Mapping human laryngeal motor cortex during vocalization

Nicole Eichert[1], Daniel Papp[1], Rogier B. Mars[1,2] & Kate E. Watkins[3]

[1] Wellcome Centre for Integrative Neuroimaging, Centre for Functional MRI of the Brain (FMRIB), Nuffield Department of Clinical Neurosciences, John Radcliffe Hospital, University of Oxford, Oxford, United Kingdom

[2] Donders Institute for Brain, Cognition and Behaviour, Radboud University Nijmegen, Nijmegen, The Netherlands

[3] Wellcome Centre for Integrative Neuroimaging, Department of Experimental Psychology, University of Oxford, Oxford, United Kingdom

Abbreviated title: human laryngeal motor cortex

Correspondence:

Nicole Eichert

Wellcome Centre for Integrative Neuroimaging

Nuffield Department of Clinical Neurosciences

John Radcliffe Hospital

University of Oxford

Oxford OX3 9DU

United Kingdom

Direct Line: +44 (0) 1865 271363

E-mail: nicole.eichert@psy.ox.ac.uk

Abstract

The representations of the articulators involved in human speech production are organized somatotopically in primary motor cortex. The neural representation of the larynx, however, remains debated. Both a dorsal and a ventral larynx representation have been previously described. It is unknown, however, whether both representations are located in primary motor cortex. Here, we mapped the motor representations of the human larynx using fMRI and characterized the cortical microstructure underlying the activated regions. We isolated brain activity related to laryngeal activity during vocalization while controlling for breathing. We also mapped the articulators (the lips and tongue) and the hand area. We found two separate activations during vocalization – a dorsal and a ventral larynx representation. Structural and quantitative neuroimaging revealed that myelin content and cortical thickness underlying the dorsal, but not the ventral larynx representation, are similar to those of other primary motor representations. This finding confirms that the dorsal larynx representation is located in primary motor cortex and that the ventral one is not. We further speculate that the location of the ventral larynx representation is in premotor cortex, as seen in other primates. It remains unclear, however, whether and how these two representations differentially contribute to laryngeal motor control.

Keywords

- Functional MRI
- Larynx
- Myelin
- Quantitative MRI
- Speech

1. Introduction

The voluntary control of highly complex speech movements is regarded as one aspect of behavior that is unique to humans (Fitch 2017). Speech production requires the fine coordination of a large number of muscles to control the supralaryngeal articulators, respiration, and the vocal folds in the larynx during voice production (Jürgens 2002). In addition to its role as main source of vocal sound production, the larynx is implicated in various biological functions such as protection of the airways and swallowing (Ludlow 2005). Several pairs of intrinsic and extrinsic muscles connect the laryngeal cartilages with each other and to the skeleton. During vocalization, these muscles are controlled in a complex fashion, so that the tension in the vocal folds allows them to be set into vibration as air from the lungs passes through them. Several lines of evidence showed that the ventral part of the precentral gyrus and the central sulcus in the human brain are involved in speech motor control (Bohland and Guenther 2006; Ackermann et al. 2014). The question as to which brain areas specifically control laryngeal activity during vocalization, however, remains debated.

Nearly 100 years ago, direct cortical stimulation of the ventral portion of the precentral gyrus in the human brain was shown to elicit vocalization (Foerster 1936; Penfield and Boldrey 1937). In more recent times, functional brain imaging studies show highly inconsistent results when mapping laryngeal activity during vocalization (reviewed in Belyk & Brown, 2017). Several studies report activity evoked by vocalization in both a ventral portion of the precentral gyrus located close to the Sylvian fissure and in a more dorsal portion of central sulcus and precentral gyrus (Terumitsu et al. 2006; Galgano and Froud 2008; Olthoff et al. 2008; Grabski et al. 2012). Some studies report vocalization-evoked activity in the dorsal location only (Sörös et al. 2006;

Brown et al. 2008; Kleber et al. 2013; Belyk and Brown 2014; Belyk et al. 2018). Only a few studies specifically isolated laryngeal activity during voice production by contrasting it with supralaryngeal articulation (Sörös et al. 2006; Terumitsu et al. 2006; Brown et al. 2009; Grabski et al. 2012), but the results were inconsistent across studies.

Another line of evidence comes from direct neurophysiological recordings from the cortical surface using implanted high-density electrode arrays in patients being prepared for epilepsy surgery. These studies also show activity in both a dorsal and a ventral region during speech production (Bouchard et al. 2013; Chang et al. 2013; Toyoda et al. 2014; Breshears et al. 2015; Dichter et al. 2018). Interpretation of electrocorticography (ECoG) studies, however, is inherently limited due to the typically small sample size and potential neurological abnormalities in the patients' brains.

In addition to these inconsistent results, an additional confound in several studies mentioned above is breathing. Voice production and respiration functionally interact during speech production and volitional expiration has been shown to activate motor areas that are located close to the putative dorsal laryngeal representation (Ramsay et al. 1993; Evans et al. 1999; McKay et al. 2003). Several previous studies, however, did not control for breathing (e.g. Brown et al., 2008; Sörös et al., 2006). Those studies that specifically studied vocalization and breathing found largely overlapping activity during both conditions and a contrast showed no difference in motor cortex indicating that the activity was not specific to larynx activity (Loucks et al. 2007; Simonyan et al. 2009).

In addition to a lack of control for confounds such as breathing and supralaryngeal articulation, most neuroimaging studies mentioned above did not assess individual differences in the activation patterns. It is common to report group-level cluster-corrected results following volumetric nonlinear image registration of task-activation maps. Interpreting group-level results, however, might obscure subject-specific features and inter-individual variability, which limits sensitivity and functional resolution (Bennett and Miller 2010; Nieto-Castañón and Fedorenko 2012; Bouchard et al. 2013; Woo et al. 2014). This lack of individual detail might have caused a failure to detect one of the cortical larynx representations in previous studies. Moreover, averaging of small sample sizes with high variability, as often performed in ECoG studies, can show two distinct larynx representations, even when individual patients show activity in only one of the regions (Bouchard et al. 2013).

Comparisons of human brains and those of other primates indicate strong species differences in the neural organization underlying laryngeal control during vocalization (Ackermann et al. 2014; Simonyan 2014; Kumar et al. 2016). Most notably, the location of the proposed human dorsal larynx representation in primary motor cortex is more dorsal-posterior than the non-human primate homolog, which is in a ventral premotor cortex area (Leyton and Sherrington 1917; Hast et al. 1974; Jürgens 1974; Simonyan and Jürgens 2002; Coudé et al. 2011).

These various findings have led to the proposal of an evolutionary 'duplication and migration' hypothesis that the larynx motor cortex comprises two structures located dorsally and ventrally in the human brain (Belyk and Brown, 2017; Jarvis, 2019; reviewed in Mars et al., 2018). This theory proposes that the dorsal larynx representation is unique to humans and that it evolved in primary motor cortex due to our especially high demands on laryngeal motor

control. The ancestral primate larynx representation in premotor cortex is presumed to be homologous to the human ventral larynx representation, which migrated posteriorly, potentially into human primary motor cortex.

The cortical areas underlying the larynx representations, however, are currently unknown and it has not been tested if the ventral larynx representation is in primary motor or in premotor cortex. A myeloarchitectonic approach, which has also become available for neuroimaging, enables us to describe some properties of cortex (Kuehn et al. 2017). Primary motor cortex, for example, is characterized by higher cortical myelin content and higher cortical thickness compared to adjacent premotor cortex and with somatosensory cortex located on the caudal bank of the central sulcus and postcentral gyrus (Fischl and Dale 2000; Glasser and van Essen 2011; Lutti et al. 2014). Describing the anatomical parcels underlying the larynx representations can inform evolutionary hypotheses and provide clues about their functional relevance.

This study sought to determine the anatomical location of larynx-related neural activity in individual subjects and to characterize the cortical structure underlying these representations. Our experimental design aimed to isolate brain activity related to voice production by controlling for breathing-related movements and movements of articulators. In one task, we identified the (supralaryngeal) articulation and the (laryngeal) vocalization component of speech during syllable production using a factorial design described in a previous study (Murphy et al. 1997). We refer to the latter 'vocalization' component as an index for laryngeal activity during voice production, while other studies have referred to it as 'phonation' or 'voicing'. In a second task, we localized the separate neural representations of lip, tongue and

154 larynx using highly controlled basic movements, while breathing movements were matched
155 across conditions.
156
157 In order to characterize the microstructural properties underlying the larynx representation,
158 we compared their myelin content and cortical thickness derived from structural and
159 quantitative MRI measurements. These quantifications give an indication of the type of cortex
160 underlying the activated regions, to inform our knowledge about the organization of the
161 human larynx motor cortex.
162
163
164 ## 2. Materials and Methods
165
166 *Subjects*
167
168 20 subjects (12 females, 18 – 40 years (27.4 ± 5.6, mean ± SD), 5 self-reported left-handers)
169 took part in the study. All subjects were self-reported native English speakers; two were raised
170 bilingually and three were fluent in a second language. All reported normal hearing, normal or
171 corrected-to-normal vision, no neurological impairments and no history or diagnoses of speech
172 disorders. The study was approved by the Central University Research Ethics Committee
173 (CUREC, R55787/RE001) in accordance with the regulatory standards of the Code of Ethics of
174 the World Medical Association (Declaration of Helsinki). All subjects gave informed consent to
175 their participation and were monetarily compensated for their participation.
176

177

178 *Experimental design and task*

179

180 We used two tasks to map the motor representations of the articulators and the larynx. Task
181 1 engaged the speech motor system in an ecologically valid way that required the participants
182 to utter a short syllable sequence in different conditions. Task 2 required the participants to
183 perform basic movements that are commonly used in other localizer studies.

184

185 Subjects practiced all tasks outside the scanner to be sure they understood the task
186 requirements. Articulator movements and vocalizations as well as breathing were
187 demonstrated by the experimenter and practiced until the subjects performed them as
188 required.

189

190

191 *Task 1 - syllable production task*

192

193 Subjects were instructed to produce the utterance "/la leɪ li la leɪ li/" in four different
194 conditions: speaking (overt speech), supralaryngeal articulation only (silent mouthing), overt
195 vowel sound production of vowel /i/ six times but no articulation (vowel production), and
196 thinking (covert speech). The vowel /i/ was chosen because it is a natural and familiar sound
197 that requires laryngeal movement during vocalization, but involves minimal movement of the
198 jaw muscles, lips and pharyngeal part of the tongue (Grabski et al. 2012). Subjects were
199 instructed not to whisper during the 'silent mouthing' or 'covert speech' conditions as this
200 would involve laryngeal motor activity.

For all conditions, the breathing pattern was explicitly instructed using the fixation symbol on the screen (Figure 1A). Subjects were instructed to inhale for 1.5 s (fixation was a square) and exhale for 4 s (fixation was a cross); the utterance in each condition was performed once during each 4-s exhalation. Each condition was performed in blocks lasting 22 s, which corresponded to four repetitions of the breathing cycle. Each block was followed by a rest period of 8 s with normal breathing (i.e. not explicitly instructed). This rest period allowed the subjects to relax their breathing patterns and to maintain a comfortable respiratory state. The four conditions were presented in a fixed pseudo-random order following a balanced Latin-square design wherein each condition was repeated five times; each condition followed and preceded each of the other conditions once and the same condition was not presented consecutively. Subjects were instructed to keep their mouth slightly open during the whole session and to keep the jaw relaxed.

Task 2 - basic functional localizer

There were four task conditions: tongue retraction, lip protrusion, production of the vowel sound /i/, and a 'breathing only' condition. The three task conditions for basic speech movements were contrasted with the 'breathing only' condition in which breathing was explicitly instructed as above. Articulator movements and vowel production were repeated during 4 s of exhalation at a rate of approximately 1 - 2 repetitions/s as described for task 1. Subjects were instructed to keep the movement rate constant throughout the scan for all movement types. Breathing instructions, task timing and randomization of the blocks were the

same as described for task 1 above, except that each condition was repeated four times during the scan run. Note that each 22-s block was followed by an 8-s period of rest with normal breathing, as above.

Hand localizer

During the scanning session, subjects performed the basic localizer first and then the syllable production task. In between these two tasks involving vocalizations, the subjects performed a phonological and semantic judgement task (Devlin et al. 2003). Participants had to indicate a yes/no response by pressing a button with the right index or the middle finger every 3 s. The task data was used here only to localize the hand representation in the left hemisphere; the language task data are not reported.

MRI Data Acquisition

MRI data were obtained at the Oxford Centre for Human Brain Activity (OHBA) using a 3-T Siemens Prisma scanner with a 32-channel head coil. Two structural images of the whole brain were acquired; a T1w image (MPRAGE sequence; 1 mm^3 isotropic resolution, TR = 1900 ms, TE = 3.97 ms, TI = 905 ms, 8° flip angle, bandwidth = 200 Hz/pixel, echo spacing = 9.2 ms, FOV = 192 × 192 × 174 mm^3) and a T2w image (SPACE turbo-spin-echo sequence; 1 mm^3 isotropic resolution, TR = 3200 ms, central TE = 451 ms, variable flip angle, bandwidth = 751 Hz/pixel,

248 echo spacing = 3.34 ms, echo train duration = 919 ms, Turbo Factor = 282, FOV = 256 × 256 ×
249 176 mm^3, GRAPPA acceleration factor 2).
250
251 For task fMRI, whole-head T2*-weighted echo-planar images (TE = 30 ms) were acquired using
252 multiband sequence (factor 6, TR = 0.8) at 2.4 mm^3 isotropic resolution. Task fMRI parameters
253 were adopted from the ABCD study (https://biobank.ctsu.ox.ac.uk/crystal/docs/brain_mri.pdf,
254 Casey et al., 2018). Two task fMRI scans were conducted lasting 8 min (600 volumes, task 2)
255 and 10 min (750 volumes, task 1). In between the two tasks, subjects performed a phonological
256 and semantic judgement task for 9 min, which did not involve vocalization.
257
258 Furthermore, a multiparameter mapping (MPM) protocol was acquired (Weiskopf et al. 2013;
259 Lutti et al. 2014). Proton density-weighted (MPM_{PDw}), magnetization transfer-weighted
260 (MPM_{MTw}) and T1-weighted (MPM_{T1w}) images were acquired using a tailored pulse sequence
261 (1 mm^3 isotropic resolution, FOV = 256 × 224 × 176 mm^3, TR = 25 ms, bandwidth = 488
262 Hz/pixels, first TE/echo spacing = 2.3/2.3 ms, 6° flip angle (MPM_{PDw}, MPM_{MTw}) or 21° (MPM_{T1w}),
263 slab rotation = 30°, and number of echoes = 8/6/8 (MPM_{PDw}/MPM_{MTw}/MPM_{T1w} respectively),
264 GRAPPA acceleration factor 2 × 2, 40 reference lines in each phase encoded direction.
265
266 Quantitative R1 (= 1 / T1) maps were estimated from the MPM_{PDw} and MPM_{T1w} images as
267 demonstrated in Weiskopf et al. (2013), which was extended by including a correction for radio
268 frequency transmit field inhomogeneities (Lutti et al. 2010) and imperfect spoiling (Preibisch
269 and Deichmann 2009). Regression of the log-signal from the signal decay over echoes across
270 all three MPM contrasts was used to calculate a map of R2* (= 1 / T2*) (Weiskopf et al. 2014).
271 The transmit field map was calculated using a 3D echo-planar imaging (EPI) spin-echo

(SE)/stimulated echo (STE) method (Lutti et al., 2012, 2010; FOV = 256 × 192 × 192 mm^3, matrix = 64 × 64 × 48 mm^3, TE = 39.06, mixing time = 33.8 ms, TR = 500 ms, nominal α varying from 115° to 65° in steps of 5°, acquisition time 4 min 24 s) and was corrected for off-resonance effects using a standard B0 field map (double gradient echo FLASH, 3 × 3 × 2 mm^3 resolution, whole-brain coverage). The MPM parameter maps took approximately 20 minutes to acquire. In addition to the MRI scans mentioned here, participants underwent a diffusion-weighted scan and a resting-state scan (data not reported here). The total duration of the entire scanning session was approximately 1.5 h, preceded by approximately 45 min of briefing and task practice.

