

Event-related functional MRI investigation of vocal pitch variation[☆]

Kyung K. Peck^{a,b,*}, Jessica F. Galgano^{c,d}, Ryan C. Branski^d, Dmitry Bogomolny^d, Margaret Ho^d, Andrei I. Holodny^a, Dennis H. Kraus^d

^a Department of Radiology, Memorial Sloan-Kettering Cancer Center, New York, NY, USA

^b Department of Medical Physics, Memorial Sloan-Kettering Cancer Center, New York, NY, USA

^c Department of Speech-Language Pathology, Teacher's College, Columbia University, New York, NY, USA

^d Department of Head and Neck Surgery, Memorial Sloan-Kettering Cancer Center, New York, NY, USA

ARTICLE INFO

Article history:

Received 27 May 2008

Revised 15 August 2008

Accepted 22 August 2008

Available online 10 September 2008

ABSTRACT

Voice production involves precise, coordinated movements of the intrinsic and extrinsic laryngeal musculature. A component of normal voice production is the modification of pitch. The underlying neural networks associated with these complex processes remains poorly characterized. However, several investigators are currently utilizing neuroimaging techniques to more clearly delineate these networks associated with phonation. The current study sought to identify the central cortical mechanism(s) associated with pitch variation during voice production using event-related functional MRI (fMRI). A single-trial design was employed consisting of three voice production tasks (low, comfortable, and high pitch) to contrast brain activity during the generation of varying frequencies. For whole brain analysis, volumes of activation within regions activated during each task were measured. Bilateral activations were shown in the cerebellum, superior temporal gyrus, insula, precentral gyrus, postcentral gyrus, inferior parietal lobe, and post-cingulate gyrus. In the left hemisphere, activations in the medial and middle frontal gyri were also observed. Regions active during high pitch production when compared to comfortable pitch were evident in the bilateral cerebellum, left inferior frontal gyrus, left cingulate gyrus, and left posterior cingulate. During low pitch generation, activations were present in the inferior frontal gyrus, insula, putamen, and cingulate gyrus in the left hemisphere. The inferior frontal gyrus in the right hemisphere produced greater activity than the area of the left hemisphere during high and low pitch generation. These results suggest that a single-trial design is sensitive enough to begin to delineate a widespread network of activations in both hemispheres associated with vocal pitch variation.

© 2008 Elsevier Inc. All rights reserved.

Introduction

Rapid pitch adjustments are associated with the complexities of voice production during both speaking and singing. Moreover, intricate coordination of the respiratory and articulatory subsystems is required to produce the inherent pitch variability associated with phonation. With regard to the larynx, pitch modulation is largely regulated by activation of the cricothyroid muscle, increasing both the length and tension of the vocal folds. This process is precisely monitored and regulated via the intrinsic laryngeal musculature (Poletto et al., 2004).

Utilizing neuroimaging techniques to more clearly delineate the neural networks underlying laryngeal control and phonation (i.e.,

voicing) has become increasingly common (Galgano et al., in press; Galgano and Froud, 2008; Loucks et al., 2007; Schulz et al., 2005). Functional MRI (fMRI) was employed to identify a larynx-specific area in the motor cortex during vocal and non-vocal laryngeal tasks (Brown et al., 2008). In addition, fMRI has shown opposite hemispheric lateralization during singing and word production (Riecker et al., 2000). Advances in technology, paradigm designs, and analysis techniques, such as those mentioned, have reduced the prevalence and negative impact of task-associated motion artifacts (Gopinath et al., in press). These advances permit more comprehensive examination of the neural correlates of tasks that involve motion within and outside of the field of view (e.g. voice, articulatory movements, and overt speech production tasks).

The neural basis of human pitch *perception* has been studied previously using functional MRI (Lattner et al., 2005) and magnetoencephalography (MEG) (Patel and Balaban, 2001). These data suggest that brain activity within the right auditory cortex is greater than a homologous region in the left auditory cortex. However, the neural correlates of laryngeal movement during vocal pitch *production* at comfortable as well as other pitch levels remain poorly

[☆] Portion of this work was presented at the International Society of Magnetic Resonance in Medicine (ISMRM), Toronto, Canada, May 2008.

* Corresponding author. Department of Medical Physics and Radiology, Memorial Sloan-Kettering Cancer Center 1275 York Avenue, New York, NY 10021, USA. Fax: +1 212 717 3010.

E-mail address: peckk@mskcc.org (K.K. Peck).

characterized. This is due, in part, to difficulties controlling for experimental confounds between the sensorimotor and cognitive-linguistic aspects of speech and phonation. In addition, complex and intricate laryngeal adjustments are required to facilitate regulation of subglottal respiratory pressures and changes in the length and tension of the vocal folds when engaging in production of various frequencies/pitches (Schulz et al., 2005). Thus, investigations aiming to isolate the neural mechanisms of voice production, and in particular, pitch modulation, are particularly challenging.

Phonation requires precise, coordinated laryngeal movements and rapid changes in intrinsic and extrinsic laryngeal musculature leading to perceptual variations in pitch. Utilization of a wide phonational frequency range contributes to the richness of human voice expression. However, people with neurologically-based laryngeal pathologies have limited pitch ranges and frequently present with other pitch generation abnormalities, such as pitch breaks (Brin et al., 1998). Therefore, successful determination of an imaging method and paradigm design which permits discovery of BOLD signal changes associated with the control of pitch variability is warranted. These findings will ultimately enhance our understanding of the pathogenesis of diseases such as laryngeal dystonia and further the development of pre-neurosurgical protocols (Grillone and Chan, 2006).

