



HHS Public Access

Author manuscript

Cortex. Author manuscript; available in PMC 2018 April 01.

Published in final edited form as:

Cortex. 2017 April ; 89: 111–119. doi:10.1016/j.cortex.2017.01.024.

Paradoxical vocal changes in a trained singer by focally cooling the right superior temporal gyrus

Kalman A. Katlowitz^{1,2,3}, Hiroyuki Oya⁴, Matthew A. Howard III⁴, Jeremy D.W. Greenlee⁴, and Michael A. Long^{1,2,3,Ψ}

¹NYU Neuroscience Institute, New York University Langone Medical Center, New York, NY 10016 USA

²Department of Otolaryngology, New York University Langone Medical Center, New York, NY 10016 USA

³Center for Neural Science, New York University, New York, NY, 10003 USA

⁴Human Brain Research Lab, Department of Neurosurgery, University of Iowa, Iowa City, IA, 52242 USA

Abstract

The production and perception of music is preferentially mediated by cortical areas within the right hemisphere, but little is known about how these brain regions individually contribute to this process. In an experienced singer undergoing awake craniotomy, we demonstrated that direct electrical stimulation to a portion of the right posterior superior temporal gyrus (pSTG) selectively interrupted singing but not speaking. We then focally cooled this region to modulate its activity during vocalization. In contrast to similar manipulations in left hemisphere speech production regions, pSTG cooling did not elicit any changes in vocal timing or quality. However, this manipulation led to an increase in the pitch of speaking with no such change in singing. Further analysis revealed that all vocalizations exhibited a cooling-induced increase in the frequency of the first formant, raising the possibility that potential pitch offsets may have been actively avoided during singing. Our results suggest that the right pSTG plays a key role in vocal sensorimotor processing whose impact is dependent on the type of vocalization produced.

Keywords

vocal production; electrical stimulation mapping; temperature; sensorimotor

ΨTo whom correspondence should be addressed: mlong@med.nyu.edu.

Author Contributions

M.A.L. and J.D.W.G. designed the research. K.K. and M.A.L. analyzed the data and wrote the paper. H.O., J.D.W.G. and M.H. assisted with the anatomical analysis and manuscript preparation.

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

1. Introduction

Despite the similarities inherent in the production of speech and song (Ozdemir et al., 2006; Patel, 2003), several lines of evidence suggest that these behaviors may be mediated in part by distinct cortical regions (Rogalsky et al., 2011; Zatorre et al., 2007). Retrospective lesion studies (Benton, 1977; Terao et al., 2006), transcranial magnetic stimulation (Epstein et al., 1999), and Wada testing (Gordon and Bogen, 1974) support a primary role for the right hemisphere in musical perception and production (Jeffries et al., 2003; Zatorre et al., 2002). Separate computational modules on the right hemisphere are thought to be involved in these behaviors (Norman-Haignere et al., 2015; Peretz and Coltheart, 2003). One of the areas in this network, the posterior superior temporal gyrus (pSTG), appears to participate as an anatomical (Garell et al., 2013) and functional (Perry et al., 1999; Tourville et al., 2008) interface between vocal perception and production. Although a variety of acoustic stimuli (Geiser et al., 2008; Poeppel et al., 2004; Zatorre and Belin, 2001) as well as covert vocalizations (Hickok, 2012; Kleber et al., 2007) can drive activity within the right pSTG, the contribution of this region to specific vocal behaviors remains poorly understood (Jeffries et al., 2003; Perry et al., 1999; Suarez et al., 2010).

In this case study, we investigated the role of the right pSTG in vocal production in a professional singer. The subject, a tenor, holds a Bachelor's degree in Voice and has performed prolifically across the United States as well as internationally over the last 25 years. We first characterized the pSTG as a site in which electrical stimulation resulted in an impairment of singing while speech was unperturbed, which is consistent with a previous observation (Suarez et al., 2010). We then used a recently validated technique, focal cortical cooling (Bakken et al., 2003; Long et al., 2016), to manipulate the pSTG during speaking and singing. Because of previous observations suggesting a specialized role for the right pSTG in rhythm (Geiser et al., 2008), we postulated that focal cooling of this region may change the timing of singing behavior, which was not the case. Instead, cooling modulated spectral properties of vocalizations in a task-dependent manner. Taken together, our stimulation and cooling results demonstrate differential roles of the right pSTG in singing and speaking.

2. Methods

2.1 Subject

The procedures undertaken as part of this study have been approved by the University of Iowa Institutional Review Board for research purposes. The subject was a patient volunteer undergoing awake craniotomy who had provided informed consent in advance of surgery to participate in this study, which did not add additional surgical risk. Cerebral dominance for speech was determined by preoperative Wada testing (Wada and Rasmussen, 1960). Conversational speech and hearing for the subject were normal prior to and during surgery, such that he could easily complete the tasks necessary for this study. Local anesthesia was administered, and intermittent intravenous sedation was used except during the research portion of the surgery.

2.2 Behavioral tasks

The subject was asked to perform specific speaking or singing tasks in response to a prompt. The two speaking tasks were either ‘Counting’ (‘21’, ‘22’, ‘23’, ‘24’, ‘25’) or reciting the ‘Days’ of the week (‘Monday’, ‘Tuesday’, ‘Wednesday’, ‘Thursday’, ‘Friday’). The first singing task (‘Scales’) involved the production of a C major scale using the word ‘La’ such that ascending and descending scales were each produced twice in succession for each trial. The second singing task (‘My Bonnie’) involved singing the first verse of the traditional Scottish folk song ‘My Bonnie Lies Over the Ocean’ in its entirety (‘My Bonnie lies over the ocean/My Bonnie lies over the sea/My Bonnie lies over the ocean/Oh bring back my Bonnie to me’). Each task was performed multiple times during cooling and electrical stimulation, and a number of control trials were collected during which time no manipulations were performed. The subject’s voice was recorded with a high-performance microphone (Beta 87C, Shure, Niles, IL) as well as an accessory Bluetooth microphone placed near the subject’s mouth. Sound files were denoised offline via spectral noise gating (Audacity, <http://audacityteam.org/>; Reduction of 24dB, Sensitivity 0 dB, Freq Smoothing 150 Hz, Attack/Decay time 0.15 s).

