

Contents lists available at [ScienceDirect](https://www.sciencedirect.com)

NeuroImage

journal homepage: [www.elsevier.com/locate/neuroimage](https://www.elsevier.com/locate/neuroimage)

## Human larynx motor cortices coordinate respiration for vocal-motor control <sup>☆</sup>

Michel Belyk <sup>a</sup>, Rachel Brown <sup>b,c</sup>, Deryk S. Beal <sup>d,e</sup>, Alard Roebroek <sup>b</sup>, Carolyn McGettigan <sup>a</sup>, Stella Guldner <sup>f</sup>, Sonja A. Kotz <sup>b,g,\*</sup>

<sup>a</sup> Department of Speech Hearing and Phonetic Sciences, University College London, London, United Kingdom

<sup>b</sup> Faculty of Psychology and Neuroscience, University of Maastricht, Maastricht, The Netherlands

<sup>c</sup> RWTH Aachen University, Institute for Psychology, Aachen Germany

<sup>d</sup> Bloorview Research Institute, Holland Bloorview Kids Rehabilitation Hospital, Toronto, Ontario, Canada

<sup>e</sup> Department of Speech-Language Pathology, University of Toronto, Toronto, Ontario, Canada

<sup>f</sup> Department of Cognitive and Clinical Neuroscience, Central Institute of Mental Health, Medical Faculty Mannheim, Heidelberg University, Mannheim, Germany

<sup>g</sup> Department of Neuropsychology, Max Planck Institute for Human and Cognitive Sciences, Leipzig, Germany

### ARTICLE INFO

#### Keywords:

Voice  
Respiration  
Larynx  
Motor cortex  
fMRI  
Song  
Whistle

### ABSTRACT

Vocal flexibility is a hallmark of the human species, most particularly the capacity to speak and sing. This ability is supported in part by the evolution of a direct neural pathway linking the motor cortex to the brainstem nucleus that controls the larynx the primary sound source for communication. Early brain imaging studies demonstrated that larynx motor cortex at the dorsal end of the orofacial division of motor cortex (dLMC) integrated laryngeal and respiratory control, thereby coordinating two major muscular systems that are necessary for vocalization. Neurosurgical studies have since demonstrated the existence of a second larynx motor area at the ventral extent of the orofacial motor division (vLMC) of motor cortex. The vLMC has been presumed to be less relevant to speech motor control, but its functional role remains unknown. We employed a novel ultra-high field (7T) magnetic resonance imaging paradigm that combined singing and whistling simple melodies to localise the larynx motor cortices and test their involvement in respiratory motor control. Surprisingly, whistling activated both 'larynx areas' more strongly than singing despite the reduced involvement of the larynx during whistling. We provide further evidence for the existence of two larynx motor areas in the human brain, and the first evidence that laryngeal-respiratory integration is a shared property of both larynx motor areas. We outline explicit predictions about the descending motor pathways that give these cortical areas access to both the laryngeal and respiratory systems and discuss the implications for the evolution of speech.

### 1. Introduction

Diverse and flexible vocal communication is a hallmark of the human species, most notably in the ability to speak and sing. These human behaviours are supported by the human capacity to flexibly add novel vocal patterns to their repertoire usually by learning through imitation (Janik and Slater, 2000). Few species of mammals have strong Vocal Production Learning (VPL) abilities and none of these are closely related to humans. Monkeys are particularly weak vocal learners (Fischer and Hammerschmidt, 2019), while non-human apes appear to have intermediate VPL abilities (Lameira et al., 2016; Wich et al., 2012) as well as some facility in controlling the respiratory drive for sound production (Lameira et al., 2013; Perlman and Clark, 2015; Wich et al., 2009).

These abilities are supported by specialisations in the motor system that have been considerably altered over course of primate evolution.

The human brain is peculiar in having two Larynx Motor Areas (LMCs) per hemisphere to control a single laryngeal organ, where only one larynx area would be expected (Belyk and Brown, 2017). The first of these LMCs to be described is located in primary motor cortex, which contains a somatotopic map of the body's muscles in the brain (Lotze et al., 2000; Penfield and Boldrey, 1937; Rao et al., 1995; Stippich et al., 2002). Early brain imaging studies identified a human-specific larynx-controlling region at the dorsal extent of the orofacial somatotopy in humans (Brown et al., 2009, 2008, 2004; Loucks et al., 2007; Simonyan et al., 2009, 2007). The dorsal Larynx Motor Cortex (dLMC) is located deep within the central sulcus in primary motor cortex and is adjacent to articulatory motor regions for the lips and

<sup>☆</sup> Preprint DOI: <https://psyarxiv.com/pc4uh/>

\* Corresponding author.

E-mail address: [sonja.kotz@maastrichtuniversity.nl](mailto:sonja.kotz@maastrichtuniversity.nl) (S.A. Kotz).

tongue (Loucks et al. 2007; Simonyan et al. 2007; Brown et al. 2008; Peck et al. 2009; Grabski et al. 2013; Belyk et al., 2018b).

The ventral Larynx Motor Cortex (vLMC) has been identified by more recent studies including neurosurgical recording (Bouchard et al., 2013; Chang et al., 2013; Dichter et al., 2018) and brain imaging (Eichert et al. 2020a; Kleber et al. 2013; Belyk et al. 2016), though earlier studies reported activation that was consistent with this region (Olthoff et al., 2008; Terumitsu et al., 2006). These studies confirmed the localization of the dLMC, near the articulatory muscles, and also observed a second, larynx-controlling region at the ventral extent of the orofacial motor cortex near the representation of the throat and swallowing (Breshears et al., 2015; Penfield and Boldrey, 1937).

The mechanisms by which these two brain areas share control over the voice, and what separate contributions they may make to voice motor control, are critical to our understanding of how humans evolved to speak. The dLMC is known to integrate laryngeal and respiratory motor control (Loucks et al., 2007; Simonyan et al., 2007), suggesting that it may have a role in coordinating these muscular systems both of which are required for vocal sound production. The dual cortical representation of the laryngeal muscles had not yet been described at the time of these studies, and it is not known whether this laryngeal-respiratory linkage is shared by the vLMC.

We hypothesize that either both larynx motor areas integrate laryngeal and respiratory function, or laryngeal-motor integration is restricted to the dLMC. Joint control of the laryngeal and respiratory musculature is likely to reduce conduction latencies as well as the metabolic cost of coordination between the muscle groups that drive the primary sound source for communication (Hallermann et al., 2012; Ju et al., 2016), which may have driven both regions to develop this feature. Alternatively, the restriction of laryngeal-respiratory integration to the dLMC would suggest a simpler adaptation in support of speech motor control.

We conducted an ultra-high field functional Magnetic Resonance Imaging (fMRI) study of voice motor control as compared to whistling (Fig. 1). Vocal imitation of wordless melodies is an effective localizer of the dorsal and ventral larynx motor areas (Belyk et al., 2016). Whistling imitation matches the experimental demands of vocal imitation (Belyk et al., 2018a) and engages a different subset of the speech relevant musculature. Whereas singing engages the laryngeal and respiratory muscles, whistling engages the articulatory and respiratory muscles (Belyk et al., 2019). Comparing the neural correlates of this combination of behaviours is a novel probe of a neural system that provides the vocal dexterity required for speech and furthers our understanding of the unique features that make the human brain speech capable.

## 2. Methods

### 2.1. Participants

Thirteen participants (9 female), with a mean age of 26.4 (SD 5.2) years participated in the study after giving their informed consent. All participants were without neurological or psychiatric illness. Twelve participants were right-handed, and one was ambidextrous. Participants had diverse linguistic backgrounds and included native speakers of English (4), German (4), Dutch (3), Spanish (1), or Catalan (1) though all were fluent in English. Participants were recruited at Maastricht University. The study was approved by the Ethical Review Committee at the Faculty of Psychology and Neuroscience at Maastricht University (ECP-161\_01\_02\_2016).

A second sample of twenty-four native speakers of British English (21 female) with mean age 21.0 (SD 3.3) was analysed to replicate resting state analyses. Volunteers were recruited from the participant pool at the Department of Psychology at Royal Holloway, University of London. This study was approved by the research ethics committee of Royal Holloway, University of London (587-2017-10-24-14-50-UXJT010).

### 2.2. Procedure

Participants performed experimental tasks in two runs of functional MRI lasting 720 s each. In the first run participants imitated simple melodies by whistling or singing. The second run was an experiment on simple speech movements not reported here. Runs followed a sparse, event-related paradigm that allowed participants to perform the experimental task without interference due to auditory noise produced by the MRI scanner during data acquisition. Imitations began 5–7 s prior to data acquisition and were followed by a 6–8 s gap to allow the BOLD (blood oxygen level dependant) response to return to baseline. Compliance to task instructions was verified by audio recordings taken during the experimental session using an MRI compatible microphone.

Participants listened to 48 simple melodies with the instruction to imitate them as accurately as possible (Fig. 1). Half of the stimuli were presented in a vocal timbre matched to the gender of the participant to be imitated by singing, and the remaining half were presented in a whistled timbre to be imitated with a bilabial whistle. Singing was performed without words as a hum with the lips gently parted to isolate vocalisation. Participants practiced singing and whistling without head movement outside the scanner on a separate set of auditory stimuli. The auditory stimuli lasted 4 s and were followed by a 5–7 s silent period during which participants imitated the stimulus.