The primary aim of the current study was to examine the efficacy of a single-trial fMRI design in identifying differences in the neural substrates of human pitch production. Recent neuroimaging studies have demonstrated a widespread network of regional activations and functional connectivity of the subsystems involved in phonation (Schulz et al., 2005; Loucks et al., 2007). However, studies specifically investigating human vocal pitch control are rare. We therefore sought to determine whether a single-trial design is satisfactory in decreasing sensitivity to the movements induced by vocal pitch production tasks, but also sensitive enough to induce BOLD signal changes when generating high, low, and modal (i.e., comfortable) pitches. A network of regions is likely differentially coupled with pitch modulation due to significant differences in cranial nerve and laryngeal musculature activity associated with the complex process of manipulating pitch (Hixon, et al., 2008). We therefore hypothesized that parallel differences in the neural areas related to the activation and monitoring of those phonatory differences could be observed using fMRI. The single-trial design does not directly address the problem of identifying regional activations that contribute to rapid pitch changes, such as those used during speaking. However, the findings may provide new information about the similarities and differences in regional activations and BOLD signal changes associated with vocal pitch production and control for three different pitches. The findings may also provide a foundation for future investigations of pitch control and variation with additional speaking or singing demands.

Subjects and methods

Subjects

Twelve healthy adults (six males and six females), ranging in age from 20 to 30 years (mean=25.4 years; SD=3.6 years) without previous history of neurological illness or voice disorder were recruited for this prospective study. All subjects were right-handed on the Edinburgh Handedness Inventory (Oldfield, 1971) and native English speakers. All subjects provided written informed consent, approved by the Institutional Review Board.

Vocal function evaluation

Evaluation of vocal function served as a secondary endpoint for comparison with the neurological correlates of voice production. A certified speech-language pathologist conducted all voice assessments, consisting of acoustic analyses. Subjects were seated comfortably

in a quiet room. The acoustic signal was obtained with a head-mounted telex unidirectional microphone and digitized, stored, and analyzed using a Kay Elemetrics model 4300B Computerized Speech Lab System. Subjects were asked to produce three “uh” productions at their lowest pitch, comfortable pitch, and high pitch. The average of the middle 3 s at each pitch was used to calculate the fundamental frequency. Immediately following the voice recording, the subjects were taken to the fMRI suite.

fMRI pitch variability tasks

A single-trial paradigm was employed in which voice generation tasks were performed briefly once every 16 s. The experimental tasks consisted of three voluntary voice production tasks rest as a baseline condition. Specifically, subjects generated 10 trials of the “uh” sound (production of “uh” using a neutral lingual position and without concomitant labial movement) for a maximum duration of 4 s followed by 16 s of rest. The “uh” task was performed at three pitches: low, comfortable, and high corresponding with their productions during the voice evaluation. These productions were followed by a fixation point for a time interval of 16 s. The cycle then repeated for a total of 10 trials. These tasks were pseudo-randomized between individuals.

Data acquisition and fMRI procedures

All images were acquired with a 3T GE Signa LX Scanner (GE Medical System, Milwaukee, WI) with a standard birdcage head coil. Twenty-six contiguous slices parallel to the anterior–posterior commissure (AC–PC line) and covering the whole brain were acquired using a gradient echo echo-planar imaging (EPI) sequence with the following parameters: 2000 ms TR; 30 ms TE; 90° flip angle; 240 mm FOV; 4.5 mm slice thickness; 128×128 matrix. An additional eight excitations without image collection (16 s) prefaced the initiation of the functional imaging run to allow the MR signal to reach equilibrium. High order shimming prior to the EPI data acquisition optimized the field homogeneity across the brain. Anatomical images were obtained using T1-weighted spin-echo images (400 ms TR; 14 ms TE; 90° flip angle; 256×256 matrix; 4.5 mm thickness; 26 slices; 240 mm FOV) and a T1-weighted 3D-spoiled GRASS sequence (6.9 ms TR; 3 ms TE; 15° flip angle; 256×256 matrix; 1.5 mm thickness; 124 slices; 240 mm FOV). The functional EPI and anatomical T1-weighted spin-echo images were acquired with the same number of sections and orientation for anatomical co-registration of observed activations.

E-Prime (Psychology Software Tools, Pittsburgh, PA) was used to control and present experimental stimuli to the subjects, control scanner initialization sequences, and coordinate timing of stimulus presentation with scanner operations. Subjects were instructed to perform the tasks, which were displayed on the screen visible with custom-designed glasses. A microphone was positioned close to the subject's mouth. An observer monitored the subject's task performance (including pitch levels, start/end timing for vocalization) by listening the pitch generation through the speaker attached E-prime equipment. The subject's brain activity and head motion were monitored using software (Brainwave, Medical Numerics, installed in GE Functool) that permits the observation in real time.

Image processing and data analysis

Image processing and statistical analysis were performed with Analysis of Functional Neuro Images (AFNI) software package (<http://afni.nimh.nih.gov>) (Cox, 1996). The reconstructed fMRI data were aligned to the volume closest in time to the high-resolution anatomical scan using a 3D rigid-body registration method. Spatial smoothing using a Gaussian blur with a full width at half maximum (FWHM) of 4 mm was applied. Normalization was applied to reduce the variability in the way subject's respond to a stimulus presentation.