2.3 Electrical stimulation and cooling

Electrical stimulation mapping (ESM) was performed to localize cortical sites involved in vocal production (Chang et al., 2016; Ojemann et al., 1989; Penfield and Boldrey, 1937). ESM consisted of 50 Hz trains of 0.2 ms biphasic, charge-balanced pulses (5–20 V, Grass SD-9, Natus Neurology, Inc., Warwick, RI) delivered directly to the cortical surface using custom built hand-held bipolar silver ball tip electrodes spaced ~5 mm apart. Stimulus intensities began at low levels (~5V) and were gradually increased until either the afterdischarge threshold or a clinical effect was reached. Continuous electrocorticography (ECoG) was obtained using strip electrode arrays adjacent to stimulated sites to rapidly and effectively detect any afterdischarges induced by stimulation. An epileptologist continually monitored the multi-channel ECoG to confirm the absence of afterdischarges at each intensity prior to brain mapping at that setting. The anesthesia team was equipped with anti-convulsant medications (e.g. levetiracetam, phenytoin), sedatives and airway adjuncts (e.g. face mask, nasal airway, laryngeal mask airway), if needed.

The cooling probe, described in a previous report (Long et al., 2016), was analyzed by the Hospital Biomedical Engineering Department at the University of Iowa and deemed safe for its intended purpose. The probe was a custom-built titanium chamber with a 1 x 1 cm square Peltier element (Custom Thermoelectric, Bishopville, MD) attached to the bottom. The brain surface side of the element was coated with a platinum film for biocompatibility. Circulating sterile chilled saline in the chamber was used to dissipate heat generated on the non-brain side of the Peltier element. Current to the device was supplied using a regulated power supply (HY1803D, Tekpower, Taiwan). The temperature of the device was continuously measured and displayed using a thermocouple (K-type, OMEGA Engineering, Inc., Stamford, CT) that was installed between the two ceramic plates housing the Peltier elements. Temperature data were recorded on a multi-channel data acquisition system (System3, Tucker-Davis Technologies, Alachua, FL). The cortical temperature change at 4 mm was modeled using a previously described procedure (Long et al., 2016). In that

calibration, we showed that temperature changes were highly depth-dependent and that the behavioral effects observed did not originate from a modulation of the contralateral hemisphere. Both the electrical stimulation sites as well as the cooling probe location were documented with multiple high resolution photographs as well as a continuous video stream of the brain surface throughout the entire procedure. An expert neuroanatomist (H. Oya) assisted in registering these locations to the patient's presurgical structural MRI.

2.4 Extraction of vocal features

We analyzed a number of features of the recorded vocalizations. First, to quantify changes in vocal timing, we manually demarcated the locations of word onsets and offsets for each sequence using custom Matlab software. This process was performed by two different experienced researchers to avoid error. Second, vocal 'quality' of each trial was evaluated using a visual analog scale (VAS) between 0 and 1 by many different anonymous raters using a previously described crowd-sourcing platform (Becker and Levine, 2013; Long et al., 2016; McAllister Byun et al., 2015). To avoid outliers, the vocal 'quality' for each trial was then defined as the median score from all raters. Third, we extracted the pitch (defined as the fundamental frequency) and the first two formants (defined as the first two relative maxima in the spectrum) of identified vowels in both singing and speaking tasks. These features were calculated using Praat via VoiceSauce (Boersma and van Heuven, 2001; Shue et al., 2009). For pitch and formants, vowels were identified as segments relative to the manual demarcations described above (minimum size = 30 ms). Timepoints in which the traces exhibited large deviations (20, 50, and 100 Hz for pitch, first formant, and second formant, respectively) were deemed artifactual and culled from the dataset. We eliminated cases in which more than half of the trace was eliminated following this process.

2.5 Analysis of vocal features

For these analyses and all others, data were presented as mean \pm standard deviation unless otherwise noted. For timing, we calculated a z-score for the durations of each word to aggregate data within each task. In pitch analysis, all instances of the vowel were linearly warped to the same length. Both pitch and formants were defined as the median value across the vowel. The mean across all instances was used to define the location of the vowel in formant space for the control and cooled conditions. Amplitude was defined as the root mean squared (RMS) value of the raw audio waveform for each word. In all described analyses, we looked at the differences in the cooled trials relative to the controls. Additionally, to test for absolute pitch errors and latency, each note in the vocal scales was compared with its respective target note. Accuracy was defined as the percentage of the note during the vowel in 'La' in which the pitch was closer to the target note than any alternative, and latency was defined as the duration needed to reach an accurate note from the beginning of 'La'. We used the Benjamini-Hochberg-Yekutieli procedure to control for type 1 errors (Benjamini and Yekutieli, 2001).

3. Results

3.1 Electrical Stimulation Mapping

The subject is a right-handed male with extensive singing experience who was 41 years old at the time of testing. The surgical procedure was intended to treat medically intractable epilepsy with an epileptic locus lateralizing to the mesial right temporal lobe. Before surgery, left-hemispheric dominance for speech was demonstrated using the Wada test; singing behavior was not measured. During the awake craniotomy procedure, electrical stimulation was applied over a large portion of the frontal and temporal lobes on the right hemisphere to functionally map the exposed brain surface (Chang et al., 2016; Ojemann et al., 1989; Penfield and Boldrey, 1937). In many cases, nearby sites were stimulated multiple times at different points during the ESM procedure, and these produced consistent results. In total, 58 unique locations were individually stimulated during singing, and 7 resulted in vocal arrest. In 4 of these sites, located along the right precentral gyrus, stimulation caused dysarthria which interrupted both singing and speaking tasks (Fig. 1A and B). In the 3 sites clustered within the pSTG (MNI coordinates 65.2, -31.5, 6.3), stimulation selectively blocked singing behavior (Fig. 1). Stimulation interrupted two distinct song types: A traditional tune ('My Bonnie', e.g., Fig. 1C) as well as simple vocal scales ('Scales'). Throughout stimulation of the pSTG, the patient retained his ability to speak without interruption (e.g., Fig. 1D) and could voluntarily move his articulator muscles in a nonverbal task.