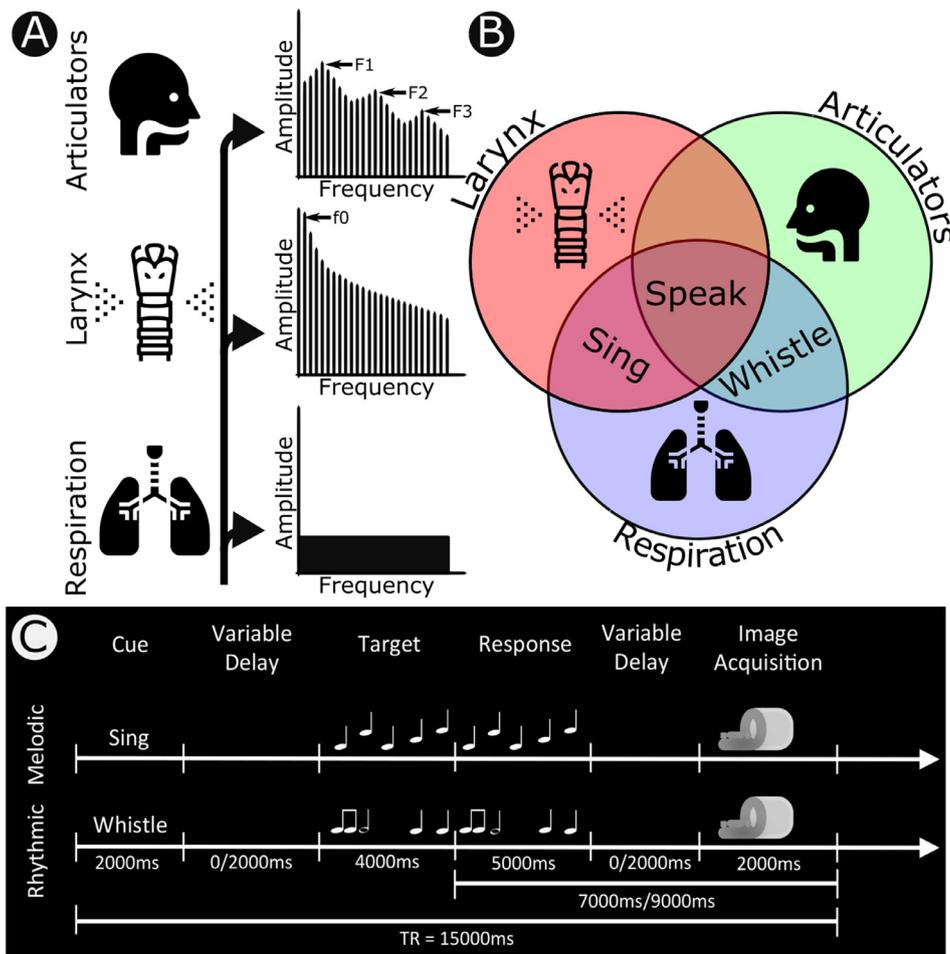
Half of the stimuli had an isochronous temporal pattern, but varied in pitch (i.e., melodies). The other half of the stimuli had a fixed pitch, but varied in temporal pattern (i.e., rhythms). Each combination of movement task (singing or whistling) and stimulus type (melodies or rhythms) was presented in two blocks of six trials for a total of 12 trials per condition. A visual cue preceded each trial indicating whether the stimulus should be sung whistled and was replaced by a fixation cross for the remainder of the trial. Blocks were presented in counterbalanced order and separated by one or two silent rest trials to a total of 12. Separate melodic and rhythmic stimuli were presented to address a separate set of hypotheses not addressed here.

Each stimulus consisted of 5 notes lasting 750 ms separated by an interval of 50 ms. Twelve melodies were composed by randomly sampling notes from a chromatic scale in the ranges A2–G3<sup>#</sup> (110–207.65 Hz) for male voices, A3–G4<sup>#</sup> (220–415.3 Hz) for female voices, and A5–G6<sup>#</sup> (880–1661.22 Hz) for whistling. Notes were sampled from a uniform distribution such that every degree of the chromatic scale appeared in equal number and interval sizes followed a normal distribution centred on zero. Rhythms were composed by splitting one standard length note into two notes of half-duration (375 ms) and combining two standard length notes into one note of double duration (1500 ms). This procedure produced temporally complex stimuli with a duration and number of notes matched to the isochronous melodies. Split and combined notes occurred with equal probability at every note position to a total of 12 rhythms. Rhythmic melodies had fixed pitches at C3, C4, or C6 (130.81, 261.62, or 1046.5 Hz) for male voices, female voices, and whistling respectively.

Stimuli for singing trials were synthesized in a vocal timbre on a neutral vowel (Leon, Zero-G Limited, Okehampton, UK). Whistled stimuli were synthesized from a sine wave multiplied by an onset envelope that was empirically estimated from bilabial whistles recorded from 10 individuals (Belyk et al., 2018a).

### 2.3. Magnetic resonance imaging

MR images were acquired with a Siemens 7T MAGNETOM ultra-high field MRI with a 32-channel head coil (Nova Medical, Wilmington, USA) at the Maastricht Brain Imaging Center in Maastricht, Netherlands. Participants' heads were firmly secured with foam pillows. Noise cancelling headphones were provided to protect against auditory scanner noise. A T<sub>1</sub>-weighted image was collected using the MP2RAGE pulse sequence (Marques et al., 2010) with 0.65 mm isotropic voxels. A run of 5 volumes with the same parameters as BOLD sensitive runs was collected



**Fig. 1. Overview of methodology.** A) The source filter theory of speech outlines three biomechanical systems (left) that contribute to speech acoustics (right). The respiratory apparatus generates mechanical force. Air passing through the larynx causes it to vibrate, setting up a harmonic structure the most notable feature of which is the pitch of the voice ( $f_0$ ). The articulatory muscles shape a series of resonant chambers in the oral cavity that selectively amplify certain frequency bands ( $F_1$ ,  $F_2$ ,  $F_3$ ), which in turn encode the vowels of speech. B) Speech engages all three muscular systems, singing wordless melodies predominantly engages the respiratory and laryngeal systems, while whistling predominantly engages the respiratory and articulatory systems. C) Participants heard novel melodies and then imitated them by singing without words or by whistling. Stimuli were either melodies or rhythms, and the figure shows an example of each. A sparse-sampling paradigm ensured that auditory scanner noise did not interfere with participant behaviour.

with the phase encoding direction inverted to support image unwarping of susceptibility induced distortions.

Functional images sensitive to the BOLD signal were collected with the Center for Magnetic Resonance Research (CMRR) multi-band echoplanar imaging sequences (Moeller et al., 2010) according to a sparse event-related sampling design (Hall et al., 1999). Samples were collected 5 or 7 s after imitation onset to eliminate auditory scanner noise during stimulus presentation and task performance, as well as to minimize movement-related artefacts during image acquisition. These jittered acquisition times were selected to ensure that data were collected near the expected maxima of the BOLD response after accounting for hemodynamic lag. Images were collected with  $TR = 15$  s,  $TA = 2$  s, spatial resolution = 1.25 mm isotropic, slice gap = 0 mm,  $FOV = 160 \times 160 \times 120$  mm, number of slices = 92, echo time = 18.6 ms, flip angle =  $70^\circ$ , multi-band acceleration factor = 2. In each run 60 volumes were collected over 15 min and 15 s. Two volumes were collected and discarded prior to each run.

An additional run of BOLD-sensitive images was collected while participants remained at rest. The resting state runs had the same parameters as task-based runs, with the exception that there were no silent gaps between volume acquisitions ( $TR = 2$  s) and 360 volumes were collected per run. Resting state runs for two participants were not collected and a third participant's resting state run was ended after 174 volumes due to specific absorption rate limitations.

A second resting state BOLD sensitive dataset from a separate set of participants was analysed as a replication sample ( $N = 24$ ). These data were collected using a Siemens Trio 3T scanner located at Royal Holloway with  $TR = 1$  s, spatial resolution = 3 mm isometric, slice gap = 0.75 mm,  $FOV = 192 \times 192 \times 127$  mm, number of slices = 34,

echo time = 30 ms, flip angle =  $78^\circ$ , multi-band acceleration factor = 2. In each run 400 volumes were collected over 6 min and 40 s.

#### 2.4. Data availability

The raw 7T imaging data, image processing pipeline, experiment code, and stimulus files are accessible through the Open Science Framework (<https://osf.io/zhb5q/>).

#### 2.5. Image analysis

##### 2.5.1. Task-based runs

Susceptibility induced distortions due to magnetic field inhomogeneity were corrected using the TOPUP algorithm in FSL v5.0.10 (Jenkinson et al., 2012; Smith et al., 2004). The remainder of the MRI data processing was performed with SPM12 and MATLAB version R2017a (Mathworks, 2017) running on an iMac (OSX 10.11.6). All images were realigned to the first echo-planar image in each functional run. Runs were co-registered to the  $T_1$ -weighted images for each individual participant, and spatially normalized to the Montreal Neurological Institute standard stereotaxic space (Fonov et al., 2009) using a transformation matrix generated during tissue class segmentation (Ashburner and Friston, 2005). Images were spatially smoothed using a Gaussian kernel with Full Width Half Maximum (FWHM) equal to three voxels. Head movement was regressed from the raw data by constructing a General Linear Model (GLM) that predicted the BOLD signal from six head motion parameters and images were masked to grey matter.

### 2.5.2. Resting state runs

In addition to the steps described above, resting state functional images were slice-time corrected and normalized for global signal intensity. The mean BOLD signal within white matter as well as cerebrospinal fluid were measured separately for each volume. The mean BOLD-signal within grey matter was not included in the regression model as the mean grey matter signal is colinear with the signal of interest to resting state functional connectivity analyses (Bright et al., 2017; Murphy and Fox, 2017). The tissue-specific measures of global signal and six degrees of head motion were modelled in a fixed-effects analysis that included data pre-whitening. The residuals from this analysis were retrieved for further analysis.

