

## Video Article

# fMRI Mapping of Brain Activity Associated with the Vocal Production of Consonant and Dissonant Intervals

Nadia González-García<sup>1</sup>, Pablo L. Rendón<sup>2</sup><sup>1</sup>Hospital Infantil de México "Federico Gómez"<sup>2</sup>Centro de Ciencias Aplicadas y Desarrollo Tecnológico, Universidad Nacional Autónoma de MéxicoCorrespondence to: Pablo L. Rendón at [pablo.rendon@ccadet.unam.mx](mailto:pablo.rendon@ccadet.unam.mx)URL: <https://www.jove.com/video/55419>DOI: [doi:10.3791/55419](https://doi.org/10.3791/55419)

Keywords: Behavior, Issue 123, Consonance and dissonance, neural correlates of music, musical interval identification, singing, functional magnetic resonance imaging (fMRI), audio-vocal integration

Date Published: 5/23/2017

Citation: González-García, N., Rendón, P.L. fMRI Mapping of Brain Activity Associated with the Vocal Production of Consonant and Dissonant Intervals. *J. Vis. Exp.* (123), e55419, doi:10.3791/55419 (2017).

## Abstract

The neural correlates of consonance and dissonance perception have been widely studied, but not the neural correlates of consonance and dissonance production. The most straightforward manner of musical production is singing, but, from an imaging perspective, it still presents more challenges than listening because it involves motor activity. The accurate singing of musical intervals requires integration between auditory feedback processing and vocal motor control in order to correctly produce each note. This protocol presents a method that permits the monitoring of neural activations associated with the vocal production of consonant and dissonant intervals. Four musical intervals, two consonant and two dissonant, are used as stimuli, both for an auditory discrimination test and a task that involves first listening to and then reproducing given intervals. Participants, all female vocal students at the conservatory level, were studied using functional Magnetic Resonance Imaging (fMRI) during the performance of the singing task, with the listening task serving as a control condition. In this manner, the activity of both the motor and auditory systems was observed, and a measure of vocal accuracy during the singing task was also obtained. Thus, the protocol can also be used to track activations associated with singing different types of intervals or with singing the required notes more accurately. The results indicate that singing dissonant intervals requires greater participation of the neural mechanisms responsible for the integration of external feedback from the auditory and sensorimotor systems than does singing consonant intervals.

## Video Link

The video component of this article can be found at <https://www.jove.com/video/55419/>

## Introduction

Certain combinations of musical pitches are generally acknowledged to be consonant, and they are typically associated with a pleasant sensation. Other combinations are generally referred to as dissonant and are associated with an unpleasant or unresolved feeling<sup>1</sup>. Although it seems sensible to assume that enculturation and training play some part in the perception of consonance<sup>2</sup>, it has been recently shown that the differences in perception of consonant and dissonant intervals and chords probably depend less on musical culture than was previously thought<sup>3</sup> and may even derive from simple biological bases<sup>4,5,6</sup>. In order to prevent an ambiguous understanding of the term consonance, Terhardt<sup>7</sup> introduced the notion of sensory consonance, as opposed to consonance in a musical context, where harmony, for example, may well influence the response to a given chord or interval. In the present protocol, only isolated, two-note intervals were used precisely to single out activations solely related to sensory consonance, without interference from context-dependent processing<sup>8</sup>.

Attempts to characterize consonance through purely physical means began with Helmholtz<sup>9</sup>, who attributed the perceived roughness associated with dissonant chords to the beating between adjacent frequency components. More recently, however, it has been shown that sensory consonance is not only associated with the absence of roughness, but also with harmonicity, which is to say the alignment of the partials of a given tone or chord with those of an unheard tone of a lower frequency<sup>10,11</sup>. Behavioral studies confirm that subjective consonance is indeed affected by purely physical parameters, such as frequency distance<sup>12,13</sup>, but a wider range of studies have conclusively demonstrated that physical phenomena cannot solely account for the differences between perceived consonance and dissonance<sup>14,15,16,17</sup>. All of these studies, however, report these differences when listening to a variety of intervals or chords. A variety of studies using Positron Emission Tomography (PET) and functional Magnetic Resonance Imaging (fMRI) have revealed significant differences in the cortical regions that become active when listening to either consonant or dissonant intervals and chords<sup>8,18,19,20</sup>. The purpose of the present study is to explore the differences in brain activity when producing, rather than listening to, consonant and dissonant intervals.