3.2 Focal cooling of pSTG preserves vocal behavior

After characterizing a region within the pSTG as a site of selective song arrest, a small cooling probe (1cm X 1cm) was placed on that location (Fig. 2A). Based on our models of the impact of cortical cooling, the estimated extent of cooling at a depth of 4 mm was $5.5 \pm 0.5^\circ\text{C}$ for a total duration of 6.5 minutes (Long et al., 2016). The subject sang both 'Scales' and 'My Bonnie' (5 control and 4 cooled for each) and recited the days of the week ('Days') as well as a list of numbers ('Numbers') (10 control and 14 cooled for each) (Fig. 2B). Tasks were prompted by the name of the trial (e.g. 'My Bonnie') without providing a reference pitch. Additionally, the subject's position was unaltered throughout the experiment, as verified by video recording, excluding the possibility that postural changes could account for vocal shifts. The subject stated that singing felt 'pretty normal' at the beginning of cooling and afterwards confirmed that his vocal production seemed 'fine' throughout. Consistent with this, both speaking (Fig. 2C) and singing (Fig. 2D) were grossly unaffected by cooling, as measured through a crowdsourced subjective rating approach that obtained quality scores from 20 different listeners for each trial (Long et al., 2016; McAllister Byun et al., 2015) ($p > 0.5$, Wilcoxon rank sum test on median rating of each trial). We also found no significant change in the mean RMS amplitude of each word in either vocalization type during cooling ($p > 0.05$, Wilcoxon signed rank test, $n = 54$ words singing and 10 speaking). We also examined the relative duration of words (e.g., Fig. 2E) for both singing and speaking, but we were unable to find a significant cooling-induced change relative to controls ($p > 0.05$, two-sample t-test) across all tasks.

3.3 pSTG cooling and spectral properties of vocalizations

We next looked for any cooling-related changes in the spectral properties of vocalizations. We examined the notes produced during the ‘Scales’ task (Fig. 3A), comparing the fundamental frequency of the cooled vocalizations to controls (Fig. 3B and C). Across a population of 29 identified vowel sounds (Fig. 3D), there was no difference in the mean pitch of the vowels produced during cooling relative to that of controls ($p>0.05$, paired t-test) (Fig. 3E, Supplementary Audio 1 and 2). We also found no difference in the mean latency to note target from the previous note ($p>0.05$, Wilcoxon signed rank test, $n=28$ latencies measured) or in the mean accuracy of each note ($p>0.05$, Wilcoxon signed rank test). Similarly, cooling did not change the pitch of individual notes in the ‘My Bonnie’ task ($p>0.05$, paired t-test, 37 identified vowel sounds) (Fig. 3D and E). We then tested whether cooling could have an impact on the pitch of speech (Fig. 3F–J), focusing on 10 vowels from each condition (‘Days’ and ‘Counting’). Surprisingly, vowels that were part of spoken words exhibited audible changes in their fundamental frequencies, with ‘Days’ increasing by 4.0 ± 1.5 Hz (Supplementary Audio 3 and 4) and ‘Counting’ increasing by 5.5 ± 2.2 Hz (paired t-test, $p<0.001$) (Fig. 3I and J).

In addition to pitch, we examined formant frequencies, which are determined by the resonances of the vocal tract (Baer et al., 1991; Fant, 1960) and have a crucial role in establishing vowel identity (Bladon and Lindblom, 1981; Peterson and Barney, 1952). We measured the first and second formants for all singing (Fig. 4A–E) and speaking (Fig. 4F–J) tasks. The first formant increased with cooling (‘Scales’: 21 ± 19 Hz; ‘My Bonnie’: 12 ± 24 Hz; ‘Days’: 14 ± 7 Hz; ‘Counting’: 14 ± 6 Hz; $p<0.005$, paired t-test), though the magnitude of change was insufficient to result in different vowel identities. Additionally, the second formant increased significantly in the ‘Scales’ task by 29 ± 44 Hz (paired t-test, $p<0.01$) (Fig. 4B and D). All parameters that changed due to cooling returned to baseline with the rewarming of the cortex.

4. Discussion

In this case study, we used two distinct methods to manipulate the pSTG on the right hemisphere during vocal production. First, we found that applying electrical stimulation directly to the right pSTG, and not the surrounding tissue, led to an interruption of singing without affecting speech. This observation lent some clarity to a previous study in which results were inconsistent across individuals, potentially due to the fact that far fewer sites were interrogated in each subject (Suarez et al., 2010). Second, we found that focal cooling of the pSTG affected vowel pitch and formant frequencies without gross changes in vocal timing or quality.

The two manipulations used in this study, electrical stimulation and cooling, have different effects on brain function. Although electrical stimulation is a commonly accepted neurosurgical technique for localizing function (Chang et al., 2016; Hamberger, 2007; Ojemann et al., 1989) in order to prevent postoperative speech deficits (Haglund et al., 1994; Ojemann and Dodrill, 1985), the large-scale influences of this stimulation are complex and poorly understood (Histed et al., 2009; Tehovnik, 1996). For instance, applied current not only excites the tissue underneath the electrode, but it also can lead to the activation of

upstream and downstream neurons (Lipski, 1981; Ranck, 1975), resulting in problems in interpretation concerning the role of an individual brain region. Additionally, the complete elimination of behavior following stimulation-induced arrest precludes further investigation into the role of that region. We therefore also used cooling to experimentally manipulate (Long and Fee, 2008; Long et al., 2016; Pires and Hoy, 1992; Tang et al., 2010) pSTG function due to the highly temperature sensitive nature of neural processing (Sabatini and Regehr, 1996; Volgushev et al., 2000). The mild temperature changes used in our study should have primarily affected the area under the cooling probe while preserving general function (Aronov and Fee, 2011; Long and Fee, 2008; Long et al., 2016). Cooling-induced effects on the acoustic properties of vocalizations supports the notion that the right pSTG may impact the frequencies of both pitch (Chang et al., 2013) and formants (Niziolek and Guenther, 2013; Tourville et al., 2008).