## 2.6. Analysis

### 2.6.1. Partial least squares analysis

In light of the high degree of dimensionality of ultra-high field MRI, we utilized non-rotated task-based Partial Least Squares (PLS) analysis (McIntosh et al., 1996; McIntosh and Lobaugh, 2004) in lieu of the standard GLM approach (Friston et al., 1994; Worsley and Friston, 1995). Unlike statistical parametric mapping, in which each voxel is evaluated independently, PLS detects latent variables that capture a network of correlated voxels. This approach mitigates the “curse of dimensionality” by taking advantage of the spatial smoothness of fMRI data: since adjacent voxels are highly correlated and the distributed network of brain regions that participate in the same network may exhibit similar responses to the experimental design, fMRI data are amenable to dimensionality reduction.

PLS is similar to principal components analysis (PCA) with the important exception that solutions may be constrained to covariance structures that are of theoretical interest (e.g., that covary with experimental conditions). We performed non-rotated task based PLS with the orthogonal contrasts 1) task versus rest and 2) whistling versus singing. Mean centring was applied at the group level. Significance was assessed via permutation test with 500 simulations at a confidence level of 95%. Bootstrapping was performed with 500 simulations with 100 split-halves to calculate bootstrap ratios (BSR) to assess the reliability

of latent variables. We report grey matter voxels with correlations that were both unlikely under the null hypothesis according to the permutation tests ( $p < 0.025$  at either tail) and stable across bootstrap iterations ( $BSR > 2$ ). This amounts to requiring that voxels are both statistically significant (as assessed by permutation tests) and reliable (as assessed by bootstrapping; Krishnan et al., 2011).

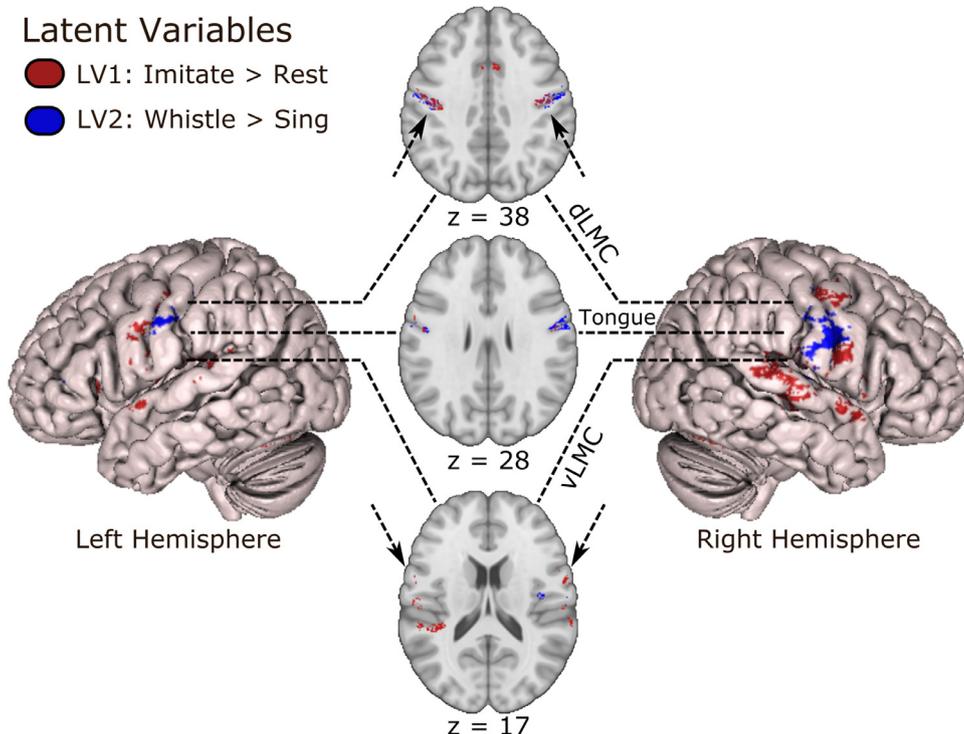
### 2.6.2. Resting state functional connectivity

Whole brain correlation maps were computed from seeds in Regions of Interest (ROIs) in the dLMC (+/-41, -16, 39), vLMC (+/-65, -4, 14), and tongue motor cortex (+/-52, -6, 28). Regions of interest were defined at 5 mm spheres around peak coordinates as localized by task based PLS and labels are based on established patterns of somatotopy (e.g., Takai et al., 2010). Group level significance was assessed using the Statistical non-Parametric Mapping (SnPM) toolbox to perform permutation tests with  $2^N$  iterations, variance smoothing with a FWHM of spanning three voxels to match the level of smoothing applied during pre-processing. Statistical maps were thresholded with a cluster-wise error rate of  $p < 0.05$  calculated from a cluster forming threshold of  $p < 0.001$ . Post-hoc tests were conducted to directly compare correlations between regions of interest to test the hypothesis that the larynx motor cortices were more strongly functionally connected with each other than with a more proximal brain region that controls a different set of muscles. Pearson correlations were calculated between each pair of ROIs within each participant to measure the degree of association between brain regions. Group level inferences were determined by Welch's paired t-tests to determine whether correlation coefficients differed from zero. The same analysis was repeated for both resting state datasets.

## 3. Results

### 3.1. PLS component 1: imitation versus rest

The first latent variable was constrained to correlate with a contrast between imitative sound production by either singing or whistling versus rest. These findings reflect regions of common activation for imitative singing and whistling (Fig. 2 and Table 1). This component loaded



**Fig. 2. Results of PLS analysis.** A) The first latent variable (red) reflects the contrast imitation > rest, regardless of whether imitation was performed by singing or by whistling and loads predominantly onto the vLMC and dLMC. The second latent variable reflects the contrast whistling > singing (blue) and loads predominantly onto tongue primary motor cortex. Lateral surface views of the brain show the spatial relationship between the vLMC and dLMC as localized by LV 1 and tongue primary motor cortex as localized by LV 2. Insets are axial slices at the level of the dLMC, tongue motor cortex, and vLMC respectively. dLMC: dorsal larynx motor cortex; LV: latent variable; vLMC: ventral larynx motor cortex.

**Table 1**

**Coordinates of activation for the first latent variable (Imitation versus Rest).** Columns indicate the name of the brain region and its anatomical division for each activation along with coordinates in MNI space. Latent Variable (LV) scores indicate the magnitude of each activation and size indicates its extent in mm. dLMC: dorsal Larynx Motor Cortex; HG: Heschel's Gyrus; PrCG: Precentral Gyrus; vLMC: ventral Larynx Motor Cortex.

Table 1: Imitation > Rest													
Label	Division	X	Y	Z	LV Score	Size (mm)	Label	BA	X	Y	Z	LV Score	Size (mm)
<i>Frontal Lobe</i>							<i>Frontal Lobe</i>						
dLMC-sulcal	BA 4	-41	-16	39	0.0072	427	dLMC-sulcal	BA 4	47	-9	38	0.0057	31
							dLMC-sulcal	BA 4	41	-14	38	0.0082	106
dLMC-gyral	BA 4	-51	-11	49	0.0057	33	dLMC-gyral	BA 4	53	-7	47	0.0087	77
							dLMC-gyral	BA 4	51	-6	49	0.0054	30
vLMC	BA 43	-59	-16	13	0.007	38	vLMC	BA 43	65	-4	14	0.0066	135
Anterior Insula	BA 13	-30	25	6	0.0054	344							
SMA	BA 6	-2	3	65	0.0057	335	SMA	BA 6	3	2	67	0.0043	80
							ACC	BA 24/32	4	14	37	0.004	117
PrCG	BA 6	-57	1	33	0.0051	49							
							PrCG	BA 4	49	-7	42	0.0075	48
PrCG	BA 4	-60	1	19	0.0045	26	PrCG	BA 4	52	-6	28	0.0036	32
							IFGop	BA 44	44	8	-1	0.0042	212
<i>Temporal Lobe</i>							<i>Temporal Lobe</i>						
HG	BA 41	-37	-35	15	0.0052	760	HG	BA 41	53	-18	7	0.0047	75
							HG	BA 41	43	-29	12	0.0041	29
aSTG	BA 22	-55	1	-3	0.007	134	aSTG	BA 22	62	2	-5	0.0057	66
							aSTG	BA 22	59	8	-7	0.0066	38
							STG	BA 22	62	-25	0	0.0066	366
							STG	BA 42	67	-28	13	0.0071	184
							STG	BA 22	65	-16	7	0.0059	82
pSTG	BA 22	-56	-39	24	0.0059	36							
<i>Subcortical</i>							<i>Subcortical</i>						
Thalamus	MGN	-2	-4	-7	0.0069	129	Thalamus	MGN	15	-27	-9	0.004	30
Thalamus	MGN	-13	-27	-10	0.0065	37							
Cerebellum	Crus VI	-24	-62	-19	0.0049	407	Cerebellum	Crus VI	31	-63	-20	0.0046	606
Cerebellum	Crus VI	-33	-58	-27	0.004	68	Cerebellum	Crus VI	15	-65	-14	0.0034	26

strongly onto voxels in the primary motor cortex in both hemispheres, including both the dLMC and vLMC. Additional regions within the motor network were also activated including the supplementary motor area (SMA), anterior cingulate cortex (ACC), and Cerebellum (Supplementary Material 1). Brain regions within the auditory network were evident, presumably due to either perceiving the target stimulus or processing auditory feedback. These brain regions included the primary auditory cortex and a large extent of the Superior Temporal Gyrus (STG), particularly in the right hemisphere, as well as the medial geniculate nucleus of the thalamus.