The study of sensory-motor control during musical production typically involves the use of musical instruments, and very often it then requires the fabrication of instruments modified specifically for their use during neuroimaging<sup>21</sup>. Singing, however, would seem to provide from the start an appropriate mechanism for the analysis of sensory-motor processes during music production, as the instrument is the human voice itself, and the vocal apparatus does not require any modification in order to be suitable during imaging<sup>22</sup>. Although the neural mechanisms associated with aspects of singing, such as pitch control<sup>23</sup>, vocal imitation<sup>24</sup>, training-induced adaptive changes<sup>25</sup>, and the integration of external

feedback<sup>25,26,27,28,29</sup>, have been the subject of a number of studies over the past two decades, the neural correlates of singing consonant and dissonant intervals were only recently described<sup>30</sup>. For this purpose, the current paper describes a behavioral test designed to establish the adequate recognition of consonant and dissonant intervals by participants. This is followed by an fMRI study of participants singing a variety of consonant and dissonant intervals. The fMRI protocol is relatively straightforward, but, as with all MRI research, great care must be taken to correctly set up the experiments. In this case, it is particularly important to minimize head, mouth, and lip movement during singing tasks, making the identification of effects not directly related to the physical act of singing more straightforward. This methodology may be used to investigate the neural mechanisms associated with a variety of activities involving musical production by singing.

## Protocol

This protocol has been approved by the Research, Ethics, and Safety Committee of the Hospital Infantil de México "Federico Gómez".

### 1. Behavioral Pretest

1. Perform a standard, pure-tone audiometric test to confirm that all prospective participants possess normal hearing (20-dB Hearing Level (HL) over octave frequencies of -8,000 Hz). Use the Edinburgh Handedness Inventory<sup>31</sup> to ensure that all participants are right-handed.
2. **Generation of interval sequences.**
  1. Produce pure tones spanning two octaves, G4-G6, using a sound-editing program.  
NOTE: Here, the free, open-source sound editing software Audacity is described. Other packages may be used for this purpose.
    1. For each tone, open a new project in the sound-editing software.
    2. Under the "Generate" menu, select "Tone." In the window that appears, select a sine waveform, an amplitude of 0.8, and a duration of 1 s. Enter the value of the frequency that corresponds to the desired note (e.g., 440 Hz for A4). Click on the "OK" button.
    3. Under the "File" menu, select "Export Audio." In the window that opens, enter the desired name for the audio file and choose WAV as the desired file type. Click "Save."
  2. Select two consonant and two dissonant intervals, according to **Table 1**, in such a way that each consonant interval is close to a dissonant interval.  
NOTE: As an example, consider the consonant intervals of a perfect fifth and an octave and the dissonant intervals of an augmented fourth (tritone) and a major seventh. These are the intervals chosen for the study conducted by the authors.
  3. Generate all possible combinations of notes corresponding to these four intervals in the range between G4 and G6.
    1. For each interval, open a new project in the sound-editing software and use "Import Audio" under the "File" menu to import the two WAV files to be concatenated.
    2. Place the cursor at any point over the second tone and click to select. Click on "Select All" under the "Edit" menu. Under the same menu, click on "Copy."
    3. Place the cursor at any point over the first tone and click. Under the "Edit" menu click on "Move Cursor to Track End" and then click "Paste" under the same menu. Export the audio as described in step 1.2.1.3.
  4. Use a random sequence generator to produce sequences consisting of 100 intervals generated pseudorandomly in such a way that each of the four different intervals occurs exactly 25 times<sup>30</sup>. To do this, use the random permutation function in the statistical analysis software (see the **Table of Materials**). Input the four intervals as arguments and create a loop that repeats this process 25 times.
  5. Use behavioral research software to generate two distinct runs. Load a sequence of 100 intervals in WAV format for each run and associate the identification of each interval with a single trial<sup>30</sup>.  
NOTE: Here, E-Prime behavioral research software is used. Other equivalent behavioral research software can be used.
3. **Explain to participants that they will listen to two sequences of 100 intervals each, where each sequence is associated with a different task and with its own set of instructions. Tell participants that, in both runs, the next interval will be played only when a valid key is pressed.**  
NOTE: Once the interval recognition sequence commences, it should not be interrupted so that the course of action should be as clear as possible to all participants.
  1. Have the participants sit down in front of a laptop computer and wear the provided headphones. Use good-quality over-the-ear headphones. Adjust the sound level to a comfortable level for each subject.
  2. If using the behavioral research software described here, open the tasks created in step 1.2.5 with E-Run. In the window that appears, enter the session and subject number and click "OK." Use the session number to distinguish between runs for each participant.  
NOTE: The instructions for the task at hand will appear on screen, followed by the beginning of the task itself.
    1. First, in a 2-alternative forced-choice task, simply have the participants identify whether the intervals they hear are consonant or dissonant. Have the participant press "C" on the computer keyboard for consonant and "D" for dissonant.  
NOTE: Since all participants are expected to have musical training at a conservatory level, they are all expected to be able to distinguish between patently consonant and patently dissonant intervals. The first task serves, in a sense, as confirmation that this is indeed the case.
    2. Second, in a 4-alternative forced-choice task, have the participants identify the intervals themselves. Have the participants press the numerals "4," "5," "7," and "8" on the computer keyboard to identify the intervals of an augmented fourth, perfect fifth, major seventh, and octave, respectively.
  3. At the end of each task, press "OK" to automatically save the results for each participant in an individual E-DataAid 2.0 file labeled with the subject and session numbers and with the extension .edat2.