Perhaps the most puzzling result is the fact that right pSTG cooling led to pitch changes in speaking but not singing. One hypothesis is that cooling the pSTG changes vocalizations by altering the drive to vocal premotor structures (Garell et al., 2013; Glasser and Rilling, 2008). However, because singing (unlike speech) has clearly defined targets – a specific ‘correct’ note within a scale or a song - the patient may have been able to use sensory feedback to compensate for any potential erroneous pitch shifts (Borden, 1979; Kelso et al., 1984; Todorov and Jordan, 2002). In agreement with this idea, acoustic properties that were not directly related to task success, such as formant frequencies, were also affected by our manipulation. An alternate hypothesis is that cooling the pSTG biased sensory feedback, potentially leading to a perceived change in fundamental frequency and formants and thus prompting a behavioral response. Similarly, artificially pitch-shifted auditory feedback can lead to a modulation of the fundamental frequency of vocalizations in songbirds (Sober and Brainard, 2009), monkeys (Eliades and Wang, 2008) and humans (Burnett et al., 1998; Elman, 1981; Fairbanks, 1954). In this case, the subject’s extensive vocal training may have allowed him to rely on an internal model during singing, independent of any contradicting feedback (Kleber et al., 2013; Zarate and Zatorre, 2005; Zarate and Zatorre, 2008) and thus avoiding the resultant pitch change.

We have decided to publish this single case due to the rare opportunity to focally and precisely manipulate the right temporal gyri in an expert singer. However, further work is necessary to both replicate these findings as well as test the aforementioned hypotheses. For example, sensory discrimination tasks could be carried out during pSTG cooling. If this manipulation leads to a shift in the perceived frequency, then it is likely that the pSTG is playing a sensory role. Additionally, this case study was performed in an individual with extensive singing experience, and the impact of this training on our behavioral results can be further explored by comparing outcomes from individuals with different levels of expertise (Kleber et al., 2010; Zarate and Zatorre, 2008). Together with our current findings, these observations will help further define the role of the right sided pSTG in functional models of vocal production (Guenther et al., 2006; Price et al., 2011; Zarate, 2013) and perception (Hickok and Poeppel, 2007).

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank the patient volunteer for his contribution to the study as well as the team of neurosurgical residents at the University of Iowa Hospitals and Clinics for their assistance with data collection. Haiming Chen and Aimee Chow helped with data analysis. Daniel Szeredi and Tara McAllister Byun helped in the implementation of our crowdsourcing procedure. We acknowledge valuable discussions with Eli Merriam and Mario Svirsky. We thank Lucia Melloni, Julia King, Daniel Bendor, and Mario Svirsky for their comments on an earlier version of this manuscript. This work was supported by the NIH (DC015260, DC004290, DC009589, NS075044, TR001447) as well as the NYSC Foundation.

References

- Aronov D, Fee MS. Analyzing the dynamics of brain circuits with temperature: design and implementation of a miniature thermoelectric device. *J Neurosci Methods*. 2011; 197:32–47. [PubMed: 21291909]
- Baer T, Gore JC, Gracco LC, Nye PW. Analysis of Vocal-Tract Shape and Dimensions Using Magnetic-Resonance-Imaging - Vowels. *J Acoust Soc Am*. 1991; 90:799–828. [PubMed: 1939886]
- Bakken H, Kawasaki H, Oya H, Greenlee JDW, Howard Ma. A device for cooling localized regions of human cerebral cortex. *Journal of neurosurgery*. 2003; 99:604–608. [PubMed: 12959453]
- Becker, M., Levine, J. Experigen – an online experiment platform. 2013.
- Benjamini Y, Yekutieli D. The control of the false discovery rate in multiple testing under dependency. *Ann Stat*. 2001; 29:1165–1188.
- Benton, A. The Amusias. In: Critchley, M., Henson, RA., editors. *Music and the Brain: Studies in the Neurology of Music*. Butterworth-Heinemann; 1977. p. 378-396.
- Bladon RAW, Lindblom B. Modeling the Judgment of Vowel Quality Differences. *J Acoust Soc Am*. 1981; 69:1414–1422. [PubMed: 7240572]
- Boersma P, van Heuven V. Speak and unSpeak with Praat. *Glott International*. 2001; 5:341–347.
- Borden GJ. An interpretation of research of feedback interruption in speech. *Brain and language*. 1979; 7:307–319. [PubMed: 455050]
- Burnett, Ta, Freedland, MB., Larson, CR., Hain, TC. Voice F0 responses to manipulations in pitch feedback. *The Journal of the Acoustical Society of America*. 1998; 103:3153–3161. [PubMed: 9637026]
- Chang EF, Breshears JD, Raygor KP, Lau D, Molinaro AM, Berger MS. Stereotactic probability and variability of speech arrest and anomia sites during stimulation mapping of the language dominant hemisphere. *J Neurosurg*. 2016:1–8.
- Chang EF, Niziolek CA, Knight RT, Nagarajan SS, Houde JF. Human cortical sensorimotor network underlying feedback control of vocal pitch. *Proc Natl Acad Sci U S A*. 2013; 110:2653–2658. [PubMed: 23345447]
- Eliades SJ, Wang X. Neural substrates of vocalization feedback monitoring in primate auditory cortex. *Nature*. 2008; 453:1102–1106. [PubMed: 18454135]
- Elman JL. Effects of Frequency-Shifted Feedback on the Pitch of Vocal Productions. *J Acoust Soc Am*. 1981; 70:45–50. [PubMed: 7264071]
- Epstein CM, Meador KJ, Loring DW, Wright RJ, Weissman JD, Sheppard S, Lah JJ, Puhlovich F, Gaitan L, Davey KR. Localization and characterization of speech arrest during transcranial magnetic stimulation. *Clin Neurophysiol*. 1999; 110:1073–1079. [PubMed: 10402094]
- Fairbanks G. Systematic research in experimental phonetics. I. A theory of the speech mechanism as a servosystem. *J Speech Hear Disord*. 1954; 19:133–139. [PubMed: 13212816]
- Fant, G. *Acoustic theory of speech production*. s'Gravenhage; Mouton: 1960.