### 3.2. PLS component 2: whistling versus singing

The second latent variable was constrained to correlate with a contrast between imitative whistling versus singing. Surprisingly, whistling activated the dLMC and vLMC more strongly than singing in addition to the primary motor lip and tongue areas that were expected given the movements required to produce a whistled sound (Fig. 2 and Table 2). Instead, singing more strongly activated regions elsewhere in the motor network including the cerebellum and basal ganglia, temporal lobe auditory regions including the STG, Superior Temporal Sulcus (STS), Middle Temporal Gyrus (MTG) and Temporal Pole. Non-audio motor regions were also activated more strongly during singing, including the Inferior Frontal Gyrus pars orbitalis (IFGorb), anterior insula, Claustrum, and amygdala (Supplementary Material 2).

### 3.3. Resting state connectivity

Seeds in the right and left dLMC and vLMC had functional connectivity throughout sensorimotor cortex, including mutual functional connectivity between the two larynx areas. However, functional connectivity between the dLMC and vLMC was no greater than between either larynx area and tongue primary motor cortex, suggesting that functional connectivity between the two larynx motor areas may be part of the broader pattern of connectivity within primary motor cortex rather than

a result of their mutual involvement in laryngeal and respiratory motor control (Supplementary Material 3). Further functional connectivity was observed with premotor cortex (BA 6), postcentral gyrus (BA 1,2,3), the insula (BA 13), anterior cingulate gyrus (BA 24), middle cingulate gyrus (BA 31), supplementary motor area (BA 6), and the inferior frontal gyrus pars opercularis (BA 44; Supplementary Material 4).

## 4. Discussion

We report the first ultra-high field fMRI comparison of singing alongside whistling to contrast sound production with and without involvement of the laryngeal sound source. Unexpectedly, given the reduced laryngeal involvement in whistling, we observed that both the dLMC and vLMC were not only engaged by whistling, but more strongly than singing. The strong expiratory drive of whistling may account for the common activation in these brain regions across modes of sound production. One previous study of whistling observed activation that may have been consistent with the vLMC (Dresel et al., 2005), but without explicit localisation of the larynx areas this correspondence could not be confirmed. We suggest that in addition to their established roles in laryngeal motor control, the dLMC and vLMC also contribute to respiratory motor control and may serve to integrate two muscular systems that require mutual coordination to support important behaviours such as speaking, singing, and airway protection.

That neither of the larynx motor areas in the human brain are specific to laryngeal motor control may be due to the placement of the larynx within the airway. Voiced sounds are produced by the vibration of the vocal folds, the tension of which determines vocal pitch (Hollien and Moore, 1960; Titze, 2008; Titze et al., 1989; Titze and Story, 2002). However, the laryngeal muscles do not actively vibrate the vocal folds, but instead they determine the configuration of the larynx while vocal fold vibration is produced passively by the passage of air (Story and Titze, 1995; Titze, 1989). In addition to its role in communication, the larynx also serves as a mechanism for airway protection (Dua et al., 1997) and participates in reflexive movements during swallowing and

**Table 2**

**Latent variable two (Whistling versus Singing).** Columns indicate the name of the brain region and its anatomical division for each activation along with coordinates in MNI space. Latent Variable (LV) scores indicate the magnitude of each activation and size indicates its extent in mm. dLMC: dorsal Larynx Motor Cortex; HG: Heschel's Gyrus; IFGorb: Inferior Frontal Gyrus pars orbitalis; IFGop: Inferior Frontal Gyrus pars opercularis; MTG: Middle Temporal Gyrus; PrCG: Precentral-Gyrus; RO: Rolandic Operculum; STG: Superior Temporal Gyrus; STS: Superior Temporal Sulcus; vLMC: ventral Larynx Motor Cortex.

Table 2a: Whistling &gt; Singing

Left Hemisphere							Right Hemisphere						
Label	Division	X	Y	Z	LV Score	Size (mm)	Label	Division	X	Y	Z	LV Score	Size (mm)
<i>Frontal Lobe</i>							<i>Frontal Lobe</i>						
dLMC-sulcal	BA 4	-49	-11	36	0.0076	385	dLMC-sulcal	BA 4	43	-13	37	0.0096	66
PrCG	BA 4	-54	-11	38	0.0062	71	vLMC	BA 43	61	-15	20	0.0067	34
							PrCG	BA 4	44	-10	37	0.0086	31
							PrCG	BA 4	65	-6	27	0.0105	865
							RO	BA 13	53	-18	13	0.0060	64
							RO	BA 13	44	-12	21	0.0059	92
<i>Parietal Lobe</i>							<i>Parietal Lobe</i>						
PoCG	BA 3	-45	-19	40	0.0063	48	–						
<i>Temporal Lobe</i>							<i>Temporal Lobe</i>						
–							Parahippocampal Gyrus	BA 28	26	-16	-24	0.0083	32
<i>Subcortical</i>							<i>Subcortical</i>						
–							Cerebellum	Crus VI	20	-62	-17	0.0045	26
Table 2b: Singing > Whistling													
<i>Frontal Lobe</i>							<i>Frontal Lobe</i>						
IFGorb	BA 47	-45	37	-12	0.0040	237	IFGorb	BA 47	48	38	-9	0.0040	32
IFGorb	BA 47	-34	29	-13	0.0040	127							
IFGorb	BA 47	-30	35	-10	0.0041	95							
IFGorb	BA 47	-42	42	-2	0.0041	64							
IFGorb	BA 47	-40	25	-5	0.0041	41							
IFGorb	BA 47	-39	25	1	0.0041	29							
IFGorb	BA 47	-28	22	-13	0.0041	26							
IFGop	BA 44	-51	12	1	0.0041	27							
							PrCG	BA 6	53	3	8	0.0040	236
							RO		58	-6	9	0.0042	42
							Subgenual ACC	BA 32	8	49	-12	0.0041	75
							<i>Temporal Lobe</i>						
Insula	BA 3	-35	10	-7	0.0040	149	aSTG	BA 22	53	2	-8	0.0040	676
Anteria Insula	BA 13	-39	17	5	0.0041	75	STS	BA 21	56	-24	-3	0.0040	236
Anteria Insula	BA 13	-32	24	-2	0.0041	37							
<i>Temporal Lobe</i>							<i>Temporal Lobe</i>						
STG	BA 22	-58	-10	-1	0.0040	164	Temporal Pole	BA 38	45	18	-22	0.0040	39
STG	BA 22	-48	-5	-8	0.0041	26							
STG	BA 22	-59	-3	-5	0.0040	73	Parahippocampal Gyrus	BA 34	23	1	-14	0.0041	36
MTG	BA 21	-60	-12	-4	0.0040	61							
MTG	BA 21	-49	-8	-15	0.0040	50							
MTG	BA 21	-58	-1	-16	0.0040	45							
MTG	BA 21	-49	0	-28	0.0041	39							
Temporal Pole	BA 38	-43	13	-29	0.0040	80							
Temporal Pole	BA 38	-49	15	-17	0.0041	76							
Temporal Pole	BA 38	-39	3	-25	0.0041	70							
Parahippocampal Gyrus	BA 34	-24	-4	-23	0.0041	138							
<i>Subcortical</i>							<i>Subcortical</i>						
Cerebellum	Crus VI	-27	-39	-29	0.0041	108	Cerebellum	Crus VI	37	-56	-22	0.0041	55
Cerebellum	Crus V	-24	-32	-17	0.0040	81							
Cerebellum	Crus I	-26	-84	-21	0.0041	42							
Clastrum		-30	10	1	0.0040	27							
Basal Ganglia	Caudate Nucleus	-2	8	-4	0.0040	79							
Amygdala		-26	0	-22	0.0041	85							

retching (Ardran and Kemp, 1952; Lang et al., 2002) to form a secondary closure below the epiglottis (Vilkman et al., 1996).

An informal laryngoscopic investigation revealed that whistled notes are often interspersed with a light closure of the glottis (see Supplementary Materials 5 and 6), which is presumed to be part of the mechanism for the cessation of airflow. Glottal closure has previously been shown to activate the dLMC, though it is unknown whether it contributes to activation of the vLMC (Brown et al., 2008). However, it is sensible to suppose that part of the LMCs joint mechanism of laryngeal and airway control should include the closure of the glottis. Regardless, this reduced involvement of the larynx during whistling is unlikely to account for the greater activation of the LMCs.

To date there have been few demonstrations that the dLMC and vLMC may have dissociable functions. While the dLMC is often reported in the absence of the vLMC, it seems likely that this merely reflects differences in the ease with which these areas are detected. This may be due to the greater abundance of large descending motor neurons in the dLMC (Brodmann, 1909; Judaš and Cepanec, 2010; Vogt, 1910), which may generate a larger BOLD signal. For instance, simple phonation (singing without words) activates only the dLMC in some studies (Belyk et al., 2018b; Brown et al., 2008) but both the dLMC and vLMC in others (Belyk et al., 2016). Furthermore, the sulci near the vLMC are more variable, which is expected to reduce rates of detection in group level analyses (Eichert et al., 2020b). Hence, while we observed stronger evidence for the involvement of the dLMC compared to the vLMC in integrating laryngeal and respiratory control, this may reflect differences in the detectability rather than function. One interesting exception is that the application of an external puff of air on the surface of the larynx may activate the vLMC, suggesting either that the vLMC has some sensory functionality or that this manipulation triggers airway protective reflexes that that involve the vLMC (Miyaji et al., 2014).

#### 4.1. A note on nomenclature

Early brain imaging studies disagreed on an appropriate label for what has come to be referred to as the dLMC (Brown et al., 2008; Loucks et al., 2007; Simonyan et al., 2007). The present findings, contribute to growing evidence that specificity of function implied by the label LMC is not borne out by the data (Brown et al., 2021, 2008; Loucks et al., 2007). While we do not undertake to reform the existing nomenclature here, we note that there is a growing need to do so.