Physiological recordings were carried out throughout scanning using a respiratory belt to measure the breathing pattern and a pulse oximeter to monitor the heart rhythm during the scan. Physiological data were recorded using a Biopac MP150 (Biopac, Goleta, CA, USA) at a sampling frequency of 500 Hz. Subjects wore ear-plugs and MRI-compatible head phones (OptoActive-II, Optoacoustics Ltd, Moshav Mazor, Israel), which reduced scanner noise using electrodynamic noise-cancelling. At the beginning of the scanning session, the headphones were calibrated and noise-cancelling performance was further monitored throughout the session. Prior to each functional scan the attenuation algorithm 'learned' the scanner noise for 16 s. During the two tasks involving vocalizations, subjects were audio-recorded using an MRI-compatible microphone with noise cancelling (Dual Channel FOMRI-III, Optoacoustics Ltd, Moshav Mazor, Israel) at a sampling rate of 22,050 Hz.

Behavioral data analysis

The subjects' vocal behavior for the tasks was manually assessed using the audio recordings. The breathing patterns during the task blocks recorded using the Biopac were inspected visually to verify that subjects complied with the breathing instruction. Individual breathing traces were cropped into segments of 30 s, which consist of 22 s of instructed breathing during the task block and 8 s of subsequent rest period with normal breathing (Figure 1C).

Structural MRI analysis

T1w and T2w scans were pre-processed using the HCP-pipeline (Glasser et al. 2013) cloned from the 'OxfordStructural' - fork (https://github.com/lennartverhagen/Pipelines). The processing pipeline includes anatomical surface reconstruction using FreeSurfer and automatic assignment of neuroanatomical labels (Fischl 2012; Jenkinson et al. 2012). The T2w image was registered to the T1w image using FSL's FLIRT using spline interpolation (Jenkinson et al. 2002).

The image of the T1w scan was divided by the image of the T2w scan to create a T1w/T2w-ratio image. The T1w/T2w-ratio was mapped onto the native midthickness surface and then resampled to the 164k standard (fs_LR) surface mesh using adaptive barycentric interpolation (approx. 164.000 vertices per hemisphere) using Workbench Command (www.humanconnectome.org/software/connectome-workbench.html). Mapping was performed with the `-volume-to-surface-mapping` command using the `-myelin-style` option.

This T1w/T2w map has been shown empirically to correlate with cortical myelin content (Glasser and van Essen 2011; Glasser et al. 2014). In addition to T1w/T2w myelin maps, the HCP-pipeline provides automatic generation of cortical thickness surface maps.

MPM parameter maps were reconstructed and pre-processed using the hMRI-toolbox (Tabelow et al. 2019) embedded in the Statistical Parametric Mapping framework (SPM12, Wellcome Trust Centre for Neuroimaging, London, UK). For one subject, MPM data were excluded due to motion-induced blurring. MPM_{MT}, MPM_{R1} and MPM_{R2*} maps were registered to the MPRAGE T1w scan using FLIRT rigid-body transformation and spline interpolation and then mapped to the surface using the same steps as for the T1w/T2w myelin map. The three MPM parameter maps have been shown to correlate to different degrees with myelin content in white matter, subcortical structures and grey matter (Draganski et al. 2011; Callaghan et al. 2014; Lutti et al. 2014; Bagnato et al. 2018). Finally, one step of surface-based smoothing (FWHM = 4 mm) was applied to the five surface maps of interest - T1w/T2w map, three MPM parameter maps and cortical thickness map.

fMRI data analysis

Functional MRI data processing was carried out using FEAT (FMRI Expert Analysis Tool) Version 6.00, part of FSL (FMRIB's Software Library, www.fmrib.ox.ac.uk/fsl). The following pre-statistics processing was applied: motion correction using MCFLIRT (Jenkinson et al. 2002); non-brain tissue removal using BET (Smith 2002); spatial smoothing using a Gaussian kernel (FWHM = 5 mm); grand-mean intensity normalization of the entire 4D dataset by a single

multiplicative factor; low-frequency drifts were removed using a temporal high-pass filter with a cut-off of 90 s for all three tasks. Motion corrected images were unwarped using a fieldmap and PRELUDE and FUGUE software running in FSL (Jenkinson 2003). Registration to the high resolution structural scan and standard 2 mm MNI-152 template was carried out using FLIRT (Jenkinson and Smith 2001; Jenkinson et al. 2002). Registration from high resolution structural to standard space was then further refined using FNIRT nonlinear registration (Andersson et al. 2007).

Time-series statistical analysis was based on a general linear model (GLM) implemented in FILM with local autocorrelation correction (Woolrich et al. 2001). Standard motion correction parameters and individual volumes that were motion outliers, determined using fsl_motion_outliers, were included as separate regressors at the first level for each subject.

For the syllable production task (task 1), the conditions were analyzed in a factorial model that allowed the (supralaryngeal) articulation and the (laryngeal) vocalization component of the task to be separated (Figure 1B). Brain activity associated with the control of articulation was defined as ('overt speech' *minus* 'vowel production') *plus* ('silent mouthing' *minus* 'covert speech') and the main contrast for vocalization was derived by the contrast ('overt speech' *minus* 'silent mouthing') *plus* ('vowel production' *minus* 'covert speech'). The 8-s periods of rest with normal breathing between condition blocks served as baseline and were not modelled as a separate explanatory variable in the GLM; they were not included in contrast to any task condition in the analysis of this factorial design.

For the basic localizer task (task 2), activity during each condition for basic speech movements was assessed relative to the 'breathing only' condition with instructed breathing. The periods of rest with normal breathing in between conditions served as baseline (i.e. they were not modelled in the GLM as described for task 1). For the hand localizer task, we derived a contrast of all task conditions involving button presses relative to the rest blocks (note, the breathing instruction was not used in the hand localizer task).

All contrasts reported in the results for both tasks were assessed relative to conditions with matched breathing. In addition, we report results of all individual task conditions relative to the resting condition with normal breathing in a supplementary figure (Figure S1).

Volumetric group average activation maps

To obtain volumetric group average maps, each individual's statistical maps were transformed to standard space (MNI152) using a nonlinear registration. Volumetric group mean activation maps were obtained using mixed effects in FLAME (FMRIB's Local Analysis of Mixed Effects, (Woolrich et al. 2004, Stage 1 only), where subjects are treated as random effects. Group-level *z*-statistic images were thresholded using a voxel-wise threshold of $z > 3.5$ ($p < 0.00025$, uncorrected). Note that the corrected threshold for a whole-brain analysis of these data is $z > 5$. Rather than use a small volume correction and ROI masking, we chose to display results at the lower threshold to better visualize the activity in the whole brain. The figures of the volumetric group-level results (Figure 2, 3) are focused, i.e. centered, on the voxel of maximal *z*-value within the right hemisphere central somatomotor strip. For the vocalization contrast

(Figure 2) and the vowel production condition (Figure 3), we generated two separate figures focused on the right dorsal and the ventral larynx representation.

Surface group count maps

In addition to volume-based analysis, task MRI activations were assessed using surface-based analyses. Surface space permits a better visualization of cortical brain activations. Furthermore, this allowed us to perform surface-based quantifications of cortical microstructures.

Each subject's *z*-statistic images were thresholded voxel-wise at $p < 0.05$ (corrected using a *z*-value determined based on data smoothness and the RESEL count). This thresholded map was projected onto the individual's native midthickness surface using the '-ribbon-enclosed' option in wb_command `-volume-to-surface-mapping`. To derive group-level surface count maps, all individual subject thresholded maps were resampled to the 32k standard (fs_LR) mesh based on the FreeSurfer registration, binarized and then summed at each vertex (Barch et al. 2013). The group count maps are shown on an inflated average brain surface, thresholded at $n > 4$ subjects. They give an indication of the between-subject variability in the areas activated by task, providing complementary information to maps of activity averaged across participants.

Individual surface activation maxima

Surface-based activations were studied further using an ROI-based approach to isolate and focus on activations in the central somatomotor strip. The two tasks provided complementary results for further analysis with Task 1 serving as a robust localizer for the larynx representations, and Task 2 providing an accurate localizer for lip and tongue representations. In order to assess intra-individual spatial variability of the motor representations, we derived the location of individual activation maxima in both hemispheres from selected task contrasts: For the larynx, we used the main contrast for vocalization during the syllable production task (task 1). For lip and tongue, we used the task contrasts from the basic localizer task (task 2) and for the hand, we used the hand localizer task (left hemisphere only).

Different ROI masks were used for the different motor representations based on individual anatomy. In short, we used an ROI of the whole central sulcus for hand, lip and tongue, a more limited portion of the central sulcus ROI for the dorsal larynx representation and a manually defined ROI for the ventral larynx. ROI definitions are described in more detail in the supplementary material.

Individual volumetric ROIs were linearly transformed from FreeSurfer's anatomical to functional space of the respective task fMRI scan. Within the ROI, the voxel of maximal intensity was determined from the uncorrected *z*-statistics image. It should be noted that for some subjects this local maximum did not achieve the corrected voxel-wise significance threshold, which was used to create the surface count maps (left hemisphere: hand $n = 3$, dorsal larynx $n = 6$, ventral larynx $n = 5$; right hemisphere: dorsal larynx $n = 5$, ventral larynx $n = 4$). Using a lower uncorrected threshold is justified given our goal to visualize and assess spatial variability of the activation maxima. Activation maxima were manually inspected in the

subject's native volume space to confirm that the systematic approach described below captured task-related activations. The activation maxima were mapped to the individual's native midthickness surface, resampled to the 32k standard (fs_LR) surface mesh using the FreeSurfer registration, smoothed (FWHM = 1 mm), and binarized to form a small circular patch.

Given that the two tasks had some conditions in common, we were able to examine within-subject reliability of the location of activation maxima for larynx and tongue as described above compared with those derived based on the vowel condition and the main contrast for articulation. We computed the geodesic distance on the individual's native midthickness surface for pairs of maxima (tongue condition and main contrast for articulation; main contrast for vocalization and vowel condition).

Cortical surface features at activation peaks

We described the cortical microstructure at the individual activation peaks based on different surface measures. In order to assess cortical myelin content, we used the T1w/T2w map and the MPM parameter maps (MPM_{MT}, MPM_{R1} and MPM_{R2*}). We extracted the mean value for each individual subject's surface map within the area defined by the circular patch around the vertex of maximal activation for hand, lip, tongue, dorsal larynx and ventral larynx representation, i.e. at the peaks shown in Figure 4A. These measures were obtained from the subject's native surface mesh prior to resampling to the standard mesh. In order to account for the different ranges in intensities and to allow direct comparison of the maps, we computed

461 *z*-scores based on the distribution of individual values in each map and each ROI. We assessed
462 the differences in *z*-scores across the different ROIs using a linear mixed effects analyses as
463 implemented in R's *lmer* function (Bates and Sarkar (2007), Core Team and Foundation for
464 Statistical Computing). The model included fixed effects for ROI (dorsal larynx, lip, tongue,
465 ventral larynx), hemisphere (left, right) and map (T1w/T2w, MPM_{MT}, MPM_{R1}, MPM_{R2*}) and
466 random effects for subject. A Shapiro-Wilk test revealed a normal distribution of the data at a
467 significance level of $p > 0.001$.

468 In addition to cortical myelin content, we also quantified cortical thickness values at the same
469 individual surface peaks. To increase sensitivity of the quantification described above, we
470 excluded activation maxima that were located in a part of cortex with lower cortical thickness
471 and thus likely to be activity in somatosensory rather than motor cortex (Fischl and Dale 2000).
472 We used a heuristic lower threshold of 2 mm to exclude sensory activation maxima and then
473 re-ran the statistical evaluation of myelin-values described above. A linear mixed effects
474 analysis for the fixed effects of ROI and hemisphere and random subject effects was performed
475 to assess cortical thickness values after exclusion of the sensory activation maxima.
476
477 In order to further characterize the differences in myelin content across the ROIs, we used an
478 additional quantification: We computed the pair-wise Manhattan distance across the ROIs
479 based on a vector of the four raw (i.e. non-*z*-transformed) values in each subject after excluding
480 the sensory activation maxima. The differences across ROIs were visualized in form of a
481 dissimilarity matrix, where we averaged Manhattan distances within each ROI first across
482 subjects and then across hemispheres.
483

A linear discriminant analysis (LDA) was run to explore which combination of *z*-transformed surface features (T1w/T2w, MPM_{MT}, MPM_{R1}, MPM_{R2*}), best discriminated the ventral from the dorsal larynx representation after excluding sensory activation maxima. The parameters of the LDA were estimated using a singular value decomposition with no shrinkage.