4. Use statistical analysis software (e.g., Matlab, SPSS Statistics, or an open-source alternative) to calculate the success rates for each task (*i.e.* the percentage of successful answers when identifying whether the intervals were consonant or dissonant, and also when identifying the intervals themselves), both individually and as a group<sup>32</sup>.

## 2. fMRI Experiment

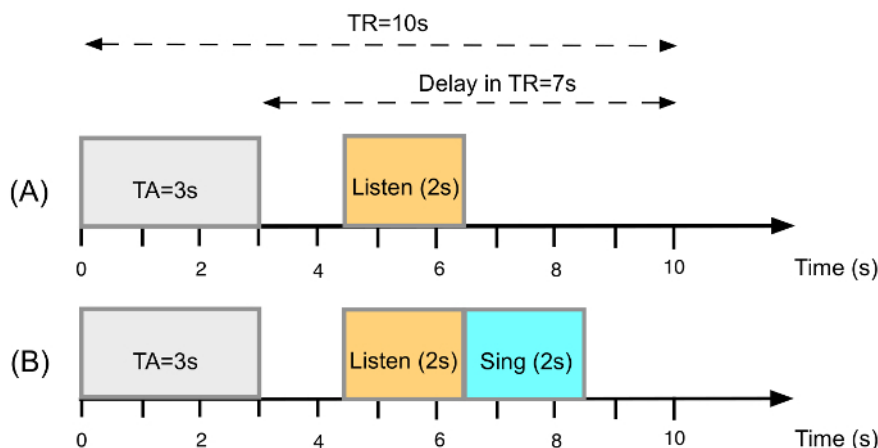
### 1. Preparation for the fMRI session.

1. Generate sequences of the same intervals as in step 1.2.3, again composed of two consecutive pure tones with a duration of 1 s each. NOTE: The vocal range of the participants must now be taken into account, and all notes must fall comfortably within the singing range of each participant.
  1. Use a random sequence generator to create a randomized sequence of 30 intervals for the listen-only trials<sup>30</sup>. For the singing trials, create a pseudorandomized sequence of 120 intervals for the participants to listen to a specific interval and then match this target interval with their singing voices. For the pseudorandomized sequence, use the same method as described in step 1.2.4, with the 4 intervals as arguments once again, but now repeating this process 30 times.
  2. Following the same procedure as in step 1.2.5, use behavioral research software to generate three distinct runs, each consisting initially of 10 silent baseline trials, followed by 10 consecutive listen-only trials, and finally by 40 consecutive singing trials. NOTE: During the baseline trials, the four intervals appear in random order, while during the singing trials, the four intervals appear in pseudorandomized order, in such a manner that each interval is eventually presented exactly 10 times. The duration of each trial is 10 s, so one whole run lasts 10 min. Since each subject goes through 3 experimental runs, the total duration of the experiment is 30 min. However, allowing for the participants to enter and exit the scanner, for time to set up and test the microphone, for time to obtain the anatomical scan, and for time between functional runs, approximately 1 h of scanner time should be allotted to each participant.
2. Explain to the participants the sequences of trials to be presented, as described in step 2.1.1.2, and respond to any doubts they might have. Instruct the participants to hum the notes without opening their mouths during the singing trials, keeping the lips still while producing an "m" sound.
3. Connect a non-magnetic, MR-compatible headset to a laptop. Adjust the sound level to a comfortable level for each subject.
4. Connect a small condenser microphone to an audio interface that is in turn connected to the laptop using a shielded twisted-triplet cable. NOTE: The microphone power supply, the audio interface, and the laptop should all be located outside the room housing the scanner.
5. Check the microphone frequency response. NOTE: The purpose of this test is to confirm that the microphone behaves as expected inside the scanner.
  1. Start a new project in the sound-editing software and select the condenser microphone as the input device.
  2. Generate a 440 Hz test tone with a duration of 10 s, as described in section 1.2.1, with the appropriate values for frequency and duration.
  3. Using the default sound reproduction software on the laptop, press "Play" to send the test tone through the headphones at locations inside (on top of the headrest) and outside (in the control room) the scanner, with the microphone placed between the sides of the headset in each case.
  4. Press "Record" in the sound-editing software to record approximately 1 s of the test tone at each location.
  5. Select "Plot Spectrum" from the "Analyze" menu for each case and compare the response of the microphone to the test tone, both inside and outside the scanner, by checking that the fundamental frequency of the signal received at each location is 440 Hz.
6. Tape the condenser microphone to the participant's neck, just below the larynx.
7. Have the participant wear the headset. Place the participant in a magnetic resonance (MR) scanner.