Table 1Volume of significant clusters (in μl) within brain regions activated during the different pitch modulations (a voxel-level corrected threshold of $p < 0.005$ was used)

Brain region	R/L	Comfortable pitch		High pitch		Low pitch	
		Coordinates	Volume (μl)	Coordinates	Volume (μl)	Coordinates	Volume (μl)
Cerebellum	R	(30 -54 15)*	133	(29 -52 -15)*	546	(31 -69 -15)*	102
	L	(-20 -59 -21)*	741	(-11 -44 -21)**	1417	(-8 -44 -21)**	111
Stg	R	(46 -39 8)	2983	(48 -27 8)	913	(41 -23 8)	737
	L	(-55 -39 8)	2364	(-58 -39 8)	374	(-60, -36, 19)	199
Insula	R	(-29 -3 18)	1989	(-32 -10 18)	398	(-36 -18 18)	732
	L	(37 -5 18)	336	(45 -3 18)	169	(36 -24 12)	3196
Ifg	R	(29 32 8)	1127	(44 23 8)	739	(28 34 8)	719
	L		116	(-46 7 15)	187	(-50 4 15)	319
Precentral gyrus	R	(58 -9 30)	3054	(53 1 30)	1148	(58 -6 30)	1084
	L	(-46 -14 27)	2973	(-48 -10 27)	2483	(-45 -13 27)	1838
Postcentral gyrus	R	(62 -9 19)	1725	(54 -28 21)	120	(55 -9 21)	605
	L	(-52 -14 21)	2390	(-51 -12 25)	658	(-48 -16 21)	1038
Ipl	R	(42 -33 26)	261	(56 -31 26)	644	(59 -31 26)	423
	L	(-46 -39 26)	692	(-45 -41 26)	187	(-55 -36 26)	1251
Thalamus	R					(15 -8 6)	104
Posterior cingulate	R	(9 -64 10)	265	(27 -55 19)	42	(19 -8 26)	247
	L	(-17 -66 10)	178	(-22 -59 12)	179	(-18 -63 11)	210
Putamen	L					(-22 -0 11)	230
Medial FG	L					(-5 -6 51)	195
Supramarginal gyrus	R	(40 -40 32)	116				
	L	(-44 -38 31)	322	(-40 -38 32)	111	(-48 -41 32)	185
Middle FG	R	(36 49 7)	790			(42 44 7)	333

The center of mass for each cluster is given in x, y, and z coordinates according to Talairach and Tournoux (1988). (Abbreviations; STG: superior temporal gyrus; IFG: inferior frontal gyrus; IPL: inferior parietal lobe; FG: frontal gyrus; R/L right/left; *:Lobule VI; **:Lobule V).

Deconvolution analysis (Saad et al., 2006) was applied to estimate the impulse response function (IRF) of the fMRI signal on a voxel-wise basis. Estimate of the IRF began at the onset of the voice generation,

and the IRF was estimated to 16 s (8 images). This analysis produced an estimated hemodynamic response relative to the baseline state without making assumptions regarding the shape, delay, or