- Garell PC, Bakken H, Greenlee JDW, Volkov I, Reale Ra, Oya H, Kawasaki H, Howard Ma, Brugge JF. Functional Connection Between Posterior Superior Temporal Gyrus and Ventrolateral Prefrontal Cortex in Human. *Cerebral Cortex*. 2013; 23:2309–2321. [PubMed: 22879355]
- Geiser E, Zaehle T, Jancke L, Meyer M. The neural correlate of speech rhythm as evidenced by metrical speech processing. *Journal of cognitive neuroscience*. 2008; 20:541–552. [PubMed: 18004944]
- Glasser MF, Rilling JK. DTI tractography of the human brain's language pathways. *Cereb Cortex*. 2008; 18:2471–2482. [PubMed: 18281301]
- Gordon HW, Bogen JE. Hemispheric lateralization of singing after intracarotid sodium amylobarbitone. *Journal of Neurology, Neurosurgery & Psychiatry*. 1974; 37:727–738.
- Guenther FH, Ghosh SS, Tourville JA. Neural modeling and imaging of the cortical interactions underlying syllable production. *Brain Lang*. 2006; 96:280–301. [PubMed: 16040108]
- Haglund MM, Berger MS, Shamseldin M, Lettich E, Ojemann Ga. Cortical Localization of Temporal Lobe Language Sites in Patients with Gliomas. *Neurosurgery*. 1994; 34:567–576. [PubMed: 7516498]
- Hamberger MJ. Cortical Language Mapping in Epilepsy: A Critical Review. *Neuropsychology Review*. 2007; 17:477–489. [PubMed: 18004662]
- Hickok G. Computational neuroanatomy of speech production. *Nat Rev Neurosci*. 2012; 13:135–145. [PubMed: 22218206]
- Hickok G, Poeppel D. The cortical organization of speech processing. *Nature Reviews Neuroscience*. 2007; 8:393–402. [PubMed: 17431404]
- Histed MH, Bonin V, Reid RC. Direct activation of sparse, distributed populations of cortical neurons by electrical microstimulation. *Neuron*. 2009; 63:508–522. [PubMed: 19709632]
- Jeffries KJ, Fritz JB, Braun aR. Words in melody: an H(2)15O PET study of brain activation during singing and speaking. *Neuroreport*. 2003; 14:749–754. [PubMed: 12692476]
- Kelso JA, Tuller B, Vatikiotis-Bateson E, Fowler CA. Functionally specific articulatory cooperation following jaw perturbations during speech: evidence for coordinative structures. *J Exp Psychol Hum Percept Perform*. 1984; 10:812–832. [PubMed: 6239907]
- Kleber B, Birbaumer N, Veit R, Trevorrow T, Lotze M. Overt and imagined singing of an Italian aria. *Neuroimage*. 2007; 36:889–900. [PubMed: 17478107]
- Kleber B, Veit R, Birbaumer N, Gruzelier J, Lotze M. The brain of opera singers: experience-dependent changes in functional activation. *Cereb Cortex*. 2010; 20:1144–1152. [PubMed: 19692631]
- Kleber B, Zeitouni AG, Friberg A, Zatorre RJ. Experience-dependent modulation of feedback integration during singing: role of the right anterior insula. *J Neurosci*. 2013; 33:6070–6080. [PubMed: 23554488]
- Lipski J. Antidromic Activation of Neurons as an Analytic Tool in the Study of the Central Nervous-System. *J Neurosci Meth*. 1981; 4:1–32.
- Long MA, Fee MS. Using temperature to analyse temporal dynamics in the songbird motor pathway. *Nature*. 2008; 456:189–194. [PubMed: 19005546]
- Long MA, Katlowitz KA, Svirsky MA, Clary RC, Byun TM, Majaj N, Oya H, Howard MA, Greenlee JDW. Functional Segregation of Cortical Regions Underlying Speech Timing and Articulation. *Neuron*. 2016; 89:1187–1193. [PubMed: 26924439]
- McAllister Byun T, Halpin PF, Szeredi D. Online crowdsourcing for efficient rating of speech: a validation study. *J Commun Disord*. 2015; 53:70–83. [PubMed: 25578293]
- Niziolek CA, Guenther FH. Vowel category boundaries enhance cortical and behavioral responses to speech feedback alterations. *J Neurosci*. 2013; 33:12090–12098. [PubMed: 23864694]
- Norman-Haignere S, Kanwisher NG, McDermott JH. Distinct Cortical Pathways for Music and Speech Revealed by Hypothesis-Free Voxel Decomposition. *Neuron*. 2015; 88:1281–1296. [PubMed: 26687225]
- Ojemann G, Ojemann J, Lettich E, Berger M. Cortical language localization in left, dominant hemisphere. *Journal of Neurosurgery*. 1989; 71:316–326. [PubMed: 2769383]