3. Results

We acquired fMRI data in 20 subjects during performance of two tasks: (1) a syllable production task required subjects to produce "/la leɪ li la leɪ li/" overtly, mouthed silently, and covertly, and to produce the vowel /i/, (2) a basic localizer task required subjects to make small repetitive movements of the lips, tongue and larynx (vowel production); an additional task was used as functional localizer for movement of the right hand.

Auditory recordings

All subjects vocalized as instructed during the conditions that involved vocalizations in both tasks. In all other conditions, subjects remained silent, as instructed.

Breathing during vocalization tasks

We recorded the breathing traces using a breath belt during both vocalization tasks to confirm that the breathing patterns were comparable across different conditions (Figure 1C). As expected for all conditions in both tasks, four breathing cycles were visible in the first 22 s, during which breathing was instructed. Note that in one subject, where an extension of the breathing belt was used, the breathing trace differed in overall shape, but the breathing cycles were still visible. Examination of these figures shows that the shape of the trace and variability were similar across all conditions, but some differences were observed. For example, in the vowel production condition of the basic localizer task, exhalation was more gradual and less rapid than in the other four conditions. All four conditions in the syllable production task showed a more gradual exhalation pattern than in the basic localizer task.

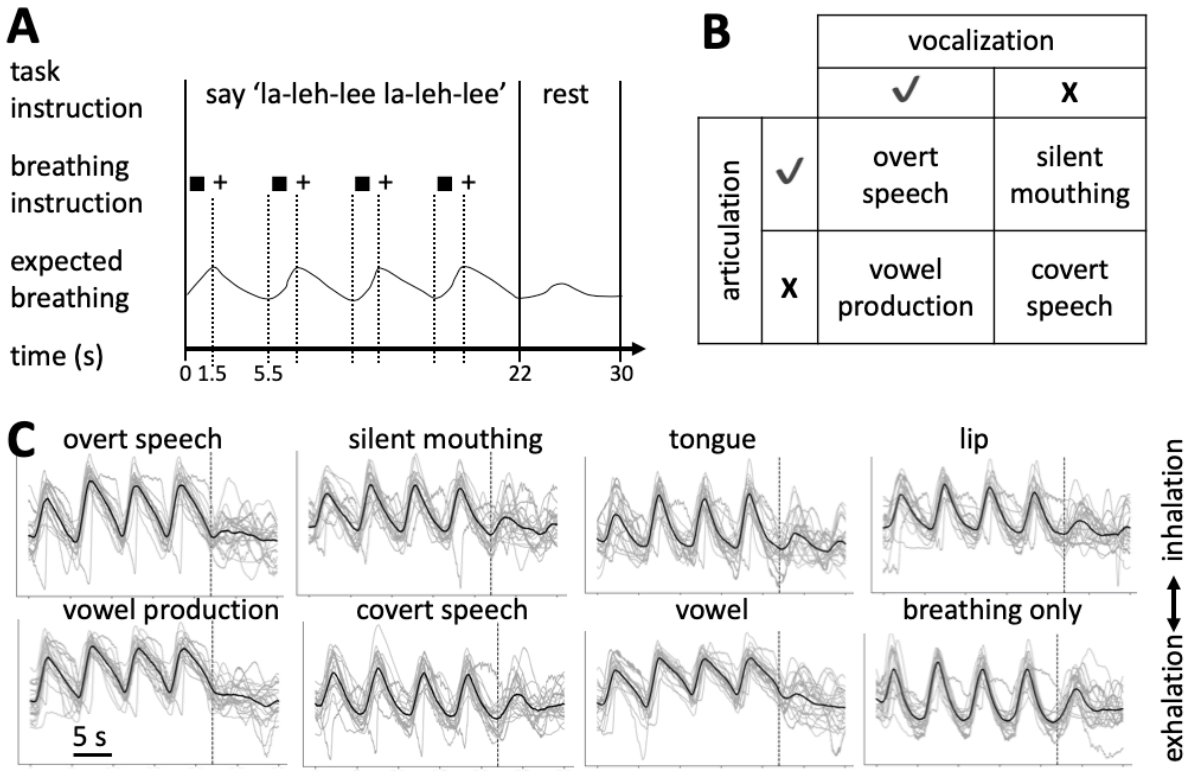


Figure 1 – Task paradigm and breathing patterns. **A**: Timing of the task paradigm for an example block of the 'overt speech' condition from the syllable production task (task 1). The utterance was repeated four times while breathing was instructed followed by a resting period with normal breathing. Square – inhale, + -

522 exhale. **B**: Factorial design used in the syllable production task. The four task conditions can be described
523 according to orthogonal main factors for laryngeal activity during vocalization and supralaryngeal
524 articulation. Breathing was instructed during all conditions as shown in A. **C**: Breathing traces during syllable
525 production task (left two panels) and basic localizer task (right two panels). Black line: group mean after
526 normalizing all individual breathing traces to the same amplitude; grey lines: average traces for individual
527 subjects (*n* = 20). The dashed vertical line indicates the end of the 22-s task block and the beginning of the
528 8-s resting period with normal breathing.

529
530

531 *Syllable production task (task 1)*

532

533 *Syllable production task - Volumetric results*

534

535 We localized the cortical activations for movement control of (supralaryngeal) articulation and
536 (laryngeal) vocalization during syllable production (Figure 2A). MNI coordinates of activation
537 maxima in the somatomotor strip are reported in the supplemental material (Table S1). The
538 main contrasts for vocalization and articulation were defined based on a factorial design that
539 combined the data from all four task conditions in the syllable production task (Figure 1B). The
540 main contrast for articulation showed activity in the mid-portion of the central gyrus. Note that
541 we also found spurious activity in prefrontal white matter, which we presumed was induced
542 by task-correlated movement.

543

544 The main contrast for vocalization showed activity in two somatomotor regions in both
545 hemispheres: one located within a dorsal region of central sulcus, and a second located in a
546 more anterior-ventral region. Portions of the superior temporal gyrus were also activated

547 during vocalization. This is presumed to reflect auditory perception of self-generated
548 vocalizations. The dorsal and ventral activations were separate and did not appear to be
549 connected.
550
551 Vocalization and articulation also activated parts of cerebellum and SMA in a somatotopic
552 fashion. Group-level volumetric results for cerebellum and SMA in both functional tasks are
553 described in the supplementary material (Figure S2). In SMA, one single representation for
554 laryngeal activity during vocalization was observed, while in cerebellum, two distinct
555 representations were found.
556
557
558 *Syllable production task - Surface results*
559
560 Group count maps of the syllable production task projected to the surface showed that several
561 vertices in the mid-portion of the central sulcus were commonly activated during articulation
562 in 19 out of 20 subjects (Figure 2B). The group count map for vocalization showed separate
563 dorsal and ventral regions robustly activated across the group. In the main contrast for
564 vocalization, there was greater variability in the exact location of activity above threshold and
565 the areas activated in individuals were smaller than for the main contrast for articulation. This
566 resulted in less overlap of activated vertices for the vocalization group maps. For some
567 subjects, vocalization-induced activity did not reach significance in relevant brain regions,
568 which also resulted in lower values in the count map.
569

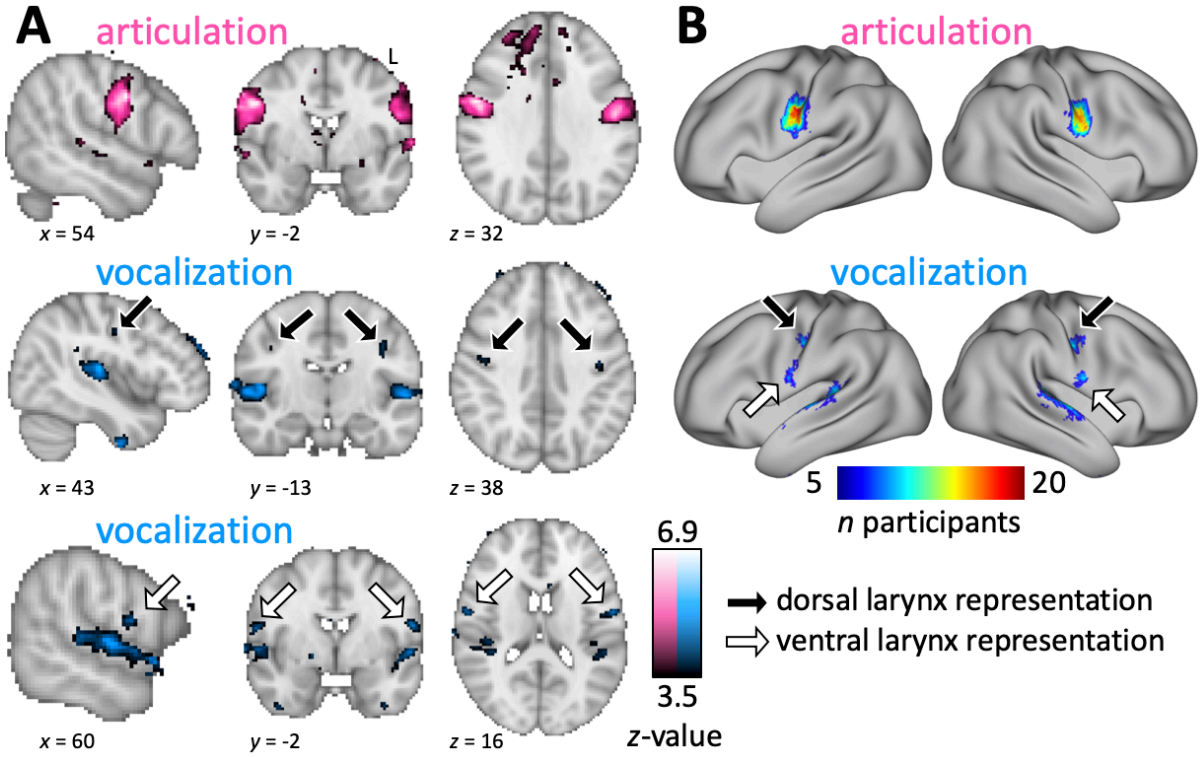


Figure 2 – Syllable production task. **A**: Whole-brain group activation maps showing areas activated during articulation and vocalization (voxel-wise threshold $z > 3.5$, $n = 20$). For the main contrast for vocalization (blue), both a dorsal and a ventral representation are shown in separate panels. Black arrow: dorsal larynx representation; white arrow: ventral larynx representation. **B**: Surface group count maps of the same contrasts. Individual surface maps were thresholded at $p < 0.05$ (corrected voxel-wise), binarized and resampled onto the fs_LR surface. These were summed across the group of 20 subjects and thresholded at $n > 4$. Note: in both analyses, vocalization also resulted in activation of the superior temporal gyrus, presumably reflecting auditory stimulation due to hearing oneself vocalizing.

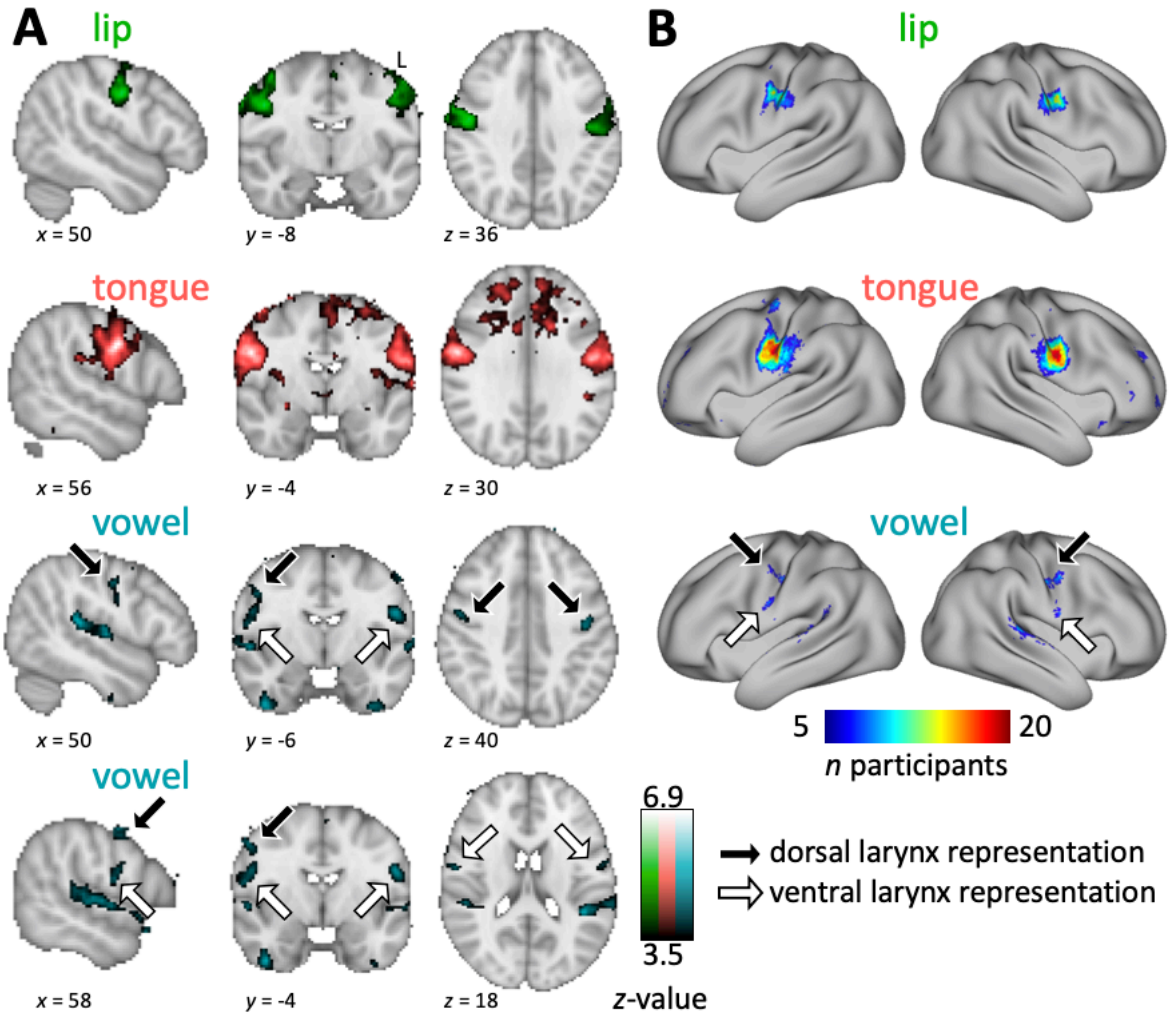
Basic localizer task (task 2)