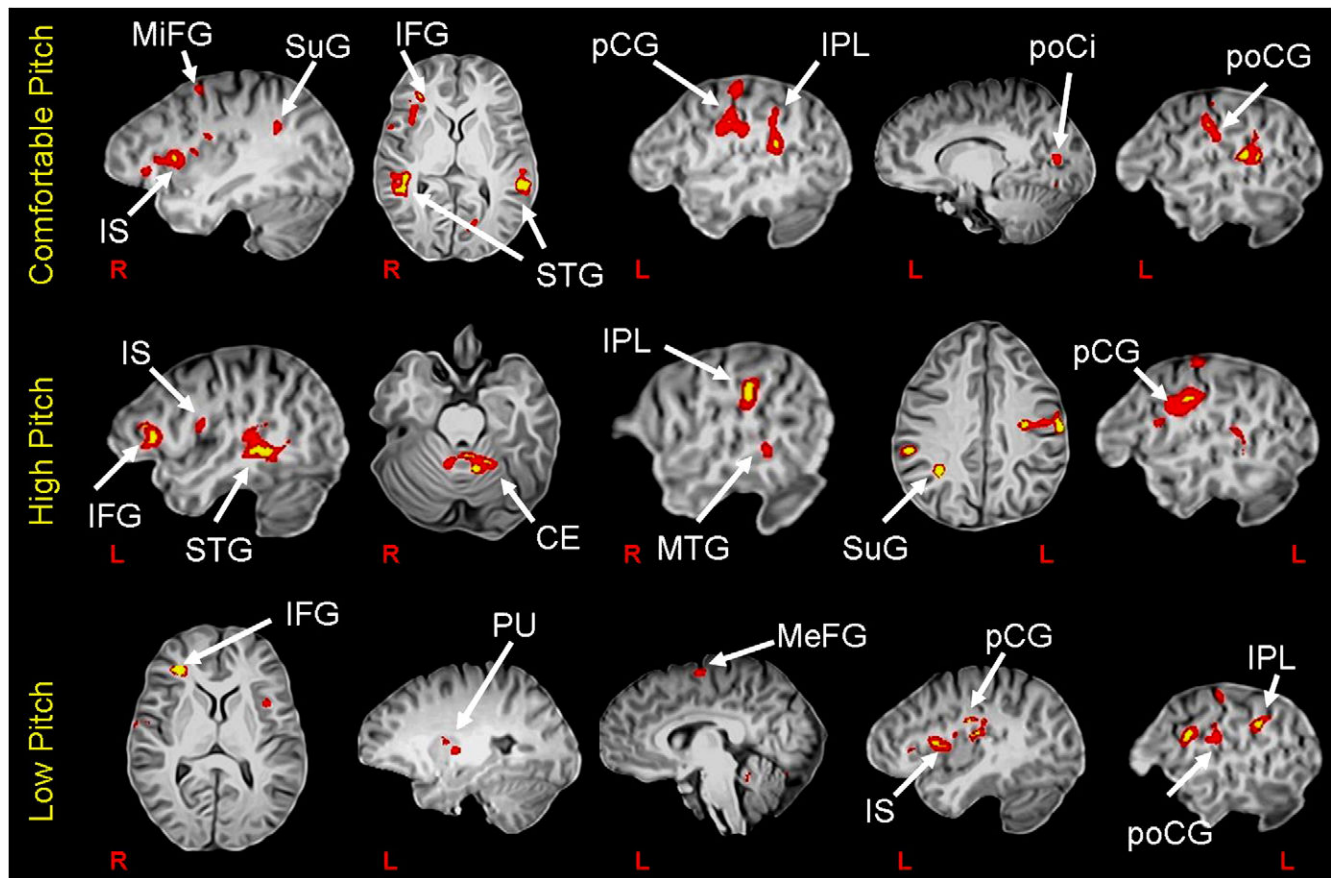


Fig. 1. Cortical responses obtained from voxel-wise comparisons of the comfortable (top row), high (middle row) and low pitch (bottom row) tasks based on averaged group data. The arrows in the axial and sagittal images indicate clusters of significant activation ($p < 0.005$, 100 mm^3). (Abbreviations: MiFG: middle frontal gyrus; IS: insula; SuG: supramarginal gyrus; IFG: inferior frontal gyrus; STG: superior temporal gyrus; pCG: precentral gyrus; IPL: inferior parietal lobe; poCi: posterior cingulate; poCG: postcentral gyrus; CE: cerebellum; MTG: middle temporal gyrus; PU: putamen; MeFG: medial frontal gyrus).

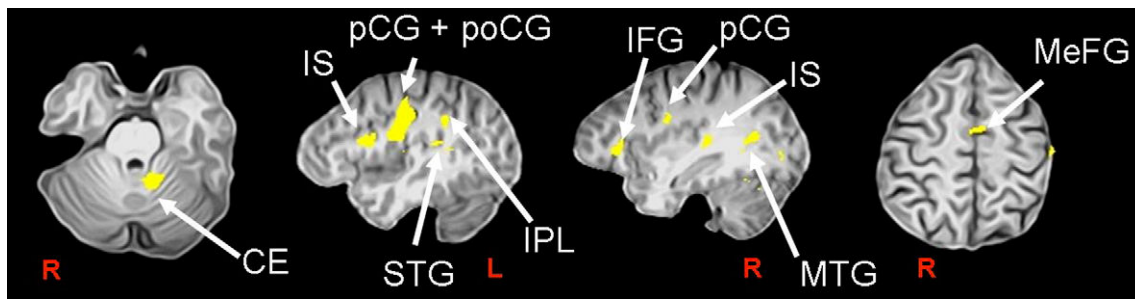


Fig. 2. Common activation during pitch tasks relative to a resting period. Spatially normalized activation was projected onto the single-subject template in the Talairach–Tournoux standard space. Yellow represents common areas within the group activated during three pitch tasks (probability threshold $p < 0.005$ and cluster size of 100 mm^3). (Abbreviations: CE: cerebellum; IS: insula; IFG: inferior frontal gyrus; STG: superior temporal gyrus; pCG: precentral gyrus; IPL: inferior parietal lobe; poCG: postcentral gyrus; MTG: middle temporal gyrus; MeFG: medial frontal gyrus).

magnitude of the impulse response function. The estimated hemodynamic response was subsequently convolved with the temporal sequence of voice generation blocks. Goodness of fit of multiple linear regressions with the original time series was measured using the coefficient of determination (Saad et al., 2006). The significance of activation was assessed through the calculation of the F -statistic for regression. The inherent difference in the time scales of blood oxygenation-level dependent (BOLD) signal and voice induced signal changes was used to minimize motion-induced susceptibility artifacts by discarding the first 1–2 images after voice production and was incorporated with the deconvolution analysis. Additionally, to reduce false positive activity due to large venous structures or other artifacts, voxels were set to zero where the standard deviation of the acquired time series exceeded eight percent of the mean signal intensity. The three dimensional anatomical data sets for each subject were spatially normalized and converted to the standard Talairach and Tournoux atlas space with 1 mm^3 voxels (Talairach and Tournoux, 1988). The functional time series data sets were then transformed into standard stereotaxic space.

For group analysis, the area under the curve was calculated based on each individual subject's data using the IRF obtained from the deconvolution analysis. To minimize intersubject spatial variation, a spatial smoothing (FWHM=4 mm) was applied. Group statistical maps were produced as t -scores of relative signal change between active and baseline tasks compared to a null hypothesis of no change with the use of a t -test. Functional activation maps for the t -statistic were averaged across all subjects. In order to address the multiple comparison correction, AlphaSim (in AFNI) using Monte-Carlo simulations (Forman et al., 1995) was performed. Using this method, a combination of clustering and thresholding produced a false positive discovery rate, α was determined. The AlphaSim output indicated that clusters smaller than 100 mm^3 should be rejected at a corrected voxel-wise p -value of $\alpha \leq 0.005$. Then, voxel-wise t -tests were applied to perform the conjunction and contrast analyses. ROI analysis was performed to calculate the volume of activation within the area. A conjunction analysis was performed to localize the overlapping regions in the brain among a set of three (low, comfortable, and high) contrasts. Contrasts showing (high vs. comfortable pitch) and