- Ojemann, Ga, Dodrill, CB. Verbal memory deficits after left temporal lobectomy for epilepsy. Mechanism and intraoperative prediction. *Journal of neurosurgery*. 1985; 62:101–107. [PubMed: 3964840]
- Ozdemir E, Norton A, Schlaug G. Shared and distinct neural correlates of singing and speaking. *NeuroImage*. 2006; 33:628–635. [PubMed: 16956772]
- Patel AD. Language, music, syntax and the brain. *Nat Neurosci*. 2003; 6:674–681. [PubMed: 12830158]
- Penfield W, Boldrey E. Somatic motor and sensory representation in the cerebral cortex of man as studied by electrical stimulation. *Brain*. 1937; 60:389–443.
- Peretz I, Coltheart M. Modularity of music processing. *Nature Neuroscience*. 2003; 6:688–691. [PubMed: 12830160]
- Perry DW, Zatorre RJ, Petrides M, Alivisatos B, Meyer E, Evans aC. Localization of cerebral activity during simple singing. *Neuroreport*. 1999; 10:3979–3984. [PubMed: 10716244]
- Peterson GE, Barney HL. Control Methods Used in a Study of the Vowels. *J Acoust Soc Am*. 1952; 24:175–184.
- Pires A, Hoy RR. Temperature coupling in cricket acoustic communication. *Journal of Comparative Physiology A*. 1992; 171:79–92.
- Poeppel D, Guillemin A, Thompson J, Fritz J, Bavelier D, Braun AR. Auditory lexical decision, categorical perception, and FM direction discrimination differentially engage left and right auditory cortex. *Neuropsychologia*. 2004; 42:183–200. [PubMed: 14644105]
- Price CJ, Crinion JT, MacSweeney M. A generative model of speech production in Broca's and Wernicke's areas. *Frontiers in Psychology*. 2011; 2:1–9. [PubMed: 21713130]
- Ranck JB. Which elements are excited in electrical stimulation of mammalian central nervous system: A review. *Brain Research*. 1975:417–440. [PubMed: 1102064]
- Rogalsky C, Rong F, Saberi K, Hickok G. Functional anatomy of language and music perception: temporal and structural factors investigated using functional magnetic resonance imaging. *J Neurosci*. 2011; 31:3843–3852. [PubMed: 21389239]
- Sabatini BL, Regehr WG. Timing of neurotransmission at fast synapses in the mammalian brain. *Nature*. 1996; 384:170–172. [PubMed: 8906792]
- Shue YL, Keating P, Vicenik C. VOICESAUCE: A program for voice analysis. *The Journal of the Acoustical Society of America*. 2009; 126:2221.
- Sober SJ, Brainard MS. Adult birdsong is actively maintained by error correction. *Nat Neurosci*. 2009; 12:927–931. [PubMed: 19525945]
- Suarez RO, Golby A, Whalen S, Sato S, Theodore WH, Kufta CV, Devinsky O, Balish M, Bromfield EB. Contributions to singing ability by the posterior portion of the superior temporal gyrus of the non-language-dominant hemisphere: First evidence from subdural cortical stimulation, Wada testing, and fMRI. *Cortex*. 2010; 46:343–353. [PubMed: 19570530]
- Tang LS, Goeritz ML, Caplan JS, Taylor AL, Fisek M, Marder E. Precise Temperature Compensation of Phase in a Rhythmic Motor Pattern. *PLoS Biology*. 2010; 8:e1000469. [PubMed: 20824168]
- Tehovnik EJ. Electrical stimulation of neural tissue to evoke behavioral responses. *J Neurosci Methods*. 1996; 65:1–17. [PubMed: 8815302]
- Terao Y, Mizuno T, Shindoh M, Sakurai Y, Ugawa Y, Kobayashi S, Nagai C, Furubayashi T, Arai N, Okabe S, et al. Vocal amusia in a professional tango singer due to a right superior temporal cortex infarction. *Neuropsychologia*. 2006; 44:479–488. [PubMed: 15982678]
- Todorov E, Jordan MI. Optimal feedback control as a theory of motor coordination. *Nat Neurosci*. 2002; 5:1226–1235. [PubMed: 12404008]
- Tourville JA, Reilly KJ, Guenther FH. Neural mechanisms underlying auditory feedback control of speech. *Neuroimage*. 2008; 39:1429–1443. [PubMed: 18035557]
- Volgushev M, Vidyasagar TR, Chistiakova M, Yousef T, Eysel UT. Membrane properties and spike generation in rat visual cortical cells during reversible cooling. *The Journal of Physiology*. 2000; 522:59–76. [PubMed: 10618152]
- Wada J, Rasmussen T. Intracarotid Injection of Sodium Amytal for the Lateralization of Cerebral Speech Dominance. *Journal of Neurosurgery*. 1960; 17:266–282.

- Zarate JM. The neural control of singing. *Front Hum Neurosci.* 2013; 7:237. [PubMed: 23761746]
- Zarate JM, Zatorre RJ. Neural substrates governing audiovocal integration for vocal pitch regulation in singing. *Ann N Y Acad Sci.* 2005; 1060:404–408. [PubMed: 16597793]
- Zarate JM, Zatorre RJ. Experience-dependent neural substrates involved in vocal pitch regulation during singing. *NeuroImage.* 2008; 40:1871–1887. [PubMed: 18343163]
- Zatorre RJ, Belin P. Spectral and Temporal Processing in Human Auditory Cortex. *Cerebral Cortex.* 2001; 11:946–953. [PubMed: 11549617]
- Zatorre RJ, Belin P, Penhune VB. Structure and function of auditory cortex: music and speech. *Trends in cognitive sciences.* 2002; 6:37–46. [PubMed: 11849614]
- Zatorre RJ, Chen JL, Penhune VB. When the brain plays music: auditory-motor interactions in music perception and production. *Nature reviews Neuroscience.* 2007; 8:547–558. [PubMed: 17585307]