#### 4.2. Respiratory motor control

The human brain has a direct projection from cortex to the nucleus ambiguus, the brainstem nucleus that exerts direct control over the laryngeal muscles (Iwatsubo et al., 1990; Kuypers, 1958a). This direct pathway is absent in the brains of monkeys (Jürgens and Ehrenreich, 2007; Simonyan and Jürgens, 2003), and less robust in the brains of non-human apes (Kuypers, 1958b). However, this novel arrangement in human laryngeal motor control is not sufficient to explain the joint cortical control of laryngeal and respiratory functions observed in the present study.

Respiratory motor control is coordinated by the nucleus retroambiguus (Figure 3) which is a brainstem nucleus of the medulla located adjacent to the nucleus ambiguus (Subramanian and Holstege, 2009). The nucleus retroambiguus contains upper motor neurons that project to lower motor neuron nuclei that innervate respiratory muscles such as the abdomen, pelvic floor, and intercostal muscles (Vanderhorst et al., 2000). The nucleus retroambiguus also projects to laryngeal motor neurons in the nucleus ambiguus (VanderHorst et al., 2001).

However, this retroambiguus-ambiguus projection appears to be unidirectional, such that cortical control over the nucleus ambiguus alone is unlikely to yield voluntary respiratory motor control. We therefore hypothesize that both human LMCs have direct projections to the nucleus retroambiguus in addition to the direct projections to the nucleus

ambiguus that have already been observed. Such a parallel projection to both nucleus ambiguus and retroambiguus has not yet been identified in humans but is observed in the analogous structures of the avian song system, which is a strong model of the human vocal motor system (Gahr, 2000; Petkov and Jarvis, 2012; Wild, 1993; Wild et al., 2000).

Observations of the direct projection to nucleus ambiguus have come from natural experiments due to cerebrovascular events (Iwatsubo et al., 1990; Kuypers, 1958a). In these studies, large cortical lesions caused the axons of upper motor neurons to degenerate, and degenerating axons were traced among more intact white matter. However, the lesions all resulted from cerebrovascular accidents of the middle cerebral artery (MCA). This artery supplies much of the speech-motor related cortex including both the vLMC and dLMC, hence the prevalence of speech-motor and swallowing disorders following MCA infarcts (Heinsius et al., 1998; Theys et al., 2011). Consequently, it is not known whether the direct connection to the nucleus ambiguus originates from one or both larynx motor areas.

We hypothesize that the direct connection to the nucleus ambiguus stems from both the vLMC and dLMC in light of the lack of strong functional connectivity between these cortical larynx areas. We have previously hypothesized that the dLMC may have a yet detected direct projection to the nucleus retroambiguus (Belyk and Brown, 2017), consistent with the analogous pathway in songbirds (Gahr, 2000; Petkov and Jarvis, 2012; Wild, 1993; Wild et al., 2000). We now further hypothesize that the human vLMC may also project directly to the nucleus retroambiguus considering the evidence provided by the current experiment that both larynx controlling regions integrate respiratory motor control.

#### 4.3. Respiratory motor control of species-typical vocalisations

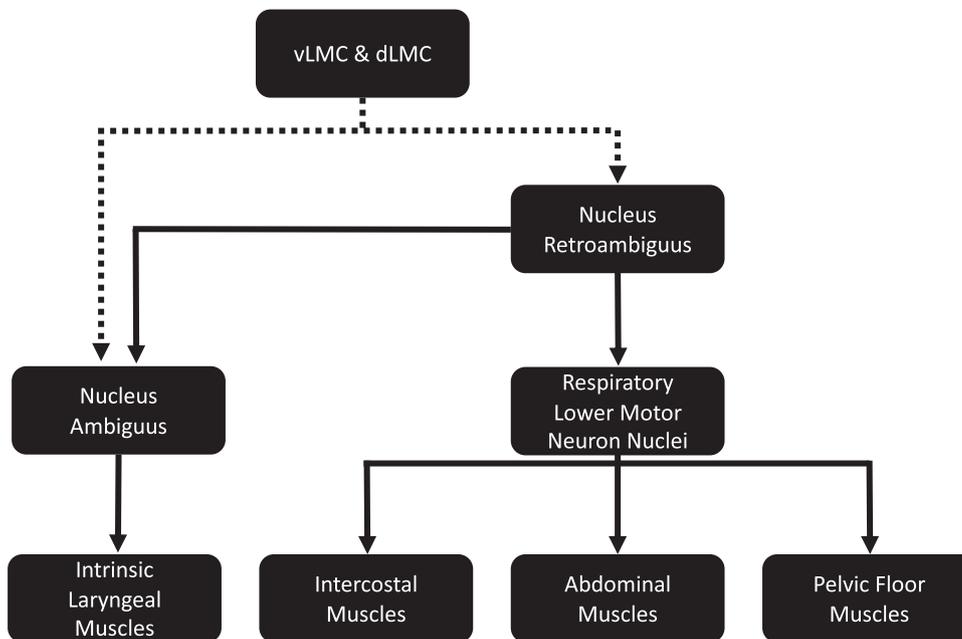
While non-human primates have relatively poor ability to learn novel vocal behaviours (Hage and Nieder, 2016; Nieder and Mooney, 2019), they can strategically select them to deploy their innate repertoire of vocalisation (Pierce, 1985). This behaviour requires some degree of voluntary control over the timing of laryngeal-respiratory action in animals that lack the mechanisms provided by the human vLMC/dLMC (Jürgens, 1974).

In human and non-human primates alike, the control of innate vocalisations is controlled separately by a pathway which circumvents motor cortex (Jürgens, 2002). The periaqueductal gray (PAG) of the brainstem organises responses to affective stimuli, and as such it receives extensive inputs from the limbic system (Dujardin and Jürgens, 2005) and has outputs to a plurality of lower motor nuclei relevant to vocalisation providing integrated control over multiple muscle groups (Thoms and Jürgens, 1987). In monkeys (*Saimiri sciureus*), electrical stimulation of the PAG elicits fully formed species-typical vocalisations, including both the laryngeal and respiratory components (Jürgens and Pratt, 1979a, 1979b).

A region of cingulate cortex provides cortical control over the PAG. Lesions to the cingulate cortex prevent the initiation of operantly conditioned vocalizations, but spare responses to stimuli that would normally elicit an innate vocalisation (Aitken, 1981; Sutton et al., 1981, 1974). The homologous pathway in humans is activated during verbal expressions of emotion (Barrett et al., 2004; Belyk and Brown, 2016; Wattendorf et al., 2013). This cingulate-PAG axis provides a mechanism for coordinated laryngeal-respiratory control of innate species-typical vocalisations, which is conserved across primates. In humans, this pathway exists alongside the novel adaptations in motor cortex which support the flexibility of human vocal behaviour.

#### 4.4. Selective pressures on respiratory motor control

Strong vocal production learning abilities are uncommon in mammals, though they have been documented in a handful of clades. The list of strong mammalian vocal learners is notably skewed towards species



**Fig. 3. Schematic of descending motor pathway.** Cortico-bulbar pathways (dotted line) have been observed projecting to the nucleus ambiguus though it is not presently known whether these axons originate from the vLMC, the dLMC or both. We have hypothesized that both areas project to the nucleus retroambiguus as well.

whose evolutionary path has placed particular constraints on respiratory motor control. VPL is most abundantly observed among aquatic mammals including cetaceans (Janik, 2014; King and Sayigh, 2013; Noad et al., 2000) and pinnipeds (Ralls et al., 1985; Ravignani et al., 2016; Stansbury and Janik, 2019) which must coordinate breathing with bouts of diving to manage the supply of oxygen, buoyancy, and ambient ocean pressure (Kooyman, 1973; Lillie et al., 2017; Roos et al., 2016). Some species of elephants have demonstrated vocal production learning (Poole et al., 2005; Stoeger et al., 2012), which may be related to the unique demands of respiratory snorkelling West (2001) as well as the possibly aquatic ancestry of these species (Gaeth et al., 1999). Bats exhibit a range of socially communicative vocalizations in addition to echolocation (Knörnschild, 2014; Vernes and Wilkinson, 2019; Vernes, 2016). Notably, bats not only integrate respiratory control with echo location calls that are used for navigation, but the muscles of respiration may interact directly with the muscular control of winged flight (Lancaster et al., 1995; Suthers et al., 1972). In this company, humans appear to be the odd-mammal out in lacking a clear selective pressure for enhanced respiratory motor control beyond its use in communication (Verhaegen et al., 2002).

#### 4.5. Limitations

The present study was based on a relatively small sample of singers due to the constraints of ultra-high field fMRI. We have mitigated the low power associated with this small sample size by utilising Partial Least Squares analysis, which considerably improves the sensitivity and reliability of statistical maps (Grady et al., 2020). PLS analyses with sample sizes in the range of the present study have reliability equivalent to a standard univariate analysis of a moderately large sample. This advantage is derived in part from analysing networks of correlated brain regions rather than a mass of independent voxels. However, Grady et al. caution that this should not be used as a substitute for thoughtful experimental design and well-motivated sample sizes. The present study tested a-priori spatially specific hypotheses such that the trade-off between sample size and spatial resolution was consistent with the aims of the experiment.

## 5. Conclusion

The dLMC and vLMC are two larynx motor areas in the human brain that are important cortical structures for the voluntary control of the voice. We observed that both areas are also active during whistling, despite the reduced laryngeal involvement in that mode of sound production. We suggest that neither the dLMC nor the vLMC are strictly laryngeal, and that both may integrate laryngeal and respiratory motor control. Some clue to the separate functions of these brain regions may be found in the complex cytoarchitecture of the vLMC, which appears to be intermediate to primary motor and primary somatosensory cortex. Regardless, coordination with respiratory motor control appears to be a ubiquitous partner to laryngeal motor control as indeed the larynx sits in the airway and any action of the larynx is likely to affect respiratory effort. This has implications for our understanding of human brain evolution to the extent that it alters our understanding of a well-documented specialisation for speech .