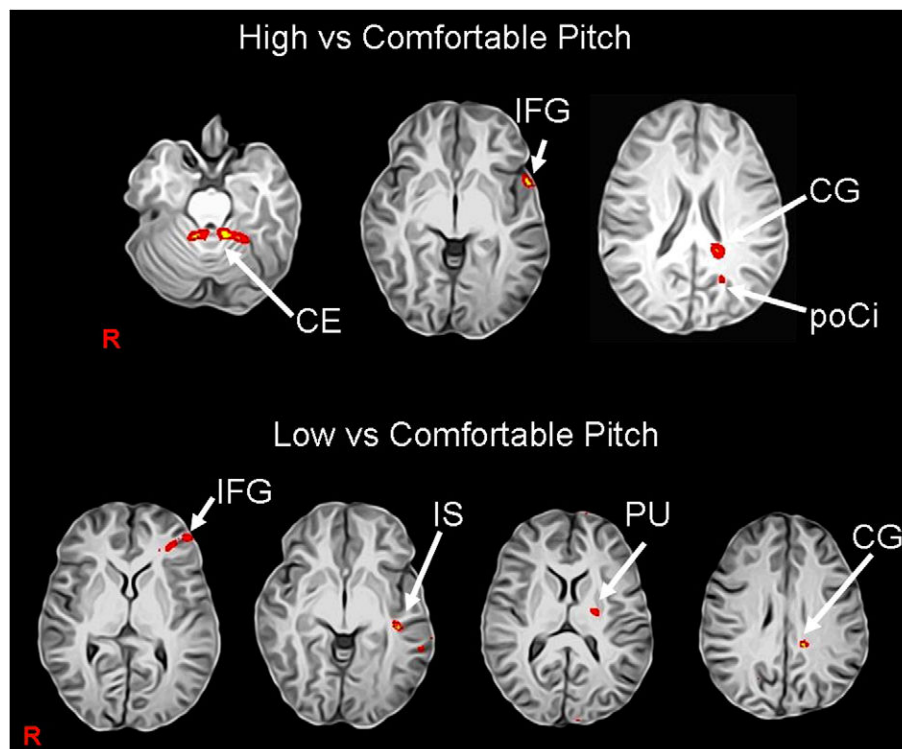


Fig. 3. Paired t -test comparisons of 1) high pitch versus comfortable pitch and 2) low pitch versus comfortable pitch (probability threshold $p < 0.005$ and cluster size of 100 mm^3). (Abbreviations: CE: cerebellum; IS: insula; IFG: inferior frontal gyrus; pCG: precentral gyrus; CG: cingular gyrus; PU: putamen).

(low vs. comfortable pitch), were calculated at voxel probability threshold $p < 0.005$ and cluster size of 100 mm³.

Results

Voice evaluation

Acoustic analyses using the Multidimensional Voice Program (MDVP; Kay Elemetrics, Lincoln Park, NJ) revealed the average, comfortable fundamental frequency to be 118.8 and 180.3 Hz for men and women, respectively. These outcomes correspond with previously published normative data regarding fundamental frequency (Stoicheff 1981; Hollien and Shipp, 1972). Subjects produced comfortable, high (average 464.7 and 500 Hz for men and women respectively), and low pitches (average 98.4 and 147.1 Hz for men and women respectively) at the time of the assessment outside the scanner and then reproduced these pitch variations in the scanner.

Activations across tasks

The areas of significant activation common to comfortable, high and low pitch tasks are shown in Table 1. Volume of activation was measured using the ROI analysis. A cluster of bilateral activation included cerebellum, superior temporal gyrus, insula, precentral gyrus, postcentral gyrus, inferior parietal lobe, post-cingulate gyrus (Table 1, Fig. 1). Activities in the medial frontal gyrus, thalamus and putamen were shown during the low pitch task only, possibly reflecting a need for greater activity to produce and control lower frequencies. Hemispheric differences in the inferior frontal gyrus and superior temporal gyrus showed more activation in the right hemisphere relative to the left hemisphere. This result was consistent across all tasks and concurred with previous reports (Hsieh et al., 2001; Wong et al., 2007).

Common activations across each condition

Conjunction analysis using paired *t*-tests showed several common activated regions across all pitch production tasks. These regions are shown in Fig. 2. Bilateral activations included the precentral gyrus, insula, superior temporal gyrus. Cerebellum, inferior parietal lobe, postcentral gyrus, and medial frontal gyrus in the left hemisphere were detected as active. Inferior frontal gyrus, middle temporal gyrus were detected in the right hemisphere.

Activations differentiating comfortable pitch

High pitch compared to comfortable pitch contrast (Fig. 3, top) showed activation in the bilateral cerebellum, left inferior frontal gyrus, left cingulate gyrus, and left posterior cingulate.

Low pitch compared to comfortable pitch contrast (Fig. 3, bottom) showed activation in the inferior frontal gyrus, insula, putamen, and cingulate gyrus in the left hemisphere.

Discussion

The neural networks associated with pitch modulation during voice production are poorly characterized. To address this issue, we employed a single-trial design. This design effectively identified significant differences in functional MR images in the presence of expected task-related subject movements during voice production of varying pitches. A network of regions was found to be differentially coupled with pitch variability. Our data suggest parallel variances in the neural areas related to different pitch productions and may reflect differences in the activation and monitoring of cranial nerve activity and laryngeal musculature during phonation. The current study does not directly address the problem of identifying regional activations

that contribute to the rapid, sequential pitch changes that occur during speaking and singing; however, the findings provide novel information regarding the differences in regional activations and BOLD signal changes associated with vocal pitch production and control for three different frequencies produced in isolation. In addition, the functional tasks used in the current study (production of “uh” at 3 vocal frequencies) successfully avoided methodological confounds between voice, speech and language tasks by utilizing a task which requiring a neutral lingual position, involving voicing only, and having no related speech or language overlay (Galgano and Froud, 2008).

The functional tasks were chosen because we sought to avoid movement artifacts involved with speech and speech-like movements such as lip-pursing for vowel-production, in particular, front/back, tense, or rounded vowels (such as /l/ or /u/), or sounds which require extensive mandibular excursion (such as /a/). Prior to fMRI scanning, these functional paradigms were practiced outside of the scanner until the subject could execute the task successfully. Although respiration-induced changes were not monitored or removed during this study, several areas of task-associated activations were found to be common across all phonatory tasks.