Basic localizer task - Volumetric fMRI results

586 Brain activation during movement of the tongue and the lips and for laryngeal activity during
587 vowel production was assessed by contrasting basic speech movements with a 'breathing only'
588 condition, which was matched for controlled breathing. We found distinct activation peaks
589 that followed the predicted somatotopic organization in the mid-portion of the central sulcus
590 with the tongue representation more ventral than the lip representation in both hemispheres
591 (Figure 3A) (Penfield and Rasmussen 1950; Grabski et al. 2012; Carey et al. 2017). The location
592 of the activity during tongue movement overlapped with the result of the main contrast for
593 articulation in the syllable production task; as previously, we found spurious activity in the
594 prefrontal lobe white matter that we presume is movement related.
595
596 Vowel production induced activity bilaterally in a central dorsal region, a ventral region anterior
597 to the central sulcus, and in superior temporal cortex. Within the dorsal region, activation is
598 found both in the central sulcus and on the precentral gyrus. As for the syllable production
599 task, activity in superior temporal cortex was presumed due to hearing oneself. In the group
600 average map, and at the statistical threshold used, the dorsal and ventral larynx
601 representations appeared to be connected, with residual activity along the ventral central
602 sulcus.
603
604
605 *Basic localizer task - Surface results*
606
607 Group count maps of the basic localizer task projected to the surface revealed the same
608 somatotopic activity pattern as seen in the volumetric results for group average activity (Figure
609 3B). The tongue condition showed highly consistent activation in the mid-portion of the central

610 sulcus during articulation (maximum overlap was achieved for all 20 subjects). The result for
611 tongue movement was highly similar to the result of the main contrast for articulation, which
612 is in line with the volumetric results. Complete overlap of activated vertices was not achieved
613 for the lip condition, but the activated region was still highly consistent across subjects
614 (maximum: 15 subjects).
615
616 For the vowel production condition, the group count maps show a dorsal and a ventral cluster,
617 similar to the pattern seen during the main contrast for vocalization (task 1). The dorsal cluster
618 extended from the central sulcus to the precentral gyrus and in the right hemisphere a distinct
619 dorsal gyral and a dorsal sulcal activation can be observed. We presumed that this dorsal gyral
620 activity, but not the sulcal activity, is a residual breathing-related effect, because it overlaps
621 with activity from the 'breathing only' condition (Figure S1). The ventral activation cluster
622 appeared to extend into a part of cortex where we expected the tongue representation, and
623 may reflect residual tongue activity during vowel production. In contrast to the volumetric
624 results for the group averaged activity, the count maps based on individually thresholded
625 activation maps show the dorsal and ventral representation activated by vowel production to
626 be clearly separate.
627
628 As described above for the main contrast for vocalization, the values in the group count maps
629 for vowel production indicate that the location of this activity is less consistent than for the
630 other conditions; the area activated during vocalization is also smaller, both dorsally and
631 ventrally, than for the articulators. The highest overlap was achieved in the right dorsal larynx
632 representation in 11 subjects.

633



634

635 Figure 3 – Basic localizer task. **A**: Whole-brain group activation maps showing areas activated during lip and
636 tongue movement and during vowel production (vowel-wise threshold $z > 3.5$, $n = 20$). For the vowel
637 production condition (turquoise), both a dorsal and a ventral representation are shown in separate panels.
638 See legend to Figure 2 for details. **B**: Surface group count maps of the same conditions. See legend to Figure
639 2 for details.

640

641

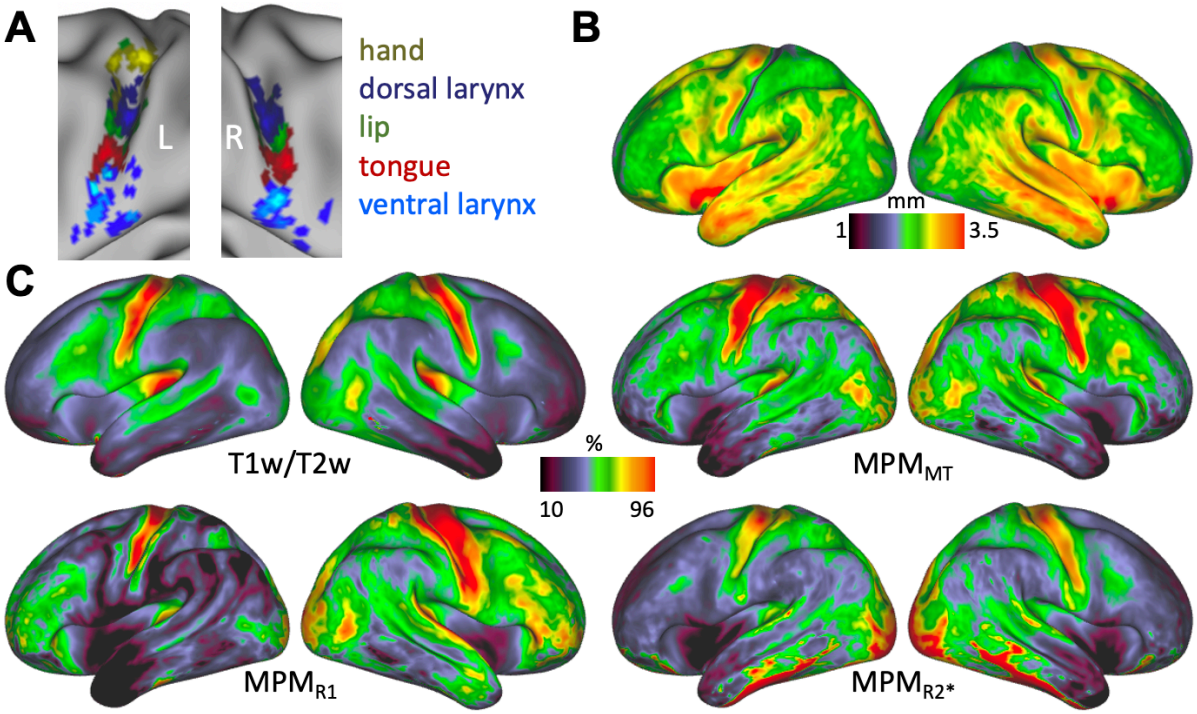
642 *Individual surface activation maxima*

643

644 In order to characterize the variability of brain activity across subjects, we derived individual
645 activation peaks for the hand movement (based on the hand-localizer, only in the left
646 hemisphere), lip and tongue movement (based on the basic localizer task) and vocalization
647 (based on the syllable production task). The main contrast for vocalization was used to localize
648 the dorsal and a ventral larynx representation. Figure 4A shows the spatial distribution of
649 individual activation peaks on an inflated brain surface. Overall, the location of the peaks for
650 the different movement types is highly consistent with similar cross-subject variability. For the
651 ventral larynx representation, particularly in the left hemisphere, however, the location of the
652 maxima appears to be more variable.
653
654 Within-subject reliability of the activation maxima across the two tasks was compared for
655 'vocalization' and 'vowel production' for both dorsal and ventral larynx representations as well
656 as for 'articulation' and 'tongue movement'. Reliability was high with a median distance across
657 subjects of less than 10 mm for the three ROIs in both hemispheres.
658
659
660 *Cortical surface maps of microstructural features*
661
662 We derived whole-brain average surface maps for different measures related to cortical
663 microstructure (Figure 4B, C). Overall, the T1w/T2w map and the three MPM parameter maps
664 (MPM_{MT}, MPM_{R1}, MPM_{R2*}) show a similar pattern of myelin content across the cortex (Figure
665 4C). In all four maps, the central sulcus as well as precentral and postcentral gyrus are
666 characteristically high in myelin, which is considered to be a defining feature of primary motor
667 and sensory cortex (Glasser and van Essen 2011). The location of the ventral boundary of the

668 somatomotor cortex, indicated by a steep gradient of myelin values, slightly varies across the
669 four maps, but this boundary is consistently more ventrally located in the right compared with
670 the left hemisphere.
671
672 The four maps that correlate with cortical myelin (T1w/T2w, MPM_{MT}, MPM_{R1} and MPM_{R2*}) are
673 sensitive to different biophysical properties of the myelin, but it should be noted that their
674 sensitivity profiles are not completely independent (Callaghan et al. 2016). Therefore, some
675 dissimilarities regarding the distribution of myelin along the cortex can be observed across the
676 maps. MPM_{MT} and MPM_{R1} have a stronger signal in the frontal lobe compared to the T1w/T2w
677 and the MPM_{R2*} map. In addition to myelin, the R2* signal is influenced by cortical properties
678 such as iron content and calcium (Wu et al. 2009; Bagnato et al. 2018). For the T1w/Tw2 map,
679 the underlying biophysical model is less well understood. The high R2* values in the ventral
680 temporal lobe are likely a susceptibility artifact caused by signal loss at the air-tissue boundary.
681
682 In the cortical thickness map, a prominent strip of low values (i.e. thinner cortex) can be
683 observed at the posterior bank of the central sulcus, indicating the location of primary sensory
684 cortex (Fischl and Dale 2000) (Figure 4B). Cortical thickness values are high (i.e. thicker cortex)
685 in the anterior bank of the central sulcus and in the precentral gyrus, which is indicating the
686 location of primary motor cortex. In the cortical thickness map, the ventral boundary of
687 primary sensory cortex can be determined by a sharp gradient at the level of the subcentral
688 gyrus. The location of this boundary is comparable with the location described above in the
689 other maps. The same hemispheric difference can be observed with the somatomotor
690 boundary being located further ventrally in the right hemisphere.
691

692 The whole brain maps demonstrate that these cortical surface measures are informative about
693 the cortical microstructure underlying our functional activation maxima, which we describe in
694 the next section. The result of the reconstructed MPM parameter maps *per se* is not a primary
695 result of the paper.
696



697

698 Figure 4 – **A**: Individual activation maxima derived for hand movement (only left hemisphere), lip movement,
699 tongue movement and vocalization (i.e. larynx activity during voice production) (*n* = 20). For vocalization,
700 activation maxima in two separate brain regions were derived to localize the dorsal and ventral larynx
701 representations. **B**: Average cortical thickness map (*n* = 20). **C**: Average T1w/T2w myelin map (*n* = 20) and
702 MPM parameter maps (*n* = 19).
703
704
705 *Microstructural properties at activation maxima*
706

Next, we determined the intensity value of the microstructural surface maps at the individual subjects' activation maxima for the different movement types (Figure 5). Figure 5A shows z-transformed intensity values for T1w/T2w, MPM_{MT}, MPM_{R1} and MPM_{R2*} maps at hand, lip, tongue as well as dorsal and ventral larynx maxima, i.e. at the peaks presented in Figure 4A. Mixed effects analyses with the factors ROI (excluding the hand ROI), hemisphere and map demonstrated that there were highly significant effects of ROI ($F(3, 600) = 60.69$, $p < 0.001$) and hemisphere ($F(1, 600) = 43.54$, $p < 0.001$) but no significant effect of map ($F(3, 600) < 1$, *n.s.*). The significant effect of ROI reflects that values for the ventral larynx are lower than in the other ROIs (post-hoc pairwise *t*-tests with Tukey adjustment, $p < 0.001$), but the values for dorsal larynx representation, lip and tongue were not significantly different. The main effect of hemisphere reflects higher values in the right hemisphere ($p < 0.001$). The interaction between ROI and hemisphere was significant ($F(3, 600) = 3.93$, $p = 0.009$) indicating that the difference between the ventral larynx representation and the other motor representations was greater on the right hemisphere than on the left.

Cortical thickness values show a high inter-individual variability for hand, dorsal larynx, tongue and lip ROIs (Figure 5B). This effect can be attributed to the fact that some maxima are located in the depth of the central sulcus, where values are low, while others are located on the anterior bank of the central sulcus, where values are high. Variability in the right hemisphere is lower, which is consistent with the location of the maxima predominantly on the anterior bank (Figure 4A). This result potentially indicates that some of derived maximal activation peaks represent activity evoked by somatosensation rather than by motor control. If maxima with a cortical thickness level below a cut-off value of 2 mm (Fischl and Dale 2000) are excluded, the comparison of cortical thickness values shows a marginally nonsignificant effect

731 of ROI ($F(3, 120) = 2.53$, $p < 0.061$), but not of hemisphere ($F(1, 120) < 1$). The marginal effect
732 of ROI is driven by the ventral larynx representation having lower cortical thickness values
733 compared to the other ROIs ($p < 0.068$).
734
735 To increase sensitivity in the quantification of myelin values described above, we re-ran the
736 linear mixed model while excluding the sensory activation maxima as defined by a cortical
737 thickness below 2 mm. The significant main effects of ROI ($F(3, 480) = 73.62$, $p < 0.001$) and
738 hemisphere ($F(1, 480) = 20.81$, $p < 0.001$) were stronger than reported above. As reported
739 above, the main effect of ROI is due to decreased values in the ventral larynx representation
740 and the hemispheric effect due to increased values in the right hemisphere. The interaction
741 effect of ROI and hemisphere ($F(3, 480) = 2.26$, $p = 0.08$)) was no longer significant, indicating
742 that the difference between ventral larynx representation and the other motor
743 representations did not differ between hemispheres.
744
745 To assess effects based on the non-transformed values of the four maps (T1w/T2w, MPM_{MT},
746 MPM_{R1}, MPM_{R2*}) rather than the *z*-scores, we also derived a 'dissimilarity matrix' based on the
747 pair-wise Manhattan distances between ROIs after excluding the sensory activation maxima
748 (Figure 5C). The values were averaged across hemispheres, because no significant interaction
749 effect was found. The dissimilarity matrix reflects the effect of the ventral larynx
750 representation being most dissimilar from the other motor representations based on these
751 quantitative measures of cortical myelin.