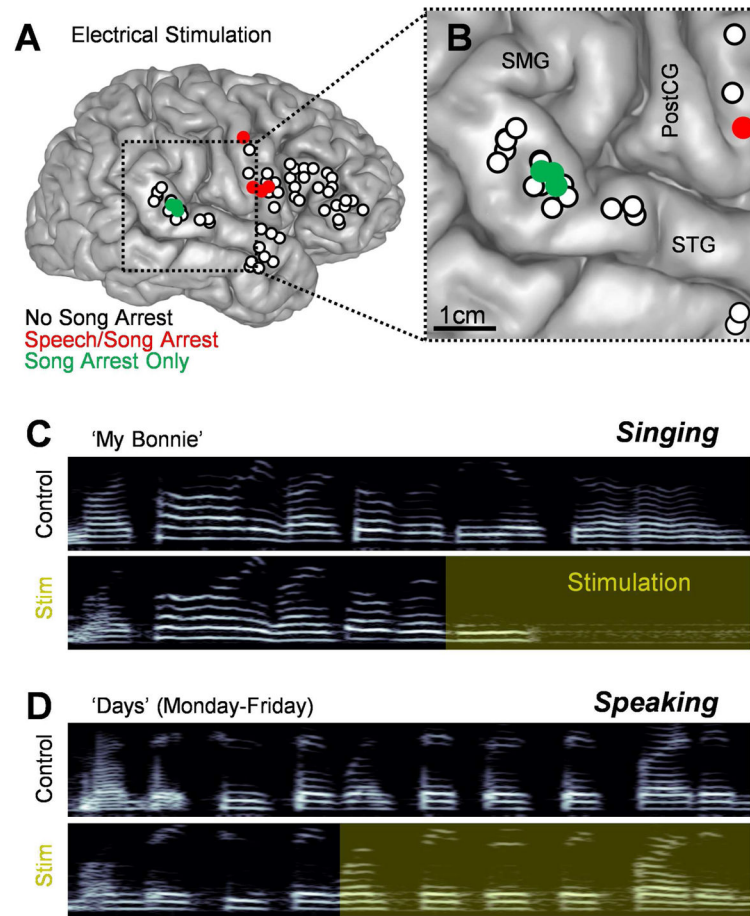


Fig. 1. Characterization of a song-selective locus via electrical stimulation mapping. (A) Anatomical reconstruction of the subject's right hemisphere highlighting 58 distinct stimulation locations. The color of each circle indicates the behavioral response to stimulation at each point. (B) A closer view of a portion of the right hemisphere including the supramarginal gyrus (SMG), superior temporal gyrus (STG), and postcentral gyrus (PostCG). (C–D) Example spectrograms (frequency range: 0–2000 Hz) showing a series of vocalizations with and without electrical stimulation administered to the pSTG. Yellow shading indicates the extent of stimulation.

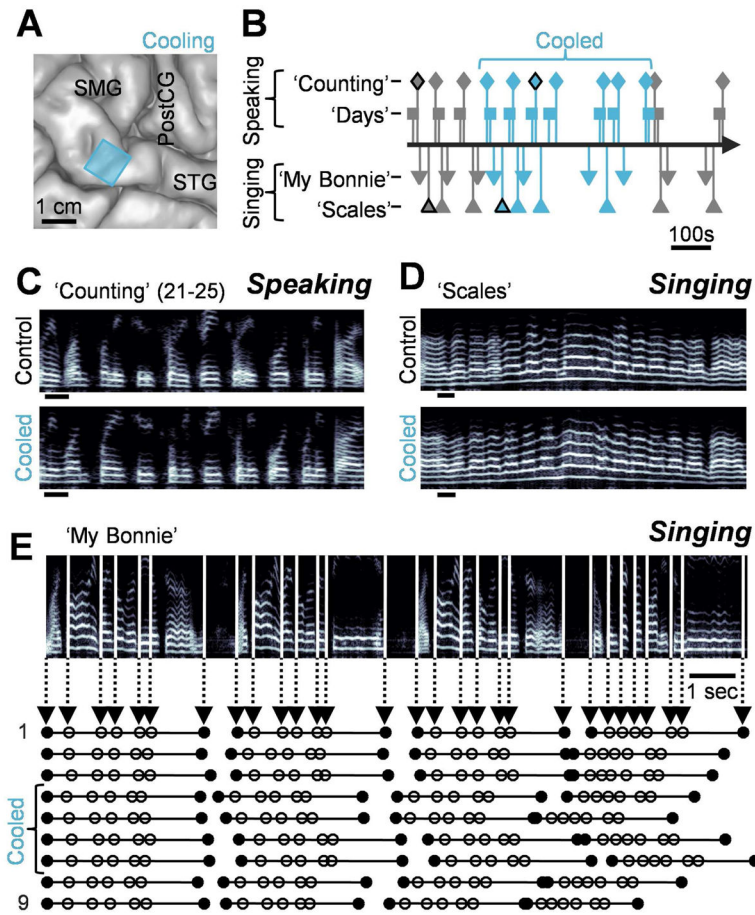


Fig. 2. Cooling a song-selective locus does not affect vocal quality or timing. (A) The subject's anatomical reconstruction with the location of the cooling probe indicated by a blue square. (B) Time course of the cooling experiment with lines indicating the onset times of each trial. For speaking, each marker denotes two successive trials. (C–D) Spectrograms of trials from each task performed in the control (top) and cooled (bottom) conditions indicating no obvious effect on articulation due to cooling. Black outlines of markers in (B) indicate examples shown in (C) and (D). Scale bars = 200 ms. (E) A spectrogram from the 'My Bonnie' task with vertical lines showing the onsets of individual words as well as the ends of each phrase (represented as circles in the plots below) for all 9 trials.