Several strategies have contributed to the development of fMRI protocols which involve overt voice production. It is well-known that task-associated non-rigid motion can cause signal artifacts that appear as false activation signals (Hajnal et al., 1994). This artifact is associated with head movement during voice production in the fMRI environment. Such artifact may also occur in response to changes in laryngeal posture associated with pitch modulation which cause movement outside the field of view. In the past, image motion artifacts during speech tasks in block-design fMRI paradigms have significantly limited the viability of fMRI for speech (Birn et al., 1998, 1999). In the present study, a fMRI paradigm using a brief overt voicing task and single-trial design was employed. The blood oxygenation-level dependent (BOLD) hemodynamic response generated in this way differs noticeably in the shape and time-course from motion-related signals, which tend to occur during the actual performance of voicing and have a sharper peak. The inherent difference in the time scales of the actual hemodynamic response and pitch-induced signal changes was incorporated with the data analysis to minimize pitch production-correlated motion artifacts by discarding the first few images after each production. However, as a limitation of analysis in the current study, omitting the first several images acquired during pitch production may not yield satisfactory outcomes if there is a significant overlap between the delayed task correlated motion signal and IRF. This overlap may be due to a reduction of task correlated motion artifact which can also mask the true BOLD signal. Therefore, during analysis, if the task is correlated, motion and IRF are not temporally separated resulting in a loss of sensitivity for true positive brain activation. In this study, the 4 s task length may be being too long to describe “pitch modulation” during speaking, in which very rapid adjustments can be made more quickly than every 4 s. However, this length allows for description of the ability to “control” laryngeal adjustments and position for the production of different frequencies. Considering these factors, the results of the current study lay the groundwork for future studies examining pitch control during speech. An additional methodological concern in the current study is that subject performance (pitch production) was not recorded during scanning. Instead, productions were confirmed by an observer (the same speech pathologist that recorded the productions during the evaluation) for qualitative comparison and judgment of pitch. This potential confound should also be addressed in future studies.

In accordance with previous findings, an event-related fMRI design to differentiate BOLD signal changes from motion-induced changes led to strong BOLD signal detection power in a paradigm involving vocal production (Birn et al., 2004). Specifically, the use of event-related, single-trial designs has been reported to be an appropriate

choice for investigations involving brief periods of speech production followed by rest periods long enough to allow for the completion of the hemodynamic response (Birn et al., 1999).

During modal pitch production, bilateral activations were observed in the cerebellum, superior temporal gyrus, insula, precentral gyrus, postcentral gyrus, inferior parietal lobe, and post-cingulate gyrus. Activations in the medial and middle frontal gyri were also observed. Regions active during high pitch production when compared to comfortable pitch were the bilateral cerebellum, left inferior frontal gyrus, left cingulate gyrus, and left posterior cingulate. During low pitch generation, activations were present in the inferior frontal gyrus, insula, putamen, and cingulate gyrus in the left hemisphere.

These findings are consistent with previous findings of studies examining human vocalization (Schulz et al., 2005; Loucks et al., 2007; Galgano et al., in press). Activity found in the middle frontal gyrus is consistent with investigations reporting activations for motor preparation and planning (e.g., Galgano et al., in press; Pedersen et al., 1998). Activation in the supplementary motor area (SMA: medial frontal gyri) is also known to play a role in the preparation of volitional movements, indicating an integral role in motor planning (Hirano, 2001).

The middle temporal gyrus, as well as the superior temporal gyrus, have also been found to be active during vocal self-monitoring (Schulz et al., 2005) and voice processing (Belin et al., 2000), respectively. Therefore, it is possible that these regions are involved in both the production and auditory-perceptual processing of voice. In addition to its involvement during voice processing, the superior temporal gyrus is reportedly involved in audiomotor integration for vocal production (Hickok et al., 2003), an important component in vocal control.

The insula was also active across all tasks. This area has been reported to be implicated in the detection of sound (Bamiou et al., 2003), an integral component of vocal-monitoring. This area was also observed during overt and covert singing and speaking tasks, and known to integrate sound structure or tonal patterns, respectively, with a speaker's emotions and attitudes (Riecker et al., 2000). Greater activity in the insula was also evident during low pitch production vs. comfortable production, possibly indicating the need for additional integration of multi-modal sensory information.

Cingulate cortex activity was also present in all conditions. Anterior cingulate cortex (ACC) activity is associated with motor control (Picard and Strick, 1996), necessary for phonation, especially during pitch modulation. The ACC has also been shown to be involved in regulating vocalization induced by the periaqueductal gray (PAG) in animal models (Jurgens and Lu, 1993; Jurgens and Zwirner, 1996). However, PAG activity was not observed in the present study.

Pitch variability during phonation was also associated with activations in the Rolandic operculum (postcentral gyrus). This area has been previously reported to be associated with pitch control via regulation of elevation and depression of the larynx (Vilkman et al., 1996). In addition, activity in this region demonstrates the need for utilization of sensory information during voicing. Inferior parietal lobe activity exhibited most likely represents a sensorimotor integration function necessary for voice production and pitch control. Regions active during high pitch production when compared to comfortable pitch showed significantly greater activity in the bilateral cerebellum, an area known to be involved in motor planning and coordination (Ito, 2000). Additionally, the left inferior frontal gyrus, involved in monitoring speech (Shergill et al., 2002) showed increased volumes of activation during high and low pitches, when compared with modal pitch. Interestingly, during high pitch production vs. comfortable production and low pitch production vs. comfortable production (Fig. 3), activation in the left hemisphere was greater than in the right, except for bilateral activation in the cerebellum. Bilateral cerebellum activation has been described previously during voice and non-voicing tasks (Brown et al., 2008) as well as word production and singing (Riecker et al., 2000). Additionally, stereotaxic coordinates for

bilateral activations in the precentral gyrus during pitch production showed good agreement with the coordinates of the ventromedial larynx/phonation area described previously (Brown et al., 2008). Meta analysis conducted by Indefrey and Levelt (2004) on single word productions against a picture naming task showed eight reliable areas, including the left anterior cingulate, right anterior insula, left lentiform nucleus, left dorsal precentral gyrus, left anterior and posterior middle frontal gyri, and left posterior medial frontal gyrus. Inconsistent unilateral activation of cortical regions during the tasks was observed in the current study. These regions include the thalamus, putamen, medial frontal gyrus and middle frontal gyrus. An inconsistent pattern of unilateral or bilateral activation shown in the current study may be due to the rapid, coordinated movements of many paired (i.e. bilateral) and unpaired laryngeal muscles during speech.

