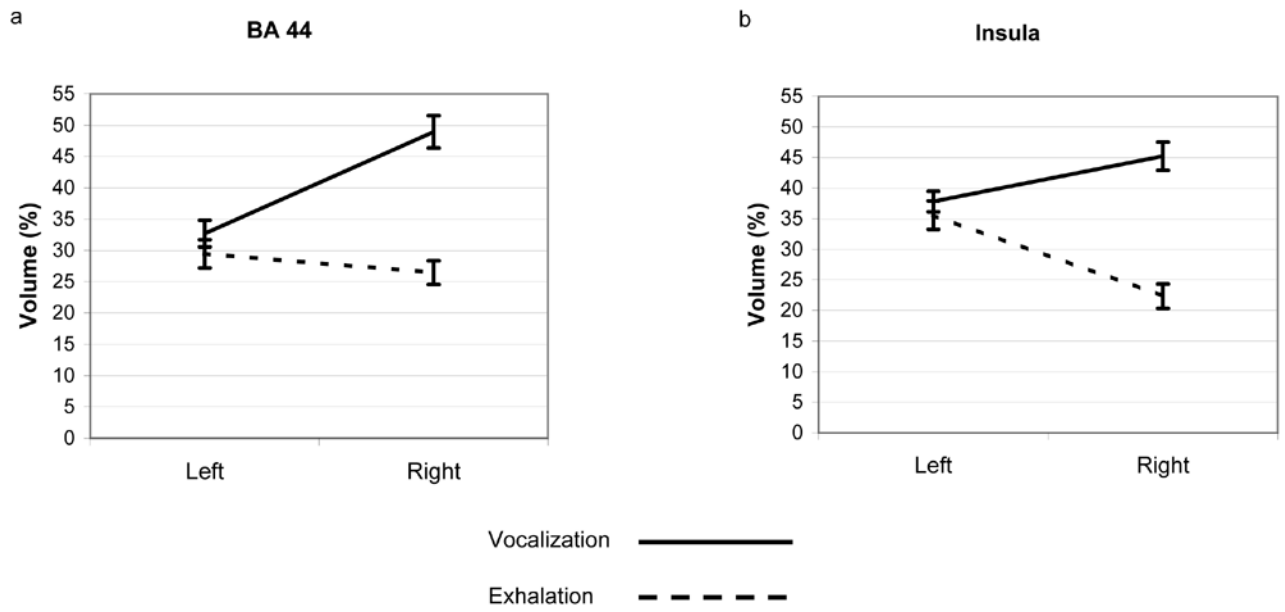


Figure 5. Average functional activation extent is shown for the task conditions (Phonation & Exhalation) and hemisphere (Left & Right): a) BA 44 and b) insula.

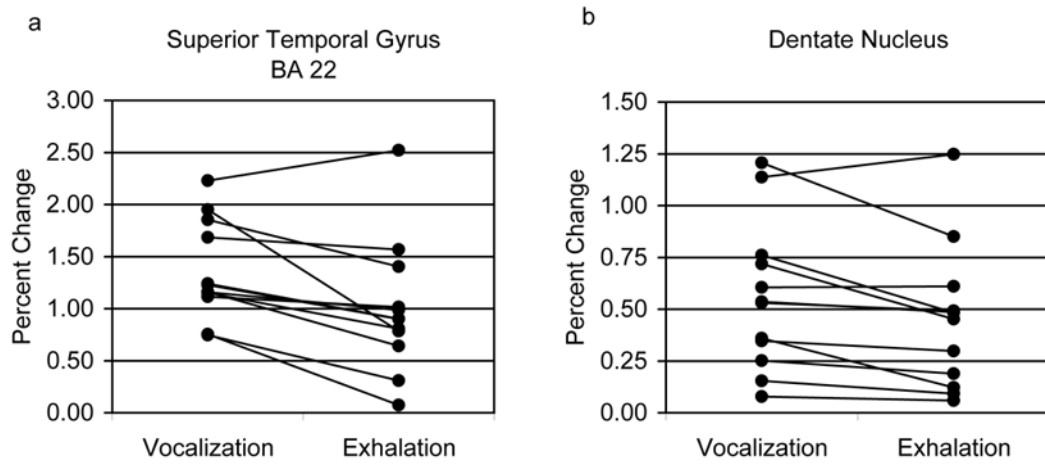


Figure 6. Task differences in mean activation amplitude (% change) from each subject is compared within specific ROI's: a) superior temporal gyrus (BA 22); b) dentate nucleus of the cerebellum.