### 2. fMRI session.

1. At the beginning of the session, open the magnetic resonance user interface (MRUI) software package. Use the MRUI to program the acquisition paradigm. NOTE: Some variation in the interface is to be expected between different models.
  1. Select the "Patient" option from the onscreen menu. Enter the participant's name, age, and weight.
  2. Click on the "Exam" button. First, choose "Head" and then "Brain" from the available options.
  3. Select "3D" and then "T1 isometric," with the following values for the relevant parameters: Repetition Time (TR) = 10.2 ms, Echo Time (TE) = 4.2 ms, Flip Angle = 90°, and Voxel Size = 1 x 1 x 1 mm<sup>3</sup>. NOTE: For each participant, a T1-weighted anatomical volume will be acquired using a gradient echo pulse sequence for anatomical reference.
  4. Click on "Program" and select EchoPlanarImage\_diff\_perf\_bold (T2\*), with the values of the relevant parameters as follows: TE = 40 ms, TR = 10 s, Acquisition Time (TA) = 3 s, Delay in TR = 7 s, Flip Angle = 90°, Field of View (FOV) = 256 mm<sup>2</sup>, and Matrix Dimensions = 64 x 64. Use the "Dummy" option to acquire 5 volumes while entering a value of "55" for the total number of volumes. NOTE: These values permit the acquisition of functional T2\*-weighted whole-head scans according to the sparse sampling paradigm illustrated in **Figure 1**, where an echo-planar imaging (EPI) "dummy" scan is acquired and discarded to allow for T1 saturation effects. Note that in some MRUIs, the value of TR should be 3 s, as it is taken to be the total time during which the acquisition takes place.
  5. Click "Copy" to make a copy of this sequence. Place the cursor at the bottom of the list of sequences and then click "Paste" twice to set up three consecutive sparse sampling sequences.
  6. Click "Start" to begin the T1-weighted anatomical volume acquisition.
  7. Present three runs to the participant, with the runs as described in step 2.1.1.2. Synchronize the start of the runs with the acquisition by the scanner using the scanner trigger-box.

1. Follow the same procedure as described in section 1.3.2 to begin each one of the three runs, differentiating between runs using the session number. Save the results of three complete runs using the same procedure described in step 1.3.3.  
NOTE: The timing of the trial presentations is systematically jittered by  $\pm 500$  ms.



**Figure 1: Sparse-sampling Design.** (A) Timeline of events within a trial involving only listening to a two-tone interval (2 s), without subsequent overt reproduction. (B) Timeline of events within a trial involving listening and singing tasks. [Please click here to view a larger version of this figure.](#)

### 3. Data Analysis

1. **Preprocess the functional data using software designed for the analysis of brain imaging data sequences following standard procedures<sup>33</sup>.**

NOTE: All of the data processing is done using the same software.

1. Use the provided menu option to realign the images to the first volume, resampled and spatially normalized (final voxel size:  $2 \times 2 \times 2$  mm<sup>3</sup>) to standard Montreal Neurological Institute (MNI) stereotactic space<sup>34</sup>.
2. Use the provided menu option to smooth the image using an isotropic, 8 mm, Full Width at Half Maximum (FWHM) Gaussian kernel.
3. To model the BOLD response, select a single-bin Finite Impulse Response (FIR) as a basis function (order 1) or boxcar function, spanning the time of volume acquisition (3 s)<sup>28</sup>.