Many of the same active regions in the present study have been described in studies investigating oral-motor movements, including speaking and swallowing (Fong et al., 2004; Hamdy et al., 1999). The non-specificity of the regions of activation found in this study are not surprising, considering the need for movement and control of many of the same structures during, for example, swallowing and voicing. Both of these activities require movement of various intrinsic and extrinsic muscles of the larynx, tongue and jaw.

Demonstration of increased activation in the left hemisphere that is associated with prosodic features of speech concur with the concept that pitch variability may be a primitive form of language expression. In our subjects, all of whom are right-handed, essentially all of the new activation caused by a change in pitch was shown in the left hemisphere, known to be dominant for language. Further studies are warranted to look further into this evidence of slight left-lateralization of activity during human pitch production.

A potential limitation of the current investigation is the possibility of active inhibition of ROI activity during pitch control causing task-induced deactivation (Shulman et al., 1997; Binder et al., 1999). A previous imaging study has shown increased magnitude of deactivation during more difficult conditions relative to the easier conditions associated with short-term memory load and target discriminability manipulations (McKiernan et al., 2003). Utilization of a limited number of trials in an attempt to limit total scanner time for each subject may have affected the number of active regions shown for each functional task examined. However, the differences in the active regional networks during high and low when compared with comfortable productions begin to delineate important differences in the neural activity underlying voice production with increased demands (e.g. high and low pitch levels).

Acknowledgments

Funding for the work described in the current manuscript was provided by the Langeloth Foundation, The Garban Fund, and Hackers for Hope.

References

- Bamiou, D.E., Musiek, F.E., Luxon, L.M., 2003. The insula (Island of Reil) and its role in auditory processing. Literature Review. Brain Res. Brain Res. Rev. 42 (2), 143–154.
- Belin, P., Zatorre, R.J., Lafaille, P., Ahad, P., Pike, B., 2000. Voice-selective areas in human auditory cortex. Nature 20 (403(6767)), 309–312.
- Binder, J.R., Frost, J.A., Hammeke, T.A., Bellgowan, P.S.F., Rao, S.M., Cox, R.W., 1999. Conceptual processing during the conscious resting state: a functional MRI study. J. Cogn. Neurosci. 11, 80–93.
- Birn, R.M., Bandettini, P.A., Cox, R.W., Jesmanowicz, A., Shaker, R., 1998. Magnetic field changes in the human brain due to swallowing or speaking. Magn. Reson. Med. 40 (1), 55–60.
- Birn, R.M., Bandettini, P.A., Cox, R.W., Shaker, R., 1999. Event-related fMRI of tasks involving brief motion. Hum. Brain Mapp. 7 (2), 106–114.
- Birn, R.M., Cox, R.W., Bandettini, P.A., 2004. Experimental designs and processing strategies for fMRI studies involving overt verbal responses. Neuroimage 23 (3), 1046–1058.