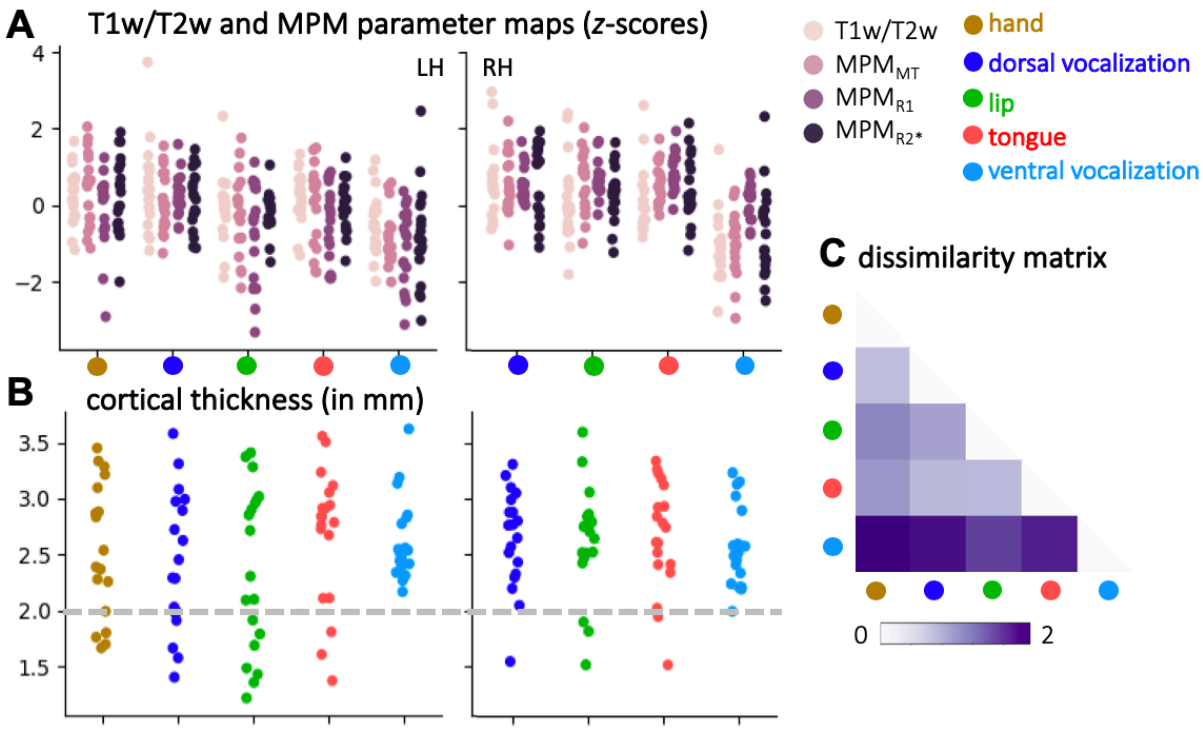


Figure 5 – Measures of cortical microstructure derived at individual activation maxima (n = 19). **A**: Microstructural values that correlate with myelin (T1w/T2w, MPM_{MT}, MPM_{R1}, MPM_{R2*}) derived at activation maxima (for individual maxima see Figure 4A). **B**: Individual values of cortical thickness. The dashed grey line indicates the threshold that was chosen to determine the exclusion of activation maxima that are presumed to be sensory due to their location in cortex < 2 mm thick. **C**: Dissimilarity matrix of the myelin values across ROIs. The matrix shows a quantitative comparison of the values shown in **A** after excluding sensory activation maxima. Pair-wise dissimilarity is defined as the Manhattan distance between each pair of ROIs.

A linear discriminant analysis (LDA) revealed that the following equation can be used to discriminate the ventral larynx representation from the dorsal larynx representation:

$$y = -0.28 \cdot T1w/T2w - 2.82 \cdot MPM_{MT} - 3.03 \cdot MPM_{R1} - 0.47 \cdot MPM_{R2*} - 1.19$$

When inputting the *z*-transformed values of the surface maps at a specific location, a resulting value of $y > 0$ indicates that the profile of values is more similar to the ventral larynx representation than to the dorsal larynx representation. The formula indicates that MPM_{MT} and MPM_{R1} are the most informative measures to discriminate the two larynx representations (largest weighting). The mean performance accuracy of the classifier is 0.94 (while 0.5 indicates chance level performance).

Taken together, these quantifications show that the ventral larynx representation is located in cortex that has lower myelin content and lower cortical thickness compared to primary motor cortex, where the other movement representations including the dorsal larynx representation, are located.

4. Discussion

The goal of this study was to characterize the cortical microstructure underlying the laryngeal representations in the human brain. We localized brain activity evoked by voluntary control of laryngeal movements during vocalization using a novel paradigm. We showed that even when controlling for breathing, vocalization elicits brain activity in two separate parts of the somatomotor cortex: a dorsal region in the central sulcus and a ventral region close to the Sylvian fissure. On an individual level, the laryngeal activations and the activations during movement of hand, lips, and tongue show a consistent somatotopic arrangement. Characterization of cortical microstructure based on structural and quantitative MRI shows that the dorsal larynx representation has a similar profile to other movement representations

primary motor cortex, while the ventral larynx representation has lower myelin content and cortical thickness. These results suggest that the dorsal larynx representation is the primary locus of laryngeal motor control in primary motor cortex, while the ventral larynx representation relates more to secondary activations in non-primary motor cortex.

Our experimental task was designed to separate laryngeal activity from that evoked by supralaryngeal articulation and breathing. We were able to dissociate the supralaryngeal and laryngeal components of syllable production using a factorial design and orthogonal task contrasts (Murphy et al. 1997). Covert speech is known to engage similar somatomotor brain areas to overt speech production (Kleber et al. 2007). For our purpose, however, using a 'covert speech' condition was essential to allow us to construct a fully factorial model aimed at isolating activity related to the execution of movements involved in articulation and vocalization while controlling for other processes such as selection and planning of movement sequences. The result of a factorial design is statistically more robust than a subtraction design, because the main contrasts include data from all task conditions improving the mathematical estimates and associated statistics. The main contrast for vocalization showed some residual activity of the tongue representation. Using the neutral vowel schwa or glottal stops instead (Gick 2002; Loucks et al. 2007; Brown et al. 2008; Grabski et al. 2012; Belyk et al. 2018), however, might have caused pharynx activity that would have been more difficult to dissociate.

In addition to the syllable production task, we used a basic localizer task that involved movement of the lips, and tongue, and vowel production. Results of the basic localizer task were consistent with the syllable production task, and within-subject variability across the comparable contrasts was low. The areas activated for the main contrast of articulation overlap

with the result from the tongue localizer, which is expected given that the syllables produced mainly rely on tongue movement. The areas activated for the main contrast for vocalization overlap with the result from the vowel production condition, but some differences could be observed. Residual tongue activity resulting from producing the vowel /i/ is more apparent in the vowel production condition, suggesting that this was better controlled for in the factorial design.

When studying vocalizations, control for breathing is essential given that human vocal speech sounds are mostly produced during exhalation. Several previous neuroimaging studies, however, did not control for breathing. Some previous studies used an instructed exhalation condition for comparison with vocalization (Loucks et al. 2007; Galgano et al. 2019), but a difference in activation of regions in motor cortex was not found consistently. One likely explanation is that explicitly instructing subjects to exhale might engage laryngeal muscles in such a way that no difference to laryngeal activity during vocalization can be observed. In this study, the breathing pattern and rate were matched for all task conditions. Inhalation and exhalation were explicitly instructed on a screen and monitored using a breath belt. Although slight deviations in the shape of the breathing trace can be observed, the overall breathing traces were comparable across subjects and conditions. The deviations in breathing patterns observed in the basic localizer task suggest that breathing is less well controlled for compared with breathing during the syllable production task, where subjects produced the same utterance in different conditions.

Our results suggest that controversial findings within the neuroimaging literature can largely be explained by differences in experimental design. Several studies focused on the dorsal

region as location of the 'laryngeal motor cortex' or 'larynx phonation area' (Sörös et al. 2006; Brown et al. 2008; Kleber et al. 2013; Belyk and Brown 2014; Belyk et al. 2018). Brown et al. (2008) suggested the presence of a 'dorsolateral' larynx representation on the crown of the precentral gyrus (x = 50, y = -2, z = 37) and a 'ventromedial' larynx representation in the central sulcus (x = 44, y = -8, z = 34). Our results suggest, however, that activity in the gyral portion of the dorsal region is a residual breathing-related signal and not specific for laryngeal activity during vocalization. When contrasting the 'breathing only' condition, during which breathing was explicitly instructed, to the resting baseline with normal breathing, we found activity in a dorsal gyral region (x = 55, y = 0, z = 43), which is presumably related to voluntary control of the diaphragm during instructed inhalation and exhalation (Figure S1B) (Ramsay et al. 1993; Olthoff et al. 2008). Similarly, we found the dorsal gyral activation when contrasting the 'covert speech' condition, which involved instructed breathing, but with no overt movement, to the resting baseline with normal breathing (Figure S1A). We also find this dorsal gyral activity during vowel production, but only when we assess it relative to the resting baseline with normal breathing (Figure S1B). When we assess vowel production relative to the 'breathing only' activity, only the sulcal portion of the dorsal activity, which we also found in the main contrast for vocalization, remains (x = 43, y = -13, z = 38). The activation that Brown et al. (2008) refer to as 'ventromedial' larynx area, overlaps most closely with the region that we focus in this paper as dorsal (sulcal) larynx representation.

In addition to activation of the dorsal larynx representation, multiple neuroimaging studies reported activation in a region that we refer to as ventral larynx representation (Terumitsu et al. 2006; Olthoff et al. 2008; Grabski et al. 2012). Strong evidence for the involvement of a ventral larynx representation in laryngeal motor control also comes from electrical recording

863 from the cortical surface during vocalization and microsimulation studies (Galgano and Froud
864 2008; Bouchard et al. 2013; Chang et al. 2013; Breshears et al. 2015; Dichter et al. 2018).
865 Electrocorticography demonstrated the presence of both a dorsal and a ventral larynx
866 representation much more consistently than the fMRI literature (reviewed in Conant et al.,
867 2014). Direct electrical stimulation of the dorsal larynx representation causes vertical larynx
868 movement that correlates with stimulation magnitude and evokes vocalization (Dichter et al.
869 2018); stimulation of the ventral larynx representation causes speech arrest (Chang et al.
870 2017).
871 In addition to our novel experimental design, this study focuses on investigating individual
872 differences in laryngeal activation patterns. The task count maps demonstrate that laryngeal
873 activity is much less consistent across subjects compared with other movement types. In some
874 subjects the dorsal or ventral activity did not reach the individual significance threshold. The
875 location of the ventral larynx representation, in particular in the left hemisphere, show a much
876 wider spatial spread across the cortex indicating large intra-individual variability. A recent
877 anatomical study based on the dataset presented here, showed that inter-individual
878 morphological variability in the ventral end of the central sulcus is higher than in the rest of
879 the central sulcus (Eichert et al. 2020). It was shown that variability in sulcal morphology can
880 explain, in part, spatial variability in functional activation peaks. As a result of the variability
881 and the low *z*-statistical values, group level fMRI analysis in previous studies might have failed
882 to detect one of the larynx representations or the task paradigm was not optimized to localize
883 the larynx representations.
884
885 The existence of two separate brain regions correlating with laryngeal activity in the central
886 sulcus raises the question as to whether both of them are involved in motor control of the

887 larynx. Given that the larynx area in the macaque is located in a ventral premotor area (Hast
888 et al. 1974; Jürgens 1974; Simonyan and Jürgens 2002; Coudé et al. 2011), an evolutionary
889 'migration and duplication' hypothesis has been proposed (Belyk and Brown, 2017; Jarvis,
890 2019; reviewed in Mars et al., 2018). The theory states that the human ventral larynx
891 representation migrated posteriorly during evolution, possibly into a different
892 cytoarchitectonic area. The dorsal larynx representation is thought to have evolved as human
893 novelty in primary motor cortex.
894
895 Additional support for a duplication theory also comes from genetic profiling analyses
896 comparing gene expressions in song-learning birds and humans (Pfenning et al. 2014;
897 Chakraborty and Jarvis 2015). Gene expression in the avian vocal nuclei, that are involved in
898 vocal learning, are similar to the expression profiles in both the dorsal and the ventral larynx
899 representation of the human brain. In the context of this more general brain pathway
900 duplication theory, it has been recently suggested that there was an additional duplication in
901 human vocal premotor cortex, leading to the emergence of a pre-larynx motor cortex (preLMC)
902 just anterior to the ventral larynx representation (Jarvis 2019). The proposed genetic
903 mechanisms for the evolutionary brain pathway duplication theory in laryngeal motor control,
904 however, remain to be verified experimentally.
905
906 An alternative interpretation of the non-human primate literature would suggest that the
907 species differences are quantitative rather than qualitative. While traditional electrical
908 mapping studies failed to find a focal larynx representation in primary motor cortex in non-
909 human primates (Bailey et al. 1950; Hast et al. 1974; Simonyan and Jürgens 2002), neuronal
910 activity correlating with laryngeal movement was observed, for example, during swallowing in

macaques (Martin et al. 1997, 1999). Furthermore, tract-tracing studies showed that the premotor larynx representation in the macaque has direct connections to primary motor cortex (Künzle 1978; Simonyan and Jürgens 2002). These findings might indicate the presence of a non-human homology of the dorsal larynx representation, which is experimentally more challenging to localize due to its smaller spatial extent.

The present study is the first to characterize the microstructural properties of cortex underlying the dorsal and ventral larynx representation. The use of high-resolution quantitative neuroimaging allows us to characterize cortical architecture noninvasively (Weiskopf et al. 2013; Lutti et al. 2014). Measures from T1w/T2w maps and MPM parameter maps (MPM_{MT}, MPM_{R1}, MPM_{R2*}) are sensitive to different microstructural properties of the cortex, but all of them have been shown to correlate with myelin to varying degrees (Glasser and van Essen 2011; Weiskopf et al. 2013; Callaghan et al. 2014, 2015; Bagnato et al. 2018). Age-related effects can confound measures of cortical microstructure (Gennatas et al. 2017; Grydeland et al. 2019), but the effects in the regions studied are expected to be minimal.