NOTE: Sparse-sampling protocols, such as this one, do not generally require the FIR to be convolved with the hemodynamic response function, as is commonly the case for event-related fMRI.

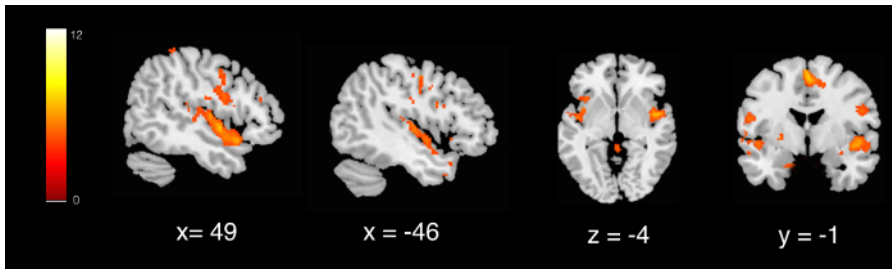
4. Apply a high-pass filter to the BOLD response for each event (1,000 s for the "singing network" and 360 s elsewhere).  
NOTE: Modelling all singing tasks together will amount to a block of 400 s<sup>35</sup>.

### Representative Results

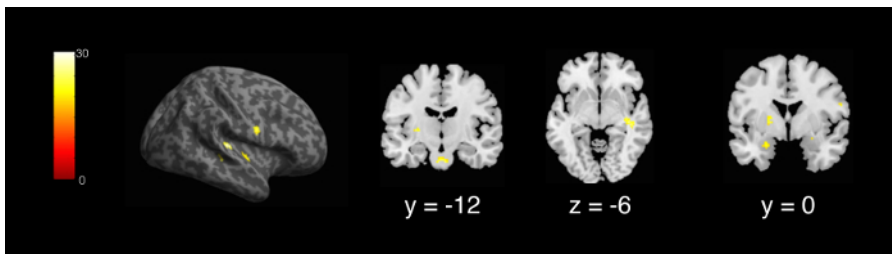
All 11 participants in our experiment were female vocal students at the conservatory level, and they performed well enough in the interval recognition tasks to be selected for scanning. The success rate for the interval identification task was  $65.72 \pm 21.67\%$ , which is, as expected, lower than the success rate when identifying dissonant and consonant intervals, which was  $74.82 \pm 14.15\%$ .

In order to validate the basic design of the study, we hoped to identify neural activity during singing in the regions known to constitute the "singing network," as defined in a number of previous studies<sup>25,26,27,28,29,30,31,32,33,34,35,36,37</sup>. The effect of singing is observed by means of the first-level linear contrast of interest, which corresponds to singing as opposed to listening. One-sample t-tests were used with clusters determined by  $Z > 3$  and a cluster significance threshold of  $p = 0.05$ , Family-Wise Error (FWE) corrected<sup>38</sup>. For anatomic labeling, the SPM Anatomy Toolbox<sup>33</sup> and the Harvard-Oxford cortical and subcortical structural atlas were used<sup>39</sup>. The brain regions where significant activation was observed were the primary somatosensory cortex (S1), the secondary somatosensory cortex (S2), the primary motor cortex (M1), the supplementary motor area (SMA), the premotor cortex (PM), Brodmann area 44 (BA 44), the primary auditory cortex (PAC), the superior temporal gyrus (STG), the temporal pole, the anterior insula, the putamen, the thalamus, and the cerebellum. These activations match those reported in the studies cited above regarding the "singing network," and they are illustrated in **Figure 2**. Note that in **Figures 2** and **3** both, the x-coordinate is perpendicular to the sagittal plane, the y-coordinate is perpendicular to the coronal plane, and the z-coordinate is perpendicular to the transverse, or horizontal plane.

Once the basic design has been validated, two further first-level linear contrasts were calculated for each participant, corresponding to singing dissonant as opposed to consonant intervals and to singing consonant as opposed to dissonant intervals. These linear contrasts were then taken to a second-level random-effects model involving a set of 2-way repeated-measures analyses of variance (ANOVA), with the factors consonance and dissonance. In this manner, the activated or deactivated areas were examined for possible interactions, with activations of interest determined according to the significance voxel threshold,  $p < 0.001$ , uncorrected for multiple comparisons<sup>28,29</sup>. For the contrast resulting from singing dissonant as opposed to consonant intervals, increased activations were observed in the right S1, right PAC, left midbrain, right posterior insula, left amygdala, and left putamen. These activations are shown in **Figure 3**. Regarding the complementary contrast, no significant changes in activation were detected during the singing of consonant intervals.