- Brin, M.F., Blitzer, A., Stewart, C., 1998. Laryngeal dystonia (spasmodic dysphonia): observations of 901 patients and treatment with botulinum toxin. *Adv. Neurol.* 78, 237–252.
- Brown, S., Ngan, E., Liotti, M., 2008. A larynx area in the human cortex. *Cereb. Cortex* 18, 837–845.
- Cox, R.W., 1996. AFNI: software for analysis and visualization of functional magnetic resonance imaging. *Comput. Biomed. Res.* 29, 162–173.
- Fong, J., Sandhu, G., Ellaway, P., Davey, N., Strutton, P., Murphy, K., Guz, A., 2004. What do we know about human cough? *Pulm. Pharmacol. Ther.* 17, 431–434.
- Forman, S.D., Cohen, J.D., Fitzgerald, M., Eddy, W.F., Mintun, M.A., Noll, D.C., 1995. Improved assessment of significant activation in functional magnetic resonance imaging (fMRI): use of a cluster-size threshold. *Magn. Reson. Med.* 33, 636–647.
- Galgano, J., Froud, K., 2008. Evidence of the voice-related cortical potential: an electroencephalographic study. *Neuroimage* 15 (41), 1313–1323.
- Galgano, J., Peck, K.K., Branski, R.C., Bogomolny, D., Mener, D., Ho, M., Holodny, A.I., Kraus, D.H., in press. Correlation between functional MRI and voice improvement following type I thyroplasty in unilateral vocal fold paralysis—a case study. *J. Voice*.
- Gopinath, K., Crosson, B., McGregor, K., Peck, K., Chang, Y., Moore, A., Sherod, M., Cavanagh, C., Wabnitz, A., Wierenga, C., White, K., Cheshkov, S., Krishnamurthy, V., Briggs, R., in press. Selective detrending method for reducing task-correlated motion artifact during speech in event-related fMRI. *Hum. Brain Mapp.*
- Grillone, G.A., Chan, T., 2006. Laryngeal dystonia. *Otolaryngol. Clin. North Am.* 39 (1), 87–100.
- Hajnal, J., Myers, R., Oatridge, A., et al., 1994. Artifacts due to stimulus correlated motion in functional imaging of the brain. *Magn. Reson. Med.* 31, 283–291.
- Hamdy, S., Mikulis, D.J., Crawley, A., Xue, S., Lau, H., Henry, S., Diamant, N.E., 1999. Cortical activation during human volitional swallowing: an event-related fMRI study. *Am. J. Physiol.* 277 (1 Pt 1), G219–G225 Jul.
- Hickok, G., Buchsbaum, B., Humphries, C., Muftuler, T., 2003. Auditory-motor interaction revealed by fMRI: speech, music, and working memory in area Spt. *J. Cogn. Neurosci.* 15 (5), 673–682.
- Hirano, S., 2001. Cortical dysfunction of the supplementary motor area in a spasmodic dysphonia patient. *Am. J. Otolaryngol.* 22 (3), 219–222.
- Hixon, T.J., Weismer, G.G., Hoit, J.D., 2008. *Preclinical Speech Science: Anatomy, Physiology, Acoustics, and Perception*. Plural Publishing, San Diego CA.
- Hsieh, L., Gandour, J., Wong, D., Hutchins, G.D., 2001. Functional heterogeneity of inferior frontal gyrus is shaped by linguistic experience. *Brain Lang.* 76 (3), 227–252.
- Hollien, H., Shipp, T., 1972. Speaking fundamental frequency and chronologic age in males. *Speech Hear. Res.* 15 (1), 155–159.
- Indefrey, P., Levelt, W.M.J., 2004. The spatial and temporal signatures of word production components. *Cognition* 92, 101–144.
- Ito, M., 2000. Mechanisms of motor learning in the cerebellum. *Brain Res.* 886, 237–245.
- Jurgens, U., Lu, C., 1993. The effects of periacqueductally injected transmitter antagonists on forebrain-elicited vocalization in the squirrel monkey. *Eur. J. Neurosci.* 5, 735–741.
- Jurgens, U., Zwirner, P., 1996. The role of periacqueductal gray in limbic and neocortical vocal fold control. *Neuroreport* 7, 2921–2923.
- Lattner, S., Meyer, M.E., Friederici, A.D., 2005. Voice perception: sex, pitch, and the right hemisphere. *Hum. Brain Mapp.* 24, 11–20.
- Loucks, T.M., Poletto, C.J., Simonyan, K., Reynolds, C.L., Ludlow, C.L., 2007. Human brain activation during phonation and exhalation: common volitional control for two upper airway functions. *Neuroimage* 36 (1), 131–143.
- McKiernan, K.A., Kaufman, J.N., Kucera-Thompson, J., Binder, J.R., 2003. A parametric manipulation of factors affecting task-induced deactivation in functional neuroimaging. *J. Cogn. Neurosci.* 15 (3), 394–408.
- Oldfield, R.C., 1971. The assessment and analysis of handedness: the Edinburgh inventory. *Neuropsychologia* 9, 97–113.
- Patel, A.D., Balaban, E., 2001. Human pitch perception is reflected in the timing of stimulus-related cortical activity. *Nat. Neurosci.* 4 (8), 839–844.
- Pedersen, J.R., Johanssen, P., Bak, C.K., Kofoed, B., Saermark, K., Gjedde, A., 2008. Origin of human motor readiness field linked to left middle frontal gyrus by MEG and PET. *NeuroImage* 8, 214–220.
- Picard, N., Strick, P.L., 1996. Motor areas of the medial wall: a review of their location and functional activation. *Cereb. Cortex* 6 (3), 342–353.
- Poletto, C.J., Verdun, L.P., Strominger, R., Ludlow, C.L., 2004. Correspondence between laryngeal vocal fold movement and muscle activity during speech and nonspeech gestures. *J. Appl. Physiol.* 97 (3), 858–866.
- Riecker, A., Ackermann, H., Wildgruber, D., Dogil, G., Grodd, W., 2000. Opposite hemispheric lateralization effects during speaking and singing at motor cortex, insula and cerebellum. *Neuroreport* 11, 1997–2000.
- Saad, Z.S., Chen, G., Reynolds, R.C., Christidis, P.P., Hammett, K.R., Bellgowan, P.S.F., Cox, R.W., 2006. FIAC Analysis According to AFNI and SUMA. *Hum. Brain Mapp.* 27, 417–424.
- Schulz, G.M., Varga, M., Jeffries, K., Ludlow, C.L., Braun, A.R., 2005. Functional neuroanatomy of human vocalization: an H215O PET study. *Cereb. Cortex* 15 (12), 1835–1847.
- Shergill, S.S., Brammer, M.J., Fukuda, R., Bullmore, E., Amaro Jr., E., Murray, R.M., McGuire, P.K., 2002. Modulation of activity in temporal cortex during generation of inner speech. *Hum. Brain Mapp.* 16(4), 219–227.
- Shulman, G.L., Fiez, J.A., Corbetta, M., Buckner, R.L., Miezin, F.M., Raichle, M.E., Petersen, S.E., 1997. Common blood flow changes across visual tasks: II. Decreases in cerebral cortex. *J. Cogn. Neurosci.* 9, 648–663.
- Stoicheff, M.L., 1981. Speaking fundamental frequency characteristics of nonsmoking female adults. *Speech Hear. Res.* 24(3), 437–441.
- Talairach, J., Tournoux, P., 1988. *A Co-Planar Stereotactic Atlas of the Human Brain*. Thieme, Stuttgart.
- Vilkman, E., Sonninen, A., Hurme, P., Korkko, P., 1996. External laryngeal frame function in voice production revisited: a review. *J. Voice* 10 (1), 78–92.
- Wong, P.C., Perrachione, T.K., Parrish, T.B., 2007. Neural characteristics of successful and less successful speech and word learning in adults. *Hum. Brain Mapp.* 28 (10), 995–1006.