The cytoarchitectonic area underlying the larynx representations, remains to be determined histologically. Our quantification, however, suggests strongly that the dorsal larynx representation is located in primary motor cortex, which is characterized by high myelin content and thicker cortex (Fischl and Dale 2000; Glasser and van Essen 2011). The ventral larynx representation has lower myelin content and thinner cortex indicating that it is located in a different cortical territory. Based on a cytoarchitectonic map of the human brain (Brodmann 1905), we suggest that the human ventral larynx representation is located in Brodmann area (BA) 6 (premotor cortex). This interpretation is consistent with the idea that

935 the primate brain contains multiple laryngeal representations in different cortical areas: The
936 human ventral larynx representation is homologous to the premotor macaque larynx
937 representation (Hast et al. 1974; Jürgens 1974; Simonyan and Jürgens 2003; Coudé et al. 2011)
938 and the human dorsal larynx representation is homologous to a macaque larynx
939 representation in primary motor cortex, which receives projections from the premotor larynx
940 representation (Künzle 1978; Simonyan and Jürgens 2002).
941
942 Many individual maxima for the ventral larynx representation were located in the subcentral
943 part of cortex. Brodmann considered this a postcentral and therefore somatosensory cortical
944 region (BA 43), subjacent and anterior to primary somatosensory cortex (BA 3, 1 and 2)
945 (Brodmann 1905). A magnetoencephalography study showed that air-puff stimulation of the
946 larynx evokes activity in a ventral region (Miyaji et al. 2014). Using fMRI, BA 43 was also found
947 to be activated by movements of the tympanic membrane associated with changes in pressure
948 in the oropharyngeal cavity such as those that occur during gustation and swallowing and in
949 vocalization (Job et al. 2011). These results suggest that the ventral larynx activity could be a
950 sensory rather than a motor representation.
951
952 The question regarding the distinct functional contributions of the two larynx representations
953 during voice productions remains unanswered. Based on our results, we formulate a
954 hypothesis regarding the causal role of the two laryngeal representations during voice
955 production: We propose that, as for vocalizations in non-human primates, the ventral larynx
956 representation is involved in basic control of the vocal folds so that a vocal sound can be
957 produced. Rapid adduction (tensioning) and abduction (relaxation) of the vocal folds at the
958 onset and offset of a vocalization is primarily modulated by the intrinsic laryngeal muscles

(Jürgens 1974; Simonyan and Jürgens 2003). Fine motor control, which is required for pitch modulations in human speech and singing, however, also relies on the vertical movement of the larynx within the trachea, which is mediated by the extrinsic muscles. We suggest that pitch control is facilitated by the dorsal larynx representation, which is located in primary motor cortex (Kleber et al. 2013; Dichter et al. 2018; Finkel et al. 2019). Activity in both intrinsic and extrinsic laryngeal muscles, however, are tightly coupled and might not be controlled by distinct brain areas (Belyk and Brown 2014). Our hypothesis is in line with an evolutionary theory suggesting that the dorsal larynx representation is unique to the human brain (Belyk and Brown 2017). Such a functional dissociation of dorsal and ventral larynx representation during vocalization, however, still needs to be tested directly using, for example, non-invasive brain stimulation.

In sum, we used neuroimaging to localize neural activity related to laryngeal motor control during vocalization, while controlling for confounding factors such as breathing and supralaryngeal articulation. We found two activated regions for laryngeal activity during vocalization, which are in anatomically distinct brain areas. Quantification of cortical microstructure suggests that the dorsal representation, but not the ventral representation, is located in primary motor cortex. It remains open, whether and how these two representations differentially contribute to laryngeal motor control.

Acknowledgements

N.E. is a Wellcome Trust Doctoral student in Neuroscience at the University of Oxford [203730/Z/16/Z]. The project was supported by the NIHR Oxford Health Biomedical Research Centre. The Wellcome Centre for Integrative Neuroimaging is supported by core funding from the Wellcome Trust [203139/Z/16/Z]. The work of R.B.M. is supported by the Biotechnology and Biological Sciences Research Council (BBSRC) UK [BB/N019814/1] and the Netherlands Organization for Scientific Research NWO [452-13-015]. The Authors would like to thank Martina Callaghan from University College London for the MPM sequence.

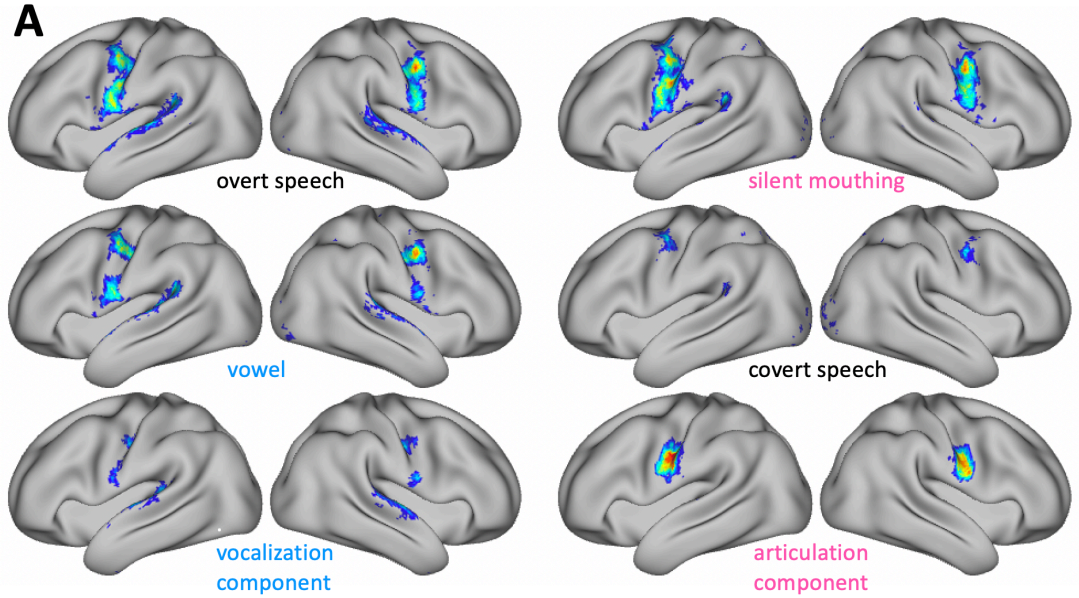
Code Availability

Upon acceptance of the manuscript, all processing code will be made available from the Wellcome Centre for Integrative Neuroimaging's GitLab at git.fmrib.ox.ac.uk/neichert/project_larynx. FSL tools, are available from fsl.fmrib.ox.ac.uk. Connectome Workbench is available at www.humanconnectome.org/software/connectome-workbench.html.

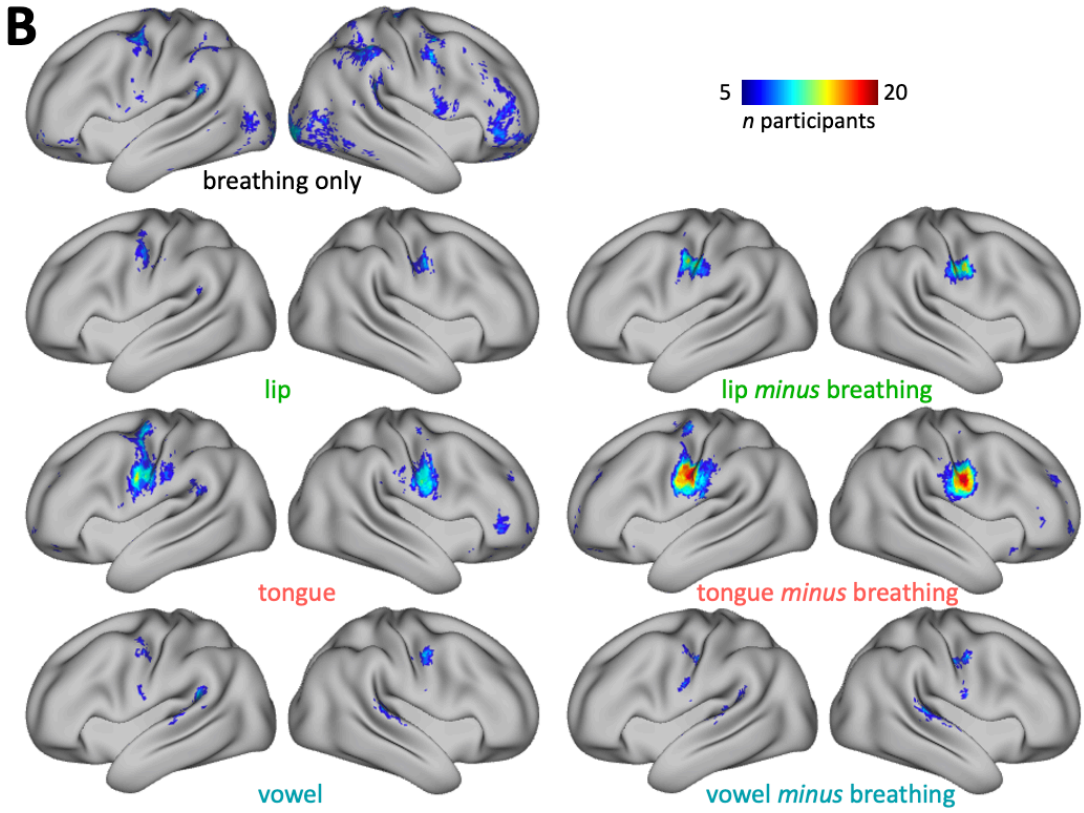
997 Supplementary Material

998

999 *Surface count maps of all individual task contrasts*



1000



1001

1002

Figure S1 – Related to Figure 2 and Figure 3. Surface count maps of all individual task contrasts. **A**: Syllable production task. Upper two rows: Four task conditions - overt speech, silent mouthing (articulation only), vowel production (vocalization only), covert speech (thinking) when compared to the resting baseline with normal breathing. Third row: Two orthogonal main contrasts for the articulation and vocalization component (these are also shown in Figure 2B). **B**: Basic localizer task. Left panel: Task conditions compared to the resting baseline with normal breathing. The 'breathing only' condition is a contrast between the breathing condition (where breathing was instructed) and the resting baseline with normal breathing. A lower threshold of $z = 1.96$ was used to threshold individual z-statistical images for this contrast due to generally lower activation. Right panel: Movement conditions compared to the breathing condition with explicit breathing instruction (these are also shown in Figure 3B).

Task activations in supplementary motor area

The two main contrasts for articulation and vocalization - at a lower voxel-wise threshold of $z > 2$ - revealed activity bilaterally in SMA (Figure S2). A somatotopic arrangement was observed with vocalization activating a more anterior and more dorsal part of cortex. The representations of the effectors during the basic localizer task also show a clear somatotopy: In dorsal-to-ventral direction we first find the representation of the larynx, which is also most anterior, then of the lip and then of the tongue. Only one single representation of the larynx was found in SMA in both the main contrast for vocalization and in the vowel condition.

The location of the representation of the speech effectors is in line with previous accounts in the literature (Picard and Strick 1996). It has been suggested that the vertical line crossing the anterior commissure, the VAC, is a landmark for a division of SMA proper and pre-SMA (Picard

and Strick 1996; Rizzolatti et al. 1996). Our results thus suggest that lip and tongue representations are located just posterior to VAC in an anterior region of SMA proper. Laryngeal activity during vocalization, however, activates a part of cortex anterior to VAC, presumably in preSMA.

Task activations in cerebellum

Movement of the articulators and laryngeal activity also evoke activation of the cerebellum in a somatotopic fashion, which is mirroring the order observed for motor cortex (Figure S2). Most ventrally, a representation of the larynx is observed, which activates during vowel production and vocalization. This is followed dorsally by a representation of the lips and the tongue and then by a second representation of the larynx.

According to a probabilistic atlas of the human cerebellum (Diedrichsen et al. 2009), all the activations observed are located in the anterior cerebellar lobule VI. This finding is consistent with previous neuroimaging studies that found activation of the anterior-superior aspect of cerebellum during speech movements (Petersen et al. 1989; Fiez and Raichle 1997). A previous resting-state functional connectivity study demonstrated that cerebellar representations mirror the topography in cerebral motor cortex (Buckner et al. 2011). It is therefore presumed that the ventral-most representation of the larynx in cerebellum is related to the dorsal larynx representation in motor cortex and vice versa.

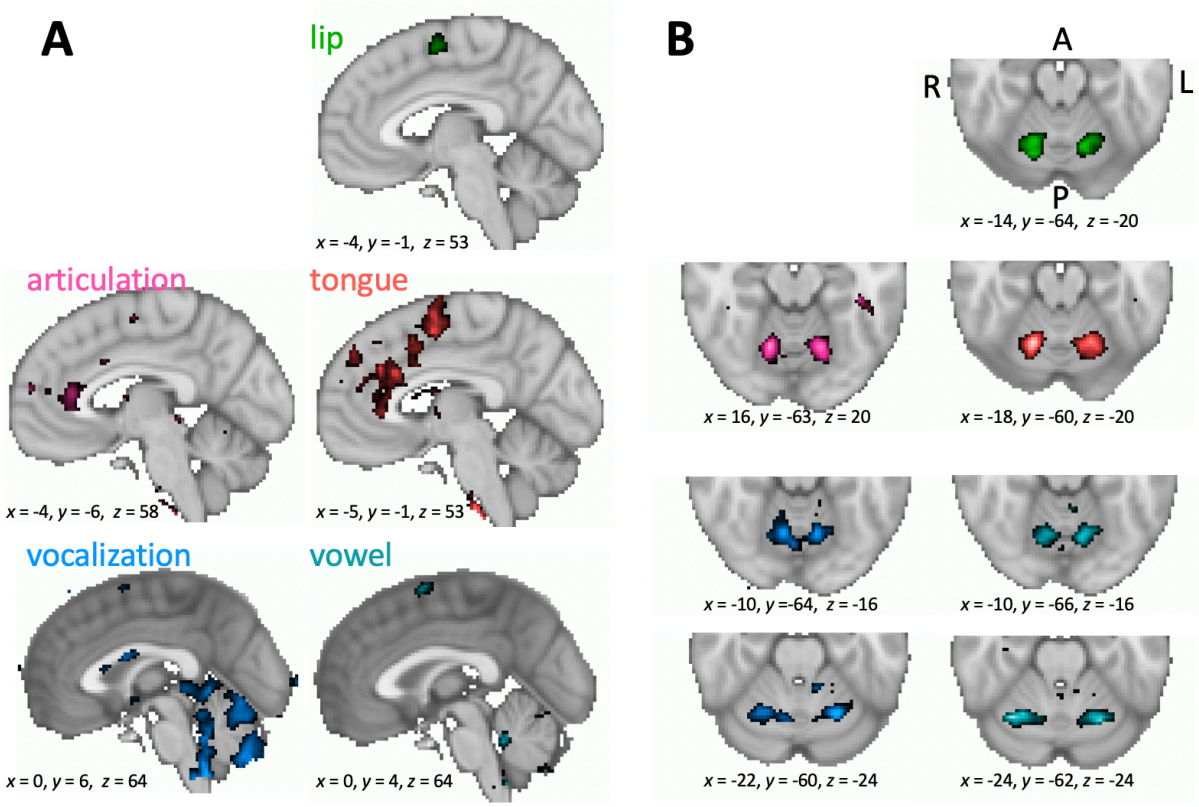


Figure S2 – Related to Figure 2 and Figure 3. Whole-brain group activation maps showing areas activated during syllable production task (left panels) and basic localizer task (right panels) (vowel-wise threshold $z > 3.5$, $n = 20$). Coordinates are given for the voxel of maximal activation in the left hemisphere. **A**: Activity on the medial brain surface centered on the voxel of maximal activity in the left supplementary motor area (SMA) (note that for the main contrast for vocalization the threshold was lowered to $z > 2$). **B**: Cerebellar activity in the same task contrasts and shown in the same colors as in **A**. For the main contrast for vocalization and for the vowel production contrast, two separate activations are shown in different slices.