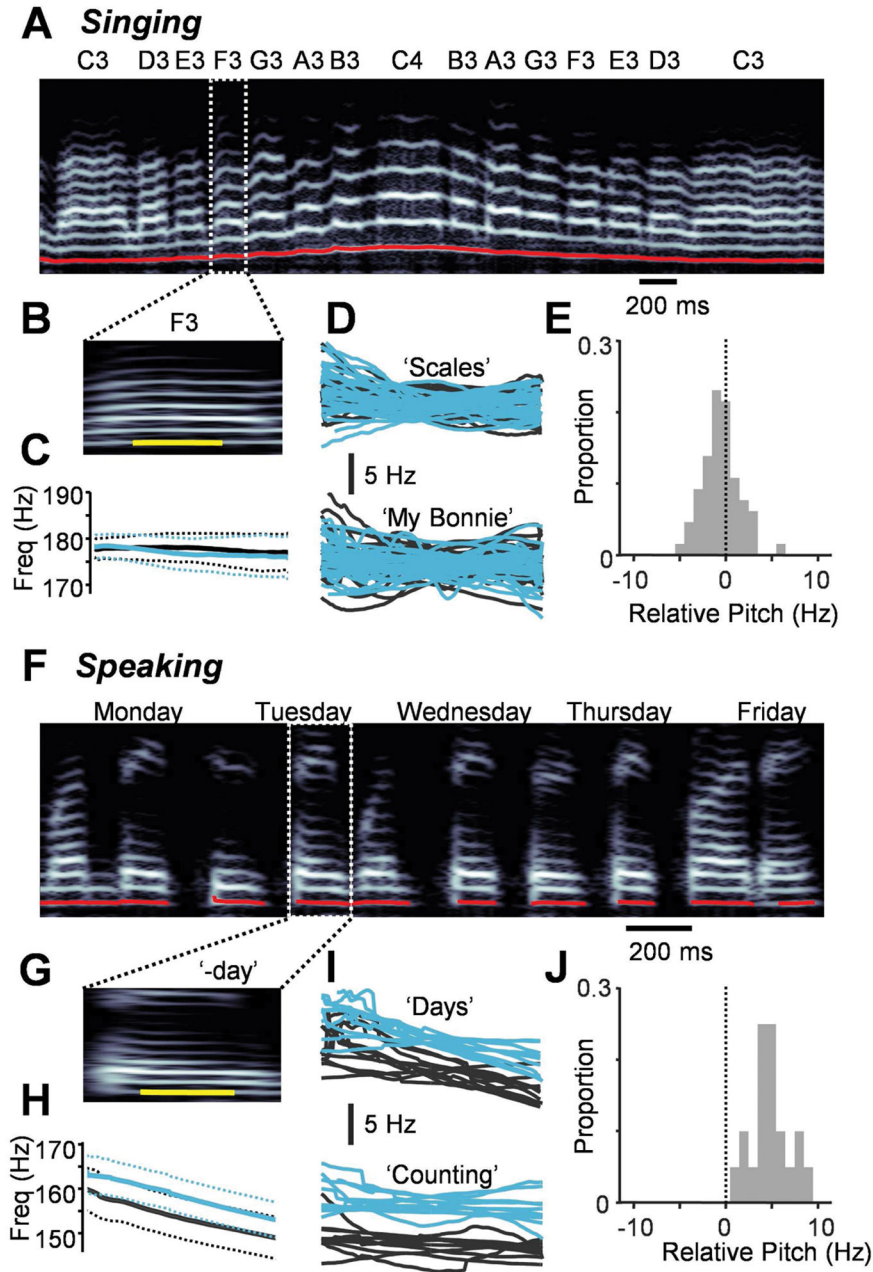


Fig. 3. Cooling causes a pitch change for speaking but not singing. (A) A spectrogram from the ‘Scales’ task with the pitch measurement overlaid in red. Target notes are shown above. (B) The ‘F3’ vocalization is shown along with the portion of the vowel analyzed (in yellow). (C) Mean pitch traces for the vowel shown in (A) for the cooled (blue, $n = 4$ trials) and control (black, $n = 5$ trials) conditions. Dotted lines = ± 2 SEM. (D) All vowels in the ‘Scales’ task ($n = 29$) and the ‘My Bonnie’ task ($n = 37$) with each trace representing an average of all trials for each unique vocal type relative to control values. (E) Change in pitch for each vowel during cooling relative to controls. (F,G) A spectrogram for the ‘Days’ task (F)

highlighting a portion of the word ‘Tuesday’ that was considered for further analysis (in yellow) (G). (H) Mean pitch traces across 24 trials (14 cooled and 10 control) of the vowel shown in (G) exhibit an upward shift during cooling. (I) All vowels in the ‘Counting’ task (n = 10) and the ‘Days’ task (n = 10). (J) Change in pitch for each vowel during cooling relative to controls.

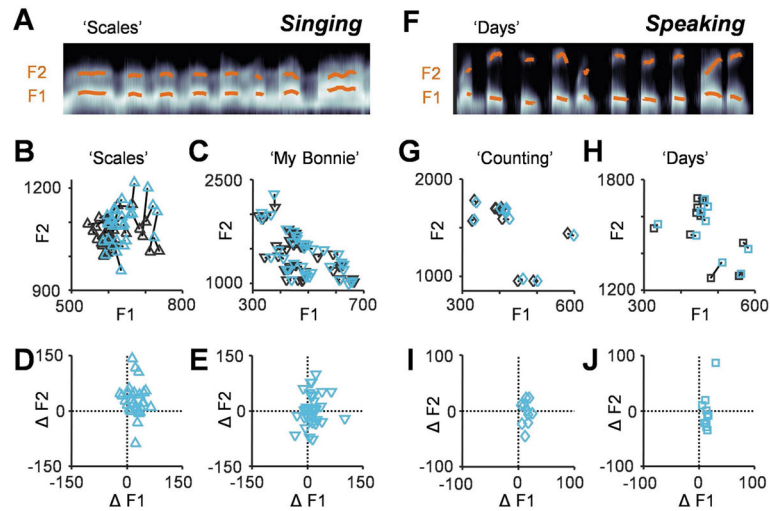


Fig. 4. Cooling changes the locations of vowels in formant space. (A) Frequency smoothed spectrogram from a portion of the ‘Scales’ task with the first and second formants overlaid in orange. (B,C) Locations of vowels in formant space for the ‘Scales’ (B) and ‘My Bonnie’ (C) tasks (control = black, cooled = blue) with lines connecting the same vowels. (D,E) The change in formants induced by cooling for each vowel in the ‘Scales’ (D) and ‘My Bonnie’ (E) tasks. (F) Spectrogram from the ‘Days’ task. (G,H) Locations of vowels in formant space for the ‘Counting’ (G) and ‘Days’ (H) tasks. (I,J) Cooling-induced formant changes in the ‘Counting’ (I) and ‘Days’ (J) tasks.