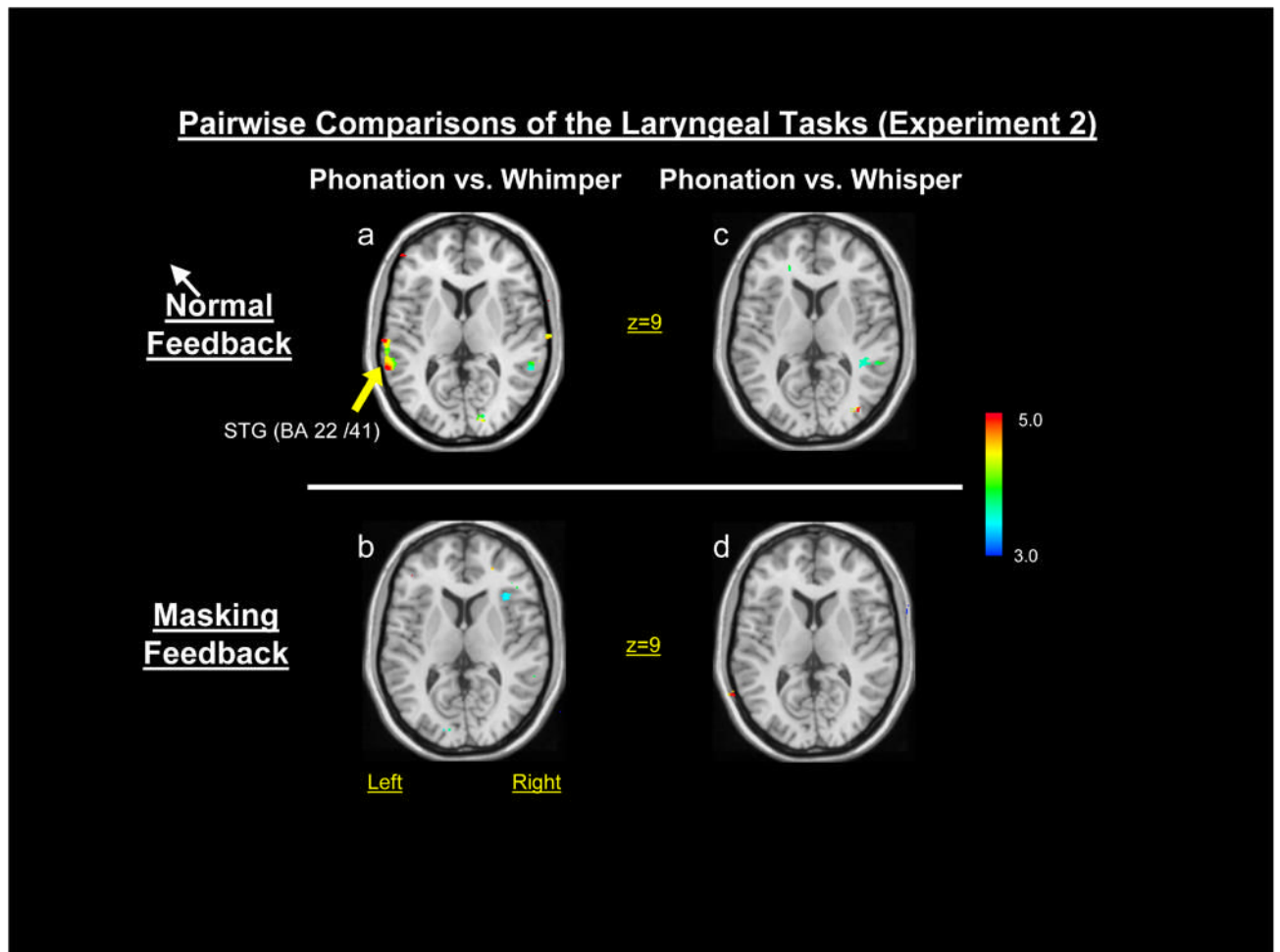


Figure 7. Pairwise comparisons (paired t-tests) of the phonatory, whimper and whisper tasks of experiment 2: a) phonation vs. whimper in the normal feedback condition; b) phonation vs. whimper in the masking feedback condition; c) phonation vs. whisper in the normal feedback condition; and, d) phonation vs. whisper in the masking feedback condition.

Table 1

Brain activation during vocalization and exhalation

Regions showing significant activation are listed for each condition and for the relevant contrasts between the conditions. For each region, the volume of activation at the p value associated with the contrast is presented, along with the mean and maximum Z score. The site of the maximum Z score is given in Talaraich & Tournoux coordinates. Only activation clusters equal to or exceeding 506 mm³ are presented.

Brain Region	Brodmann No.	Volume mm ³	Mean Z-score	Max Z-score	x	y	z
Repetitive Vocalization > Continuous Vocalization							
Lt SMA	6	531	3.3	4.0	-10	-2	56
Vocalization > Rest							
Lt ventrolateral cortex	1-4, 6, 22	6596	5.0	6.6	-59	-28	18
Lt pre/post central gyri	3-4, 6	1873	4.9	5.8	-44	-10	39
Rt precentral gyrus	4, 6	664	4.9	5.9	36	-8	28
Rt supramarginal gyrus	40	1001	5.1	6.2	56	-41	42
Midline SMA & ACC	6, 32	886	4.8	5.5	-5	-8	47
Rt inferior lingual gyrus	19	518	5.3	7.0	39	-72	-10
Rt cerebellum		3902	4.9	6.2	7	-71	-13
Rt thalamus VA, VL, VPm, MD		1183	4.9	5.5	9	-4	18
Exhalation > Rest							
Lt ventrolateral cortex	1-4, 6, 22, 40	4390	4.7	6.2	-42	-4	10
Lt pre/post central gyri	3-4, 6	1418	4.6	5.6	-44	-12	38
Rt supramarginal gyrus	40	653	4.8	6.0	56	-41	41
Rt inferior lingual gyrus	19	591	5.0	6.8	39	-70	-6
Rt cerebellum		3034	4.6	6.2	7	-70	-13
Rt thalamus - VPLn, LPN		1933	4.5	5.8	21	22	11
Vocalization > Exhalation							
Rt STG & insula	21, 22	5331	3.4	5.2	64	7	-2
Lt STG	40, 42	2020	3.2	4.1	-59	-28	18
Lt STG	22	759	3.3	4.2	-48	-9	4

Abbreviations: MFG - medial frontal gyrus, STG - superior temporal gyrus, SMA - supplementary motor area, ACC - anterior cingulate cortex; Thalamus: VA - ventral anterior, VL - ventral lateral, VPm - ventral posteromedial, LPn - lateral posterior, MD - medial dorsal. Lt - Left, Rt - Right