Regions-of-interest to derive maxima during task activation

An example of the volumetric ROIs in an individual is shown in Figure S3. The central sulcus ROI used for the hand, lip and tongue was defined using FreeSurfer's automatic volumetric labelling based on the Destrieux Atlas.

For the larynx, we identified two activation maxima in separate ROIs: One for the dorsal and one for the ventral larynx representation. The dorsal larynx ROI was a portion of the same central sulcus ROI used above from *z*-coordinates in MNI space of 50 - 30. The limits were determined empirically, so that ROI did not capture the ventral larynx representation or an unrelated supra-dorsal activation in the trunk area, which was observed in some individuals (Foerster 1931).

The ventral larynx representation lay outside the central sulcus and was located ventrally in the subcentral part of cortex. Due to the high intra-individual morphological variability in this region (Eichert et al. 2020), the ventral larynx ROI was derived manually based on individual anatomy in surface space. A liberal surface ROI was drawn on each individual's midthickness surface covering the ventral part of the central sulcus and adjacent gyri (Figure S3A). Anteriorly, the ROI was delineated by the inferior portion of the precentral sulcus and posteriorly the ROI spanned the postcentral gyrus. If present, the lateral portion of the ascending sulcus in the subcentral gyrus was included within the ROI. The dorsal limit of the ROI was defined by a horizontal plane across the gyrus at the level of the usual location of the posterior ramus of the inferior precentral sulcus. The ventral larynx surface ROI was converted into a volumetric ROI covering the underlying cortical ribbon using wb_command. We checked that the ventral larynx ROI did not overlap with subjacent auditory cortex in the temporal lobe or inferior frontal cortex.

In some subjects, the main contrast for vocalization in the syllable production task had additional activity related to articulation of the tongue. To remove this, we transformed the coordinates for each individual's maximal voxel from the tongue contrast (from the basic localizer task) to the functional space of the syllable production task (task 1) using rigid-body transformation and then derived a spherical ROI (7 voxels diameter) around it. This sphere was used to mask the *z*-statistic image of the main contrast for vocalization prior to localizing the maxima for laryngeal activity in the dorsal and ventral ROIs described above.

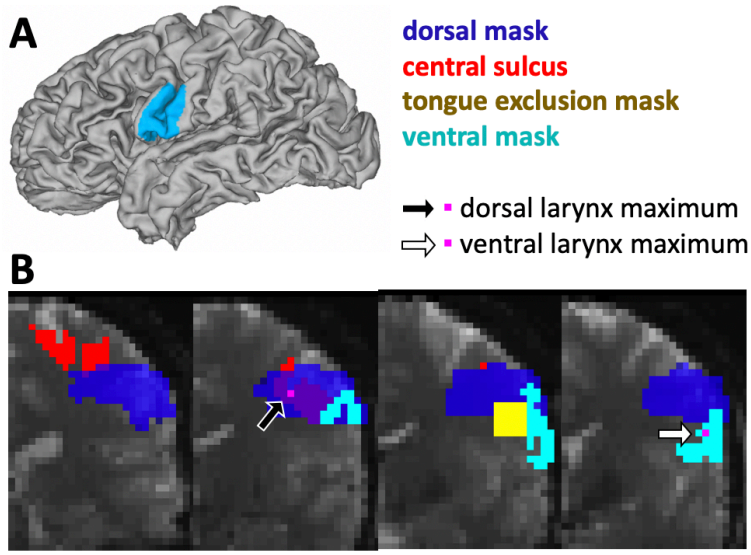


Figure S3 – Related to Methods. **A**: Example of a manually-drawn surface ROI for the ventral larynx representation. **B**: Volumetric ROIs overlaid onto an individual's functional scan. The central sulcus ROI (red) was used to derive the voxel of maximal activation for the hand, lip and tongue. The intersection of the central sulcus ROI and a dorsal mask (dark blue) was used to derive the voxel of maximal activation for the dorsal larynx representation (black arrow, pink voxel). The ventral mask was projected from surface to volume space (bright blue) to derive the voxel of maximal activation in the ventral larynx representation (white arrow, pink voxel). A spherical ROI around the voxel of maximal activation in the tongue contrast was masked out from the dorsal and ventral mask (yellow).

Group-level task activation maxima

Table S1: Group-level task activation maxima. Reported are MNI coordinates of maxima in somatomotor cortex for the main contrasts reported in the manuscript (LH, RH: left and right hemisphere). For the main contrast for vocalization and vowel production, a dorsal and a ventral maxima are reported separately.

task	main contrast	LH			RH		
		x	*y*	*z*	*x*	*y*	*z*
hand localizer	hand	-38	-21	57			
syllable production	articulation	-50	-12	32	54	-2	32
	vocalization (dorsal)	-40	-16	36	43	-13	38
	vocalization (ventral)	-58	-2	18	60	-2	16
basic localizer	lip	-48	-10	40	50	-8	36
	tongue	-52	-11	28	56	-4	30
	vowel (dorsal)	-42	-16	40	50	-6	40
	vowel (ventral)	-58	-2	20	58	-4	18

1113 References

1114

1115 Ackermann H, Hage S, Ziegler W. 2014. Brain mechanisms of acoustic communication in
1116 humans and nonhuman primates: An evolutionary perspective. Behavioral and Brain
1117 Sciences. 37:529–604.

1118 Andersson JLR, Jenkinson M, Smith SM. 2007. Non-linear optimisation FMRIB technial eport
1119 TR07JA1.

1120 Bagnato F, Hametner S, Boyd E, Endmayr V, Shi Y, Ikonomidou V, Chen G, Pawate S,
1121 Lassmann H, Smith S, Brian Welch E. 2018. Untangling the R2* contrast in multiple
1122 sclerosis: A combined MRI-histology study at 7.0 Tesla. PLoS ONE. 13:e0193839.

1123 Bailey P, Bonin G von, McCulloch WS. 1950. The Isocortex of the Chimpanzee. 440.

1124 Barch DM, Burgess GC, Harms MP, Petersen SE, Schlaggar BL, Corbetta M, Glasser MF, Curtiss
1125 S, Dixit S, Feldt C, Nolan D, Bryant E, Hartley T, Footer O, Bjork JM, Poldrack R, Smith S,
1126 Johansen-Berg H, Snyder AZ, Van Essen DC. 2013. Function in the human connectome:
1127 Task-fMRI and individual differences in behavior. NeuroImage. 80:169–189.

1128 Bates DM, Sarkar D. 2007. Bates, Douglas M., and Deepayan Sarkar. "lme4: linear mixed-
1129 effects models using S4 classes. R package version 099875-6.

1130 Belyk M, Brown S. 2014. Somatotopy of the extrinsic laryngeal muscles in the human
1131 sensorimotor cortex. Behavioural Brain Research. 270:364–371.

1132 Belyk M, Brown S. 2017. The origins of the vocal brain in humans. Neuroscience and
1133 Biobehavioral Reviews. 77:177–193.

1134 Belyk M, Lee YS, Brown S. 2018. How does human motor cortex regulate vocal pitch in
1135 singers ? Royal Society Open Science. 5:172208.

1136 Bennett CM, Miller MB. 2010. How reliable are the results from functional magnetic

1137 resonance imaging? Annals of the New York Academy of Sciences.

1138 Bohland JW, Guenther FH. 2006. An fMRI investigation of syllable sequence production.

1139 NeuroImage. 32:821–841.

1140 Bouchard KE, Mesgarani N, Johnson K, Chang EF. 2013. Functional organization of human

1141 sensorimotor cortex for speech articulation. Nature. 495:327–332.

1142 Breshears JD, Molinaro AM, Chang EF. 2015. A probabilistic map of the human ventral

1143 sensorimotor cortex using electrical stimulation. Journal of Neurosurgery. 123:340–349.

1144 Brodmann K. 1905. Beitrage zur histologischen Lokalisation der Grosshirnrinde: III.

1145 Mitteilung. Die Rindenfelder der niederen Affen. J Psychol Neurol. 4:177–206.

1146 Brown S, Laird AR, Pfordresher PQ, Thelen SM, Turkeltaub P, Liotti M. 2009. The somatotopy

1147 of speech: Phonation and articulation in the human motor cortex. Brain and Cognition.

1148 70:31–41.

1149 Brown S, Ngan E, Liotti M. 2008. A larynx area in the human motor cortex. Cerebral Cortex.

1150 18:837–845.

1151 Buckner RL, Krienen FM, Castellanos A, Diaz JC, Yeo BTT, Thomas Yeo BT. 2011. The

1152 organization of the human cerebellum estimated by intrinsic functional connectivity.

1153 Journal of Neurophysiology. 106:2322–2345.

1154 Callaghan MF, Freund P, Draganski B, Anderson E, Cappelletti M, Chowdhury R, Diedrichsen J,

1155 FitzGerald THB, Smittenaar P, Helms G, Lutti A, Weiskopf N. 2014. Widespread age-

1156 related differences in the human brain microstructure revealed by quantitative

1157 magnetic resonance imaging. Neurobiology of Aging. 35:1862–1872.

1158 Callaghan MF, Helms G, Lutti A, Mohammadi S, Weiskopf N. 2015. A general linear

1159 relaxometry model of R1 using imaging data. Magnetic Resonance in Medicine.

1160 73:1309–1314.

1161 Callaghan MF, Mohammadi S, Weiskopf N. 2016. Synthetic quantitative MRI through
1162 relaxometry modelling. NMR in Biomedicine. 29:1729–1738.
1163 Carey D, Krishnan S, Callaghan MF, Sereno MI, Dick F. 2017. Functional and quantitative MRI
1164 mapping of somatomotor representations of human supralaryngeal vocal tract. Cerebral
1165 Cortex. 27:265–278.
1166 Casey BJ, Cannonier T, Conley MI, Cohen AO, Barch DM, Heitzeg MM, Soules ME, Teslovich T,
1167 Dellarco D V., Garavan H, Orr CA, Wager TD, Banich MT, Speer NK, Sutherland MT,
1168 Riedel MC, Dick AS, Bjork JM, Thomas KM, Chaarani B, Mejia MH, Hagler DJ, Daniela
1169 Cornejo M, Sicat CS, Harms MP, Dosenbach NUF, Rosenberg M, Earl E, Bartsch H, Watts
1170 R, Polimeni JR, Kuperman JM, Fair DA, Dale AM. 2018. The Adolescent Brain Cognitive
1171 Development (ABCD) study: Imaging acquisition across 21 sites. Developmental
1172 Cognitive Neuroscience. 32:43–54.
1173 Chakraborty M, Jarvis ED. 2015. Brain evolution by brain pathway duplication. Philosophical
1174 transactions of the Royal Society of London Series B, Biological sciences. 370:20150056-.
1175 Chang EF, Breshears JD, Raygor KP, Lau D, Molinaro AM, Berger MS. 2017. Stereotactic
1176 probability and variability of speech arrest and anomia sites during stimulation mapping
1177 of the language dominant hemisphere. Journal of Neurosurgery. 126:114–121.
1178 Chang EF, Niziolek CA, Knight RT, Nagarajan SS, Houde JF. 2013. Human cortical sensorimotor
1179 network underlying feedback control of vocal pitch. Proceedings of the National
1180 Academy of Sciences. 110:2653–2658.
1181 Conant D, Bouchard KE, Chang EF. 2014. Speech map in the human ventral sensory-motor
1182 cortex. Current Opinion in Neurobiology. 24:63–67.
1183 Coudé G, Ferrari PF, Rodà F, Maranesi M, Borelli E, Veroni V, Monti F, Rozzi S, Fogassi L. 2011.
1184 Neurons controlling voluntary vocalization in the macaque ventral premotor cortex.

1185 PLoS ONE. 6:1–10.

1186 Devlin JT, Matthews PM, Rushworth MFS. 2003. Semantic Processing in the Left Inferior
1187 Prefrontal Cortex: A Combined Functional Magnetic Resonance Imaging and
1188 Transcranial Magnetic Stimulation Study. Journal of Cognitive Neuroscience. 15:71–84.

1189 Dichter BK, Breshears JD, Leonard MK, Chang EF. 2018. The Control of Vocal Pitch in Human
1190 Laryngeal Motor Cortex. Cell. 174:21-31.e9.

1191 Diedrichsen J, Balsters JH, Flavell J, Cussans E, Ramnani N. 2009. A probabilistic MR atlas of
1192 the human cerebellum. NeuroImage. 46:39–46.

1193 Draganski B, Ashburner J, Hutton C, Kherif F, Frackowiak RSJ, Helms G, Weiskopf N. 2011.
1194 Regional specificity of MRI contrast parameter changes in normal ageing revealed by
1195 voxel-based quantification (VBQ). NeuroImage. 55:1423–1434.

1196 Eichert N, Watkins KE, Mars RB, Petrides M. 2020. Morphological and functional variability in
1197 central and subcentral motor cortex of the human brain. bioRxiv. 2020.03.17.995035.

1198 Evans KC, Shea SA, Saykin AJ. 1999. Functional MRI localisation of central nervous system
1199 regions associated with volitional inspiration in humans. Journal of Physiology. 520:383–
1200 392.