## Acknowledgements

We thank Aaron Low and The Voice Clinic for supporting laryngoscopic investigations.

## Funding

This work was supported by grants from the Natural Sciences and Engineering Research Council of Canada (grant no. PDF502954-2017), the Kimmel Family Opportunity Fund, the Biotechnical and Biological Sciences Research Council of the UK (grant no. BB/M009742/1), an ERC Starting Grant (MULTICONNECT, #639938, a Dutch science foundation (NWO) VIDI Grant (#14637), and the Faculty of Psychology and Neuroscience, Maastricht Brain Imaging Center fund.

## Author's contributions

MB conceived of and designed the study, analysed the data and prepared the manuscript. RB contributed to experimental design, and collected the 7T dataset. DSB consulted on data analysis, edited the



manuscript, and organised laryngoscopic investigations. AR consulted on experimental design and data analysis. CM provided critical comments on the manuscript. RG contributed the 3T dataset. SAK consulted on experimental design, edited, and provided critical comments on the manuscript.

## Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.neuroimage.2021.118326.

## References

- Aitken, P.G., 1981. Cortical control of conditioned and spontaneous vocal behavior in rhesus monkeys. *Brain Lang.* 13, 171–184.
- Ardran, G., Kemp, F., 1952. The protection of the laryngeal airway during swallowing. *Br. J. Radiol.* 25, 406–416.
- Ashburner, J., Friston, K.J., 2005. Unified segmentation. *Neuroimage* 26, 839–851. doi:10.1016/j.neuroimage.2005.02.018.
- Barrett, J., Pike, G.B., Paus, T., 2004. The role of the anterior cingulate cortex in pitch variation during sad affect. *Eur. J. Neurosci.* 19, 458–464.
- Belyk, M., Brown, S., 2017. The origins of the vocal brain in humans. *Neurosci. Biobehav. Rev.* 77, 177–193. doi:10.1016/j.neubiorev.2017.03.014.
- Belyk, M., Brown, S., 2016. Pitch underlies activation of the vocal system during affective vocalization. *Soc. Cogn. Affect. Neurosci.* 11, 1078–1088. doi:10.1093/scan/nsv074.
- Belyk, M., Johnson, J.F., Kotz, S., 2018a. Poor neuro-motor tuning of the human larynx: A comparison of sung and whistled pitch imitation. *R. Soc. Open Sci.* 5, 171544. doi:10.1098/rsos.171544.
- Belyk, M., Lee, Y.S., Brown, S., 2018b. How does human motor cortex regulate vocal pitch in singers? *R. Soc. Open Sci.* 5, 172208. doi:10.1098/rsos.172208.
- Belyk, M., Pfordresher, P.Q., Liotti, M., Brown, S., 2016. The neural basis of vocal pitch imitation in humans. *J. Cogn. Neurosci.* 28, 621–635. doi:10.1162/jocn.
- Belyk, M., Schultz, B.G., Correia, J., Beal, D.S., Kotz, S.A., 2019. Whistling shares a common tongue with speech: bioacoustics from real-time MRI of the human vocal tract. *Proc. R. Soc. B* 286, 20191116. doi:10.1098/rspb.2019.1116.
- Bouchard, K.E., Mesgarani, N., Johnson, K., Chang, E.F., 2013. Functional organization of human sensorimotor cortex for speech articulation. *Nature* 495, 327–332. doi:10.1038/nature11911.
- Breshears, J.D., Molinaro, A.M., Chang, E.F., 2015. A probabilistic map of the human ventral sensorimotor cortex using electrical stimulation. *J. Neurosurg.* 123, 340–349. doi:10.3171/2014.11.JNS14889.
- Bright, M.G., Tench, C.R., Murphy, K., 2017. Potential pitfalls when denoising resting state fMRI data using nuisance regression. *Neuroimage* 154, 159–168. doi:10.1016/j.neuroimage.2016.12.027.
- Brodman, K., 1909. *Localisation in the Cerebral Cortex*, 3rd ed. Springer, New York.
- Brown, S., Laird, A.R., Pfordresher, P.Q., Thelen, S.M., Turkeltaub, P., Liotti, M., 2009. The somatotopy of speech: Phonation and articulation in the human motor cortex. *Brain Cogn* 70, 31–41. doi:10.1016/j.bandc.2008.12.006.
- Brown, S., Martinez, M.J., Hodges, D.A., Fox, P.T., Parsons, L.M., 2004. The song system of the human brain. *Cogn. Brain Res* 20, 363–375. doi:10.1016/j.cogbrainres.2004.03.016.
- Brown, S., Ngan, E., Liotti, M., 2008. A larynx area in the human motor cortex. *Cereb. Cortex* 18, 837–845. doi:10.1093/cercor/bhm131.
- Brown, S., Yuan, Y., Belyk, M., 2021. Evolution of the speech-ready brain: The voice/jaw connection in the human motor cortex. *J. Comp. Neurol.* 529, 1018–1028. doi:10.1002/cne.24997.
- Chang, E.F., Niziolek, C.A., Knight, R.T., Nagarajan, S.S., Houde, J.F., 2013. Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proc. Natl. Acad. Sci.* 110, 2653–2658. doi:10.1073/pnas.1216827110.
- Dichter, B.K., Breshears, J.D., Leonard, M.K., Chang, E.F., 2018. The control of vocal pitch in human laryngeal motor cortex. *Cell* 174, 21–31. doi:10.1016/j.cell.2018.05.016.e9.
- Dresel, C., Castrop, F., Haslinger, B., Wohlschlaeger, A.M., Hennenlotter, A., Ceballos-Baumann, A.O., 2005. The functional neuroanatomy of coordinated orofacial movements: sparse sampling fMRI of whistling. *Neuroimage* 28, 588–597. doi:10.1016/j.neuroimage.2005.06.021.
- Dua, K.S., Ren, J., Bardan, E., Xie, P., Shaker, R., 1997. Coordination of deglutitive glottal function and pharyngeal bolus transit during normal eating. *Gastroenterology* 112, 73–83.
- Dujardin, E., Jürgens, U., 2005. Afferents of vocalization-controlling periaqueductal regions in the squirrel monkey. *Brain Res.* 1034, 114–131. doi:10.1016/j.brainres.2004.11.048.
- Eichert, N., Papp, D., Mars, R.B., Watkins, K.E., 2020a. Mapping human laryngeal motor cortex during vocalization. *Cereb. Cortex* 30, 6254–6269. doi:10.1101/2020.02.20.958314.
- Eichert, N., Watkins, K.E., Mars, R.B., Petrides, M., 2020b. Morphological and functional variability in central and subcentral motor cortex of the human brain. *Brain Struct. Funct.* doi:10.1007/s00429-020-02180-w.
- Fischer, J., Hammerschmidt, K., 2019. Towards a new taxonomy of primate vocal learning. *Philos. Trans. R. Soc. B Biol. Sci.* 375, 20199945. doi:10.1098/rstb.2019.0045.
- Fonov, V.V., Evans, A.C.A., Botteron, K., Almlí, C.R., McKinstry, R.R.C., Collins, D.L., Almlí, C.R., Collins, D.L., 2009. Unbiased nonlinear average age-appropriate brain templates from birth to adulthood. *Neuroimage* 47, S102. doi:10.1016/j.neuroimage.2010.07.033.
- Friston, K., Holmes, A., Worsley, K., Frackowiak, R., 1994. Statistical parametric maps in functional imaging: A general linear approach. *Hum. Brain Mapp* 2, 189–210.
- Gaeth, A.P., Short, R.V., Renfree, M.B., 1999. The developing renal, reproductive, and respiratory systems of the African elephant suggest an aquatic ancestry. *Proc. Natl. Acad. Sci. U. S. A.* 96, 5555–5558. doi:10.1073/pnas.96.10.5555.
- Gahr, M., 2000. Neural song control system of hummingbirds: Comparison to swifts, vocal learning (songbirds) and nonlearning (suboscines) passerines, and vocal learning (budgerigars) and nonlearning (dove, owl, gull, quail, chicken) nonpasserines. *J. Comp. Neurol.* 426, 182–196. doi:10.1002/1096-9861(20001016)426:2<182::AID-CNE2>3.0.CO;2-M.
- Grabski, K., Schwartz, J.-L., Lamalle, L., Vilain, C., Vallée, N., Baciu, M., Le Bas, J.-F., Sato, M., 2013. Shared and distinct neural correlates of vowel perception and production. *J. Neurolinguistics* 26, 384–408. doi:10.1016/j.jneuroling.2012.11.003.
- Grady, C.L., Rieck, J.R., Nichol, D., Rodrigue, K.M., Kennedy, K.M., 2020. Influence of sample size and analytic approach on stability and interpretation of brain-behavior correlations in task-related fMRI data. *Hum. Brain Mapp.* 1–16. doi:10.1002/hbm.25217.
- Hage, S.R., Nieder, A., 2016. Dual neural network model for the evolution of speech and language. *Trends Neurosci.* 39, 813–829. doi:10.1016/j.tins.2016.10.006.
- Hall, D.A., Haggard, M.P., Akeroyd, M.A., Palmer, A., Summerfield, A.Q., Elliott, M.R., Gurney, E.M., Bowtell, R.W., 1999. Sparse temporal sampling in auditory fMRI. *Hum. Brain Mapp* 7, 213–223.
- Hallermann, S., De Kock, C.P.J., Stuart, G.J., Kole, M.H.P., 2012. State and location dependence of action potential metabolic cost in cortical pyramidal neurons. *Nat. Neurosci.* 15, 1007–1014. doi:10.1038/nn.3132.
- Heinsius, T., Bogousslavsky, J., Van Melle, G., 1998. Large infarcts in the middle cerebral artery territory: etiology and outcome patterns. *Neurology* 50, 341–350. doi:10.1212/WNL.50.2.341.
- Hollien, H., Moore, P., 1960. Measurements of the vocal folds during changes in pitch. *J. Speech, Lang. Hear. Res.* 3, 157–165. doi:10.1044/jshr.0302.157.
- Iwatsubo, T., Kuzuhara, S., Kanemitsu, A., 1990. Corticofugal projections to the motor nuclei of the brainstem and spinal cord in humans. *Neurology* 40, 309–312. doi:10.1212/WNL.40.2.309.
- Janik, V.M., 2014. Cetacean vocal learning and communication. *Curr. Opin. Neurobiol.* 28, 60–65. doi:10.1016/j.conb.2014.06.010.
- Janik, V.M., Slater, P.J.B., 2000. The different roles of social learning in vocal communication. *Anim. Behav.* 60, 1–11. doi:10.1006/anbe.2000.1410.
- Jenkinson, M., Beckmann, C.F., Behrens, T.E.J., Woolrich, M.W., Smith, S.M., 2012. FSL. *Neuroimage* 62, 782–790. doi:10.1016/j.neuroimage.2011.09.015.
- Ju, H., Hines, M.L., Yu, Y., 2016. Cable energy function of cortical axons. *Sci. Rep.* 6, 1–13. doi:10.1038/srep29686.
- Judaš, M., Cepanec, M., 2010. Oskar Vogt: The first myeloarchitectonic map of the human frontal cortex. *Transl. Neurosci.* 1, 72–94. doi:10.2478/v10134-010-0005-z.
- Jürgens, U., 2002. Neural pathways underlying vocal control. *Neurosci. Biobehav. Rev.* 26, 235–258. doi:10.1016/S0149-7634(01)00068-9.
- Jürgens, U., 1974. On the elicibility of vocalization from the cortical larynx area. *Brain Res* 81, 564–566. doi:10.1016/0006-8993(74)90853-1.
- Jürgens, U., Ehrenreich, L., 2007. The descending motorcortical pathway to the laryngeal motoneurons in the squirrel monkey. *Brain Res.* 1148, 90–95. doi:10.1016/j.brainres.2007.02.020.
- Jürgens, U., Pratt, R., 1979a. The singular vocalization pathway in the squirrel monkey. *Exp. Brain Res.* 510, 499–510.
- Jürgens, U., Pratt, R., 1979b. Role of the periaqueductal grey in vocal expression of emotion. *Brain Res* 167, 367–378.
- King, S., Sayigh, L., 2013. Vocal copying of individually distinctive signature whistles in bottlenose dolphins. *Proc. R. Soc. B* 280, 1–9. doi:10.1098/rspb.2013.0053.
- Kleber, B., Zeitouni, A.G., Friberg, A., Zatorre, R.J., 2013. Experience-dependent modulation of feedback integration during singing: role of the right anterior insula. *J. Neurosci.* 33, 6070–6080. doi:10.1523/JNEUROSCI.4418-12.2013.
- Knörnschild, M., 2014. Vocal production learning in bats. *Curr. Opin. Neurobiol.* 28, 80–85.
- Kooyman, G.L., 1973. Respiratory adaptations in marine mammals. *Integr. Comp. Biol.* 13, 457–468. doi:10.1093/icb/13.2.457.
- Krishnan, A., Williams, L.J., McIntosh, A.R., Abdi, H., 2011. Partial least squares (PLS) methods for neuroimaging: a tutorial and review. *Neuroimage* 56, 455–475. doi:10.1016/j.neuroimage.2010.07.034.
- Kuypers, H.G.J.M., 1958a. Corticobulbar connexions to the pons and lower brain-stem in man. *Brain* 81, 364–388. doi:10.1093/brain/81.3.364.
- Kuypers, H.G.J.M., 1958b. Some projections from the peri-central cortex to the pons and lower brain stem in monkey and chimpanzee. *J. Comp. Neurol.* 110, 221–255. doi:10.1002/cne.901100205.
- Lameira, A.R., Hardus, M.E., Kowalsky, B., de Vries, H., Spruijt, B.M., Sterck, E.H.M., Shumaker, R.W., Wich, S.A., 2013. Orangutan (*Pongo spp.*) whistling and implications for the emergence of an open-ended call repertoire: a replication and extension. *J. Acoust. Soc. Am.* 134, 2326–2335. doi:10.1121/1.4817929.
- Lameira, A.R., Hardus, M.E., Mielke, A., Wich, S.A., Shumaker, R.W., 2016. Vocal fold control beyond the species-specific repertoire in an orangutan. *Sci. Rep.* 6, 1–10. doi:10.1038/srep30315.
- Lancaster, W.C., Henson, O.W., Keating, A.W., 1995. Respiratory muscle activity in relation to vocalization in flying bats. *J. Exp. Biol.* 198, 175–191.
- Lang, I.M., Dana, N., Medda, B.K., Shaker, R., 2002. Mechanisms of airway protection during retching, vomiting, and swallowing. *Am. J. Physiol. - Gastrointest. Liver Physiol.* 283, G529–G536. doi:10.1152/ajpgi.00062.2002.
- Lillie, M.A., Vogl, A.W., Raverty, S., Haukena, M., McLellan, W.A., Stenson, G.B., Shadwick, R.E., 2017. Controlling thoracic pressures in cetaceans during a