1201 Fiez JA, Raichle ME. 1997. Linguistic Processing. In: International Review of Neurobiology. p.
1202 233–254.

1203 Finkel S, Veit R, Lotze M, Friberg A, Vuust P, Soekadar S, Birbaumer N, Kleber B. 2019.
1204 Intermittent theta burst stimulation over right somatosensory larynx cortex enhances
1205 vocal pitch-regulation in nonsingers. Human Brain Mapping. 1–14.

1206 Fischl B. 2012. FreeSurfer. NeuroImage.

1207 Fischl B, Dale AM. 2000. Measuring the thickness of the human cerebral cortex from
1208 magnetic resonance images. Proceedings of the National Academy of Sciences of the

1209 United States of America. 97:11050–11055.

1210 Fitch WT. 2017. Empirical approaches to the study of language evolution. Psychonomic

1211 Bulletin and Review. 24:3–33.

1212 Foerster O. 1931. THE CEREBRAL CORTEX IN MAN. Lancet. 309–312.

1213 Foerster O. 1936. Symptomatologie der Erkrankungen des Grosshirns. Motorische Felder und

1214 Bahnen. Sensible corticale Felder, Handbuch der Neurologie 6. Berlin: Springer.

1215 Galgano J, Froud K. 2008. Evidence of the voice-related cortical potential: An

1216 electroencephalographic study. NeuroImage. 41:1313–1323.

1217 Galgano J, Pantazatos S, Allen K, Yanagihara T, Hirsch J. 2019. Functional connectivity of PAG

1218 with core limbic system and laryngeal cortico-motor structures during human

1219 phonation. Brain Research. 1707:184–189.

1220 Gennatas ED, Avants BB, Wolf DH, Satterthwaite TD, Ruparel K, Ciric R, Hakonarson H, Gur

1221 RE, Gur RC. 2017. Age-related effects and sex differences in gray matter density,

1222 volume, mass, and cortical thickness from childhood to young adulthood. Journal of

1223 Neuroscience. 37:5065–5073.

1224 Gick B. 2002. An X-ray investigation of pharyngeal constriction in American English Schwa.

1225 Phonetica. 59:38–48.

1226 Glasser MF, Goyal MS, Preuss TM, Raichle ME, Van Essen DC. 2014. Trends and properties of

1227 human cerebral cortex: Correlations with cortical myelin content. NeuroImage.

1228 Glasser MF, Sotiropoulos SN, Wilson JA, Coalson TS, Fischl B, Andersson JL, Xu J, Jbabdi S,

1229 Webster M, Polimeni JR, Van Essen DC, Jenkinson M. 2013. The minimal preprocessing

1230 pipelines for the Human Connectome Project. NeuroImage. 80:105–124.

1231 Glasser MF, van Essen DC. 2011. Mapping human cortical areas in vivo based on myelin

1232 content as revealed by T1- and T2-weighted MRI. Journal of Neuroscience. 31:11597–

11616.

Grabski K, Lamalle L, Vilain C, Schwartz JL, Vallée N, Tropres I, Baciu M, Le Bas JF, Sato M. 2012. Functional MRI assessment of orofacial articulators: Neural correlates of lip, jaw, larynx, and tongue movements. Human Brain Mapping. 33:2306–2321.

Grydeland H, Vértes PE, Váša F, Romero-Garcia R, Whitaker K, Alexander-Bloch AF, Bjørnerud A, Patel AX, Sederevičius D, Tamnes CK, Westlye LT, White SR, Walhovd KB, Fjell AM, Bullmore ET. 2019. Waves of maturation and senescence in micro-structural MRI markers of human cortical myelination over the lifespan. Cerebral Cortex. 29:1369–1381.

Hast MH, Fischer JM, Wetzel AB, Thompson VE. 1974. Cortical motor representation of the laryngeal muscles in Macaca mulatta. Brain Research. 73:229–240.

Jarvis ED. 2019. Evolution of vocal learning and spoken language. Science. 366:50–54.

Jenkinson M. 2003. Fast, automated, N-dimensional phase-unwrapping algorithm. Magnetic Resonance in Medicine. 49:193–197.

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

Jenkinson M, Beckmann CF, Behrens TEJ, Woolrich MW, Smith SM. 2012. FSL. NeuroImage. 62:782–790.

Jenkinson M, Smith S. 2001. A global optimisation method for robust affine registration of brain images. Medical Image Analysis. 5:143–156.

Job A, Paucod JC, O'Beirne GA, Delon-Martin C. 2011. Cortical representation of tympanic membrane movements due to pressure variation: An fMRI study. Human Brain Mapping. 32:744–749.

1257 Jürgens U. 1974. On the elicitability of vocalization from the cortical larynx area. Brain
1258 Research. 81:564–566.
1259 Jürgens U. 2002. Neural pathways underlying vocal control. Neuroscience and Biobehavioral
1260 Reviews.
1261 Kleber B, Birbaumer N, Veit R, Trevorrow T, Lotze M. 2007. Overt and imagined singing of an
1262 Italian aria. NeuroImage. 36:889–900.
1263 Kleber B, Zeitouni AG, Friberg A, Zatorre RJ. 2013. Experience-Dependent Modulation of
1264 Feedback Integration during Singing: Role of the Right Anterior Insula. Journal of
1265 Neuroscience. 33:6070–6080.
1266 Kuehn E, Dinse J, Jakobsen E, Long X, Schäfer A, Bazin PL, Villringer A, Sereno MI, Margulies
1267 DS. 2017. Body Topography Parcellates Human Sensory and Motor Cortex. Cerebral
1268 Cortex. 27:3790–3805.
1269 Kumar V, Croxson PL, Simonyan K. 2016. Structural Organization of the Laryngeal Motor
1270 Cortical Network and Its Implication for Evolution of Speech Production. Journal of
1271 Neuroscience. 36:4170–4181.
1272 Künzle H. 1978. Cortico-cortical efferents of primary motor and somatosensory regions of the
1273 cerebral cortex in Macaca fascicularis. Neuroscience. 3:25–39.
1274 Leyton ASF, Sherrington CS. 1917. Observations on the Excitable Cortex of the Chimpanzee,
1275 Orang-Utan, and Gorilla. Quarterly Journal of Experimental Physiology. 11:135–222.
1276 Loucks TMJ, Poletto CJ, Simonyan K, Reynolds CL, Ludlow CL. 2007. Human brain activation
1277 during phonation and exhalation: Common volitional control for two upper airway
1278 functions. NeuroImage. 36:131–143.
1279 Ludlow CL. 2005. Central nervous system control of the laryngeal muscles in humans.
1280 Respiratory Physiology and Neurobiology.

1281 Lutti A, Dick F, Sereno MI, Weiskopf N. 2014. Using high-resolution quantitative mapping of
1282 R1 as an index of cortical myelination. NeuroImage. 93:176–188.
1283 Lutti A, Hutton C, Finsterbusch J, Helms G, Weiskopf N. 2010. Optimization and validation of
1284 methods for mapping of the radiofrequency transmit field at 3T. Magnetic Resonance in
1285 Medicine. 64:229–238.
1286 Lutti A, Stadler J, Josephs O, Windischberger C, Speck O, Bernarding J, Hutton C, Weiskopf N.
1287 2012. Robust and fast whole brain mapping of the RF transmit field B1 at 7T. PLoS ONE.
1288 7:1–7.
1289 Mars RB, Eichert N, Jbabdi S, Verhagen L, Rushworth MFS. 2018. Connectivity and the search
1290 for specializations in the language-capable brain. Current Opinion in Behavioral
1291 Sciences. 21:19–26.
1292 Martin RE, Kemppainen P, Masuda Y, Yao D, Murray GM, Sessle BJ. 1999. Features of
1293 cortically evoked swallowing in the awake primate (Macaca fascicularis). Journal of
1294 Neurophysiology. 82:1529–1541.
1295 Martin RE, Murray GM, Kemppainen P, Masuda Y, Sessle BJ. 1997. Functional properties of
1296 neurons in the primate tongue primary motor cortex during swallowing. Journal of
1297 Neurophysiology. 78:1516–1530.
1298 McKay LC, Evans KC, Frackowiak RSJ, Corfield DR. 2003. Neural correlates of voluntary
1299 breathing in humans. Journal of Applied Physiology. 95:1170–1178.
1300 Miyaji H, Hironaga N, Umezaki T, Hagiwara K, Shigeto H, Sawatsubashi M, Tobimatsu S,
1301 Komune S. 2014. Neuromagnetic detection of the laryngeal area: Sensory-evoked fields
1302 to air-puff stimulation. NeuroImage. 88:162–169.
1303 Murphy K, Corfield DR, Guz A, Fink GR, Wise RJ, Harrison J, Adams L. 1997. Cerebral areas
1304 associated with motor control of speech in humans. Journal of applied physiology

(Bethesda, Md : 1985). 83:1438–1447.

Nieto-Castañón A, Fedorenko E. 2012. Subject-specific functional localizers increase sensitivity and functional resolution of multi-subject analyses. NeuroImage. 63:1646–1669.

Olthoff A, Baudewig J, Kruse E, Dechent P. 2008. Cortical Sensorimotor Control in Vocalization: A Functional Magnetic Resonance Imaging Study. The Laryngoscope. 118:2091–2096.

Penfield W, Boldrey E. 1937. Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. Brain. 60:389–443.

Penfield W, Rasmussen T. 1950. The Cerebral Cortex of Man: Clinical Study of Localization of Function. Academic Medicine. 25:375.

Petersen SE, Fox PT, Posner MI, Mintun M, Raichle ME. 1989. Positron emission tomographic studies of the processing of single words. Journal of Cognitive Neuroscience. 1:153–170.

Pfenning AR, Hara E, Whitney O, Rivas M V, Wang R, Roulhac PL, Howard JT, Wirthlin M, Lovell P V, Ganapathy G, Mountcastle J, Moseley MA, Thompson JW, Soderblom EJ, Iriki A, Kato M, Gilbert MTP, Zhang G, Bakken T, Bongaarts A, Bernard A, Lein E, Mello C V, Hartemink AJ, Jarvis ED. 2014. Convergent transcriptional specializations in the brains of humans and song-learning birds. Science. 346:1256846–1256846.

Picard N, Strick PL. 1996. Motor areas of the medial wall: A review of their location and functional activation, Cerebral Cortex. Narnia.

Preibisch C, Deichmann R. 2009. Influence of RF spoiling on the stability and accuracy of T1 mapping based on spoiled FLASH with varying flip angles. Magnetic Resonance in Medicine. 61:125–135.

Ramsay SC, Adams L, Murphy K, Corfield DR, Grootoonk S, Bailey DL, Frackowiak RS, Guz A.

1329 1993. Regional cerebral blood flow during volitional expiration in man: A comparison
1330 with volitional inspiration. The Journal of Physiology. 461:85–101.
1331 Rizzolatti G, Luppino G, Matelli M. 1996. The classic supplementary motor area is formed by
1332 two independent areas. Advances in neurology. 70:45–56.
1333 Simonyan K. 2014. The laryngeal motor cortex: Its organization and connectivity. Current
1334 Opinion in Neurobiology. 28:15–21.
1335 Simonyan K, Jürgens U. 2002. Cortico-cortical projections of the motorcortical larynx area in
1336 the rhesus monkey. Brain Research. 949:23–31.
1337 Simonyan K, Jürgens U. 2003. Efferent subcortical projections of the laryngeal motorcortex in
1338 the rhesus monkey. Brain Research. 974:43–59.
1339 Simonyan K, Ostuni J, Ludlow CL, Horwitz B. 2009. Functional But Not Structural Networks of
1340 the Human Laryngeal Motor Cortex Show Left Hemispheric Lateralization during Syllable
1341 But Not Breathing Production. Journal of Neuroscience. 29:14912–14923.
1342 Smith SM. 2002. Fast robust automated brain extraction. Human Brain Mapping. 17:143–
1343 155.
1344 Sörös P, Sokoloff LG, Bose A, McIntosh AR, Graham SJ, Stuss DT. 2006. Clustered functional
1345 MRI of overt speech production. NeuroImage. 32:376–387.
1346 Tabelow K, Balteau E, Ashburner J, Callaghan MF, Draganski B, Helms G, Kherif F, Leutritz T,
1347 Lutti A, Phillips C, Reimer E, Ruthotto L, Seif M, Weiskopf N, Ziegler G, Mohammadi S.
1348 2019. hMRI – A toolbox for quantitative MRI in neuroscience and clinical research.
1349 NeuroImage. 194:191–210.
1350 Terumitsu M, Fujii Y, Suzuki K, Kwee IL, Nakada T. 2006. Human primary motor cortex shows
1351 hemispheric specialization for speech. Neuroreport. 17:1091–1095.
1352 Toyoda G, Brown EC, Matsuzaki N, Kojima K, Nishida M, Asano E. 2014. Electrocorticographic

correlates of overt articulation of 44 English phonemes: Intracranial recording in children with focal epilepsy. Clinical Neurophysiology. 125:1129–1137.

Weiskopf N, Callaghan MF, Josephs O, Lutti A, Mohammadi S. 2014. Estimating the apparent transverse relaxation time (R2*) from images with different contrasts (ESTATICS) reduces motion artifacts. Frontiers in Neuroscience. 8.

Weiskopf N, Suckling J, Williams G, Correia M. MM, Inkster B, Tait R, Ooi C, Bullmore T. ET, Lutti A. 2013. Quantitative multi-parameter mapping of R1, PD*, MT, and R2* at 3T: A multi-center validation. Frontiers in Neuroscience. 7:95.

Woo CW, Krishnan A, Wager TD. 2014. Cluster-extent based thresholding in fMRI analyses: Pitfalls and recommendations. NeuroImage. 91:412–419.

Woolrich MW, Behrens TEJ, Beckmann CF, Jenkinson M, Smith SM. 2004. Multilevel linear modelling for FMRI group analysis using Bayesian inference. NeuroImage. 21:1732–1747.

Woolrich MW, Ripley BD, Brady M, Smith SM. 2001. Temporal autocorrelation in univariate linear modeling of FMRI data. NeuroImage. 14:1370–1386.

Wu Z, Mittal S, Kish K, Yu Y, Hu J, Haacke EM. 2009. Identification of calcification with MRI using susceptibility-weighted imaging: A case study. Journal of Magnetic Resonance Imaging. 29:177–182.