- breath-hold dive: Importance of the diaphragm. *J. Exp. Biol.* 220, 3464–3477. doi:10.1242/jeb.162289.
- Lotze, M., Erb, M., Flor, H., Huelsmann, E., Godde, B., Grodd, W., 2000. fMRI evaluation of somatotopic representation in human primary motor cortex. *Neuroimage* 11, 473–481. doi:10.1006/nimg.2000.0556.
- Loucks, T.M.J., Poletto, C.J., Simonyan, K., Reynolds, C.L., Ludlow, C.L., 2007. Human brain activation during phonation and exhalation: common volitional control for two upper airway functions. *Neuroimage* 36, 131–143. doi:10.1016/j.neuroimage.2007.01.049.
- Marques, J.P., Kober, T., Krueger, G., van der Zwaag, W., Van de Moortele, P.F., Gruetter, R., 2010. MP2RAGE, a self bias-field corrected sequence for improved segmentation and T1-mapping at high field. *Neuroimage* 49, 1271–1281. doi:10.1016/j.neuroimage.2009.10.002.
- Mathworks, 2017. MATLAB, version R2017a.
- McIntosh, A.R., Bookstein, F.L., Haxby, J.V., Grady, C.L., 1996. Spatial pattern analysis of functional brain images using partial least squares. *Neuroimage* 3, 143–157. doi:10.1006/nimg.1996.0016.
- McIntosh, A.R., Lobaugh, N.J., 2004. Partial least squares analysis of neuroimaging data: applications and advances. *Neuroimage* 23, 250–263. doi:10.1016/j.neuroimage.2004.07.020.
- Miyajiri, H., Hironaga, N., Umezaki, T., Hagiwara, K., Shigeto, H., Sawatsubashi, M., Tobimatsu, S., Komune, S., 2014. Neuromagnetic detection of the laryngeal area: Sensory-evoked fields to air-puff stimulation. *Neuroimage* 88, 162–169. doi:10.1016/j.neuroimage.2013.11.008.
- Moeller, S., Yacoub, E., Olman, C.A., Auerbach, E., Strupp, J., Harel, N., Ugurbil, K., 2010. Multiband multislice GE-EPI at 7 tesla, with 16-fold acceleration using partial parallel imaging with application to high spatial and temporal whole-brain fMRI. *Magn. Reson. Med.* 63, 1144–1153. doi:10.1002/mrm.22361.
- Murphy, K., Fox, M.D., 2017. Towards a consensus regarding global signal regression for resting state functional connectivity MRI. *Neuroimage* 154, 169–173. doi:10.1016/j.neuroimage.2016.11.052.
- Nieder, A., Mooney, R., 2019. The neurobiology of innate, volitional and learned vocalizations in mammals and birds. *Philos. Trans. R. Soc. B Biol. Sci.* 375, 20190054. doi:10.1098/rstb.2019.0054.
- Noad, M.J., Cato, D.H., Bryden, M.M., Jenner, M.N., Jenner, K.C., 2000. Cultural revolution in whale songs. *Nature* 408, 537. doi:10.1038/35046199.
- Olthoff, A., Baudewig, J., Kruse, E., Dechent, P., 2008. Cortical sensorimotor control in vocalization: a functional magnetic resonance imaging study. *Laryngoscope* 118, 2091–2096. doi:10.1097/MLG.0b013e31817fd40f.
- Peck, K.K., Galgano, J.F., Branski, R.C., Bogomolny, D., Ho, M., Holodny, A.I., Kraus, D.H., 2009. Event-related functional MRI investigation of vocal pitch variation. *Neuroimage* 44, 175–181. doi:10.1016/j.neuroimage.2008.08.034.
- Penfield, W., Boldrey, E., 1937. Somatic motor and sensory representations in the cerebral cortex of man as studied by electrical stimulation. *Brain* 60, 389–443. doi:10.1192/bjp.84.352.868-a.
- Perlman, M., Clark, N., 2015. Learned vocal and breathing behavior in an enculturated gorilla. *Anim. Cogn.* 18, 1165–1179. doi:10.1007/s10071-015-0889-6.
- Petkov, C.I., Jarvis, E.D., 2012. Birds, primates, and spoken language origins: Behavioral phenotypes and neurobiological substrates. *Front. Evol. Neurosci.* 4, 1–24. doi:10.3389/fnevo.2012.00012.
- Pierce, J.D., 1985. A review of attempts to condition operantly allprimate vocalizations. *Primates* 26, 202–213. doi:10.1007/BF02382019.
- Poole, J.H., Tyack, P.L., Stoeger-Horwath, A.S., Watwood, S., 2005. Elephants are capable of vocal learning. *Nature* 434, 455–456. doi:10.1029/2001GL014051.
- Ralls, K., Fiorelli, P., Gish, S., 1985. Vocalizations and vocal mimicry in captive harbor seals, *Phoca vitulina*. *Can. J. Zool.* 63, 1050–1056. doi:10.1139/z85-157.
- Rao, S.M., Binder, J.R., Hammeke, T.A., Bandettini, P.A., Bobholz, M.A., Frost, J.A., Myklebust, B.M., Jacobson, R.D., Hyde, J.S., 1995. Somatotopic mapping of the human primary motor cortex with functional magnetic resonance imaging. *Neurology* 45, 919–924. doi:10.1212/WNL.45.5.919.
- Ravignani, A., Fitch, W.T., Hanke, F.D., Heinrich, T., Hurgitsch, B., Kotz, S.A., Scharff, C., Stoeger, A.S., de Boer, B., 2016. What pinnipeds have to say about human speech, music, and the evolution of rhythm. *Front. Neurosci.* 10, 274. doi:10.3389/fnins.2016.00274.
- Roos, M.M.H., Wu, G.M., Miller, P.J.O., 2016. The significance of respiration timing in the energetics estimates of free-ranging killer whales (*Orcinus orca*). *J. Exp. Biol.* 219, 2066–2077. doi:10.1242/jeb.137513.
- Simonyan, K., Jürgens, U., 2003. Efferent subcortical projections of the laryngeal motorcortex in the rhesus monkey. *Brain Res* 974, 43–59. doi:10.1016/S0006-8993(03)02548-4.
- Simonyan, K., Ostuni, J., Ludlow, C.L., Horwitz, B., 2009. Functional but not structural networks of the human laryngeal motor cortex show left hemispheric lateralization during syllable but not breathing production. *J. Neurosci.* 29, 14912–14923. doi:10.1523/JNEUROSCI.4897-09.2009.
- Simonyan, K., Saad, Z.S., Loucks, T.M.J., Poletto, C.J., Ludlow, C.L., 2007. Functional neuroanatomy of human voluntary cough and sniff production. *Neuroimage* 37, 401–409. doi:10.1016/j.neuroimage.2007.05.021.
- Smith, S.M., Jenkinson, M., Woolrich, M.W., Beckmann, C.F., Behrens, T.E.J., Johansen-Berg, H., Bannister, P.R., De Luca, M., Drobnjak, I., Flitney, D.E., Niazzy, R.K., Saunders, J., Vickers, J., Zhang, Y., De Stefano, N., Brady, J.M., Matthews, P.M., 2004. Advances in functional and structural MR image analysis and implementation as FSL. *Neuroimage* 23, 208–219. doi:10.1016/j.neuroimage.2004.07.051.
- Stansbury, A.L., Janik, V.M., 2019. Formant modification through vocal production learning in Gray Seals. *Curr. Biol.* 29, 1–6. doi:10.1016/j.cub.2019.05.071.
- Stippich, C., Ochmann, H., Sartor, K., 2002. Somatotopic mapping of the human primary sensorimotor cortex during motor imagery and motor execution by functional magnetic resonance imaging. *Neurosci. Lett.* 331, 50–54. doi:10.1016/S0304-3940(02)00826-1.
- Stoeger, A.S., Mietchen, D., Oh, S., de Silva, S., Herbst, C.T., Kwon, S., Fitch, W.T., 2012. An Asian elephant imitates human speech. *Curr. Biol.* 22, 2144–2148. doi:10.1016/j.cub.2012.09.022.
- Story, B.H., Titze, I.R., 1995. Voice simulation with a body-cover model of the vocal folds. *J. Acoust. Soc. Am.* doi:10.1121/1.412234.
- Subramanian, H.H., Holstege, G., 2009. The nucleus retroambiguus control of respiration. *J. Neurosci.* 29, 3824–3832. doi:10.1523/JNEUROSCI.0607-09.2009.
- Suthers, R.A., Thomas, S.P., Suthers, B.J., 1972. Respiration, wing-beat and ultrasonic pulse emission in an echo-locating bat. *J. Exp. Biol.* 56, 37–48.
- Sutton, D., Larson, C., Lindeman, R.C., 1974. Neocortical and limbic lesion effects on primate phonation. *Brain Res* 71, 61–75.
- Sutton, D., Trachy, R.E., Lindeman, R.C., 1981. Primate phonation: unilateral and bilateral cingulate lesion effects. *Behav. Brain Res.* 3, 99–114.
- Takai, O., Brown, S., Liotti, M., 2010. Representation of the speech effectors in the human motor cortex: Somatotopy or overlap? *Brain Lang* 113, 39–44. doi:10.1016/j.bandl.2010.01.008.
- Terumitsu, M., Fujii, Y., Suzuki, K., Kwee, I.L., Nakada, T., 2006. Human primary motor cortex shows hemispheric specialization for speech. *Neuroreport* 17, 1091–1095. doi:10.1097/01.wnr.0000224778.97399.c4.
- THEYS, C., van Wieringen, A., Sunaert, S., Thijs, V., De Nil, L.F., 2011. A one year prospective study of neurogenic stuttering following stroke: Incidence and co-occurring disorders. *J. Commun. Disord.* 44, 678–687. doi:10.1016/j.jcomdis.2011.06.001.
- Thoms, G., Jürgens, U., 1987. Common input of the cranial motor nuclei involved in phonation in squirrel monkey. *Exp. Neurol.* 95, 85–99. doi:10.1016/0014-4886(87)90009-4.
- Titze, I.R., 2008. Nonlinear source-filter coupling in phonation: theory. *J. Acoust. Soc. Am.* 123, 2733. doi:10.1121/1.2832337.
- Titze, I.R., 1989. On the relation between subglottal pressure and fundamental frequency in phonation. *J. Acoust. Soc. Am.* 85, 901–906. doi:10.1121/1.397562.
- Titze, I.R., Luschei, E.S., Hirano, M., 1989. Role of the thyroarytenoid muscle in regulation of fundamental frequency. *J. Voice* 3, 213–224. doi:10.1016/S0892-1997(89)80003-7.
- Titze, I.R., Story, B.H., 2002. Rules for controlling low-dimensional vocal fold models with muscle activation. *J. Acoust. Soc. Am.* 112, 1064–1076. doi:10.1121/1.1496080.
- VanderHorst, V.G.J.M., Terasawa, E., Ralston, H.J., 2001. Monosynaptic projections from the nucleus retroambiguus region to laryngeal motoneurons in the rhesus monkey. *Neuroscience* 107, 117–125. doi:10.1016/S0306-4522(01)00343-8.
- Vanderhorst, V.G.J.M., Terasawa, E., Ralston, H.J., Holstege, G., 2000. Monosynaptic projections from the nucleus retroambiguus to motoneurons supplying the abdominal wall, axial, hindlimb, and pelvic floor muscles in the female rhesus monkey. *J. Comp. Neurol.* 424, 233–250. doi:10.1002/1096-9861(20000821)424:2<233::AID-CNEA>3.0.CO;2-C.
- Verhaegen, M., Puech, P.F., Munro, S., 2002. Aquaroreal ancestors? *Trends Ecol. Evol.* 17, 212–217. doi:10.1016/S0169-5347(02)02490-4.
- Vernes, S., Wilkinson, G., 2019. Behaviour, biology, and evolution of vocal learning in bats. *Philos. Trans. R. Soc. Lond. B Biol. Sci.* 375, 20190061. doi:10.1101/646703.
- Vernes, S.C., 2016. What bats have to say about speech and language. *Psychon. Bull. Rev.* 1–7. doi:10.3758/s13423-016-1060-3.
- Vilkman, E., Sonninen, A., Hurme, P., Körkkö, P., 1996. External laryngeal frame function in voice production revisited: a review. *J. Voice* 10, 78–92. doi:10.1016/S0892-1997(96)80021-X.
- Vogt, O., 1910. Die myeloarchitektonische Felderung des menschlichen Stirnhirns. *J. für Psychol. und Neurol.* 15, 221–232.
- Wattendorf, E., Westermann, B., Fiedler, K., Kaza, E., Lotze, M., Celio, M.R., 2013. Exploration of the neural correlates of ticklish laughter by functional magnetic resonance imaging. *Cereb. Cortex* 23, 1280–1289. doi:10.1093/cercor/bhs0.
- West, J.B., 2001. Snorkel breathing in the elephant explains the unique anatomy of its pleura. *Respir. Physiol.* 126, 1–8. doi:10.1016/S0034-5687(01)00203-1.
- Wich, S.A., Krützen, M., Lameira, A.R., Nater, A., Arora, N., Bastian, M.L., Meulman, E., Morrogh-Bernard, H.C., Utami Atmoko, S.S., Pamungkas, J., Perwitasari-Farajallah, D., Hardus, M.E., van Noordwijk, M., van Schaik, C.P., 2012. Call cultures in orang-utans? *PLoS One* 7, 1–9. doi:10.1371/journal.pone.0036180.
- Wich, S.A., Swartz, K.B., Hardus, M.E., Lameira, A.R., Stromberg, E., Shumaker, R.W., 2009. A case of spontaneous acquisition of a human sound by an orangutan. *Primates* 50, 56–64. doi:10.1007/s10329-008-0117-y.
- Wild, J.M., 1993. Descending projections of the songbird nucleus robustus archistriatalis. *J. Comp. Neurol.* 338, 225–241. doi:10.1002/cne.903380207.
- Wild, J.M., Williams, M.N., Suthers, R.A., 2000. Neural pathways for bilateral vocal control in songbirds. *J. Comp. Neurol.* 426, 413–426.
- Worseley, K., Friston, K., 1995. Analysis of fMRI time-series revisited-again. *Neuroimage* 2, 173–181.