**Figure 2: Activation in Regions that Constitute the "Singing Network."** Activation maps are presented with a cluster significance threshold of  $p = 0.05$ , family-wise error (FWE) corrected. BOLD responses are reported in arbitrary units. [Please click here to view a larger version of this figure.](#)



**Figure 3: Contrast between the Singing of Dissonant and Consonant Intervals.** Activation maps are presented, uncorrected for multiple comparisons, with a cluster significance threshold of  $p = 0.001$ . BOLD responses are reported in arbitrary units. [Please click here to view a larger version of this figure.](#)

| Interval              | Number of semitones | Ratio of fundamentals |
|-----------------------|---------------------|-----------------------|
| <b>Unison</b>         | <b>0</b>            | <b>1:1</b>            |
| <i>Minor second</i>   | <i>1</i>            | <i>16:15</i>          |
| Major second          | 2                   | 9:8                   |
| Minor third           | 3                   | 6:5                   |
| <b>Major third</b>    | <b>4</b>            | <b>5:4</b>            |
| <b>Perfect fourth</b> | <b>5</b>            | <b>4:3</b>            |
| <i>Tritone</i>        | <i>6</i>            | <i>45:32</i>          |
| <b>Perfect fifth</b>  | <b>7</b>            | <b>3:2</b>            |
| Minor sixth           | 8                   | 8:5                   |
| <b>Major sixth</b>    | <b>9</b>            | <b>5:3</b>            |
| Minor seventh         | 10                  | 16:9                  |
| <i>Major seventh</i>  | <i>11</i>           | <i>15:8</i>           |
| <b>Octave</b>         | <b>12</b>           | <b>2:1</b>            |

**Table 1: Consonant and Dissonant Musical Intervals.** Consonant intervals appear in boldface, while dissonant intervals appear in italics. Observe that the more consonant an interval, the smaller the integers that appear in the frequency ratio used to represent it. For an in-depth discussion of consonance and dissonance as a function of frequency ratios, see Bidelman & Krishnan<sup>40</sup>.

## Discussion

This work describes a protocol in which singing is used as a means of studying brain activity during the production of consonant and dissonant intervals. Even though singing provides what is possibly the simplest method for the production of musical intervals<sup>22</sup>, it does not allow for the production of chords. However, although most physical characterizations of the notion of consonance rely, to some degree, on the superposition

of simultaneous notes, a number of studies have shown that intervals constructed with notes that correspond to consonant or dissonant chords are still perceived as consonant or dissonant, respectively<sup>4,6,15,41,42</sup>.

The behavioral interval perception task is used to establish, before participants can go through to the scanning session, if they are able to adequately distinguish the intervals. Thus, they perform well once inside the magnetic resonator. Any participants unable to meet a predetermined threshold when performing these identification tasks should not proceed to the fMRI experiment. The main purpose of this selection process is to ensure that differences in performance between the participants are not due to deficient perceptual abilities. Chosen participants should have similar degrees of vocal and musical training, and also, if possible, similar tessituras. If they have vocal ranges that vary significantly, the range of intervals presented during the singing tasks they will need to perform must be personalized.

The microphone setup is critical for the acquisition to be reliable and artifact-free. The type of microphone itself is very important, and although it is possible to use optical<sup>28</sup> or specially designed, MR-compatible<sup>29</sup> microphones, it has been shown that the sensitivity of condenser microphones is not affected by the presence of the intense magnetic fields in the imaging environment<sup>43</sup>. Indeed, a small Lavalier condenser microphone may be used in this context, provided a shielded twisted-triplet cable is used to connect the microphone to the preamplifier, which must be placed outside the room where the MR scanner is housed. This arrangement will prevent the appearance of imaging artifacts<sup>44</sup>, but researchers should also ensure that the scanner does not interfere with the performance of the microphone. To this end, a test tone can be sent through the MR-compatible headphones to the microphone placed inside the MR scanner, and the signal obtained in this manner can then be compared to that obtained by sending the same tone to the microphone now placed outside the scanner. The sound pressure levels inside the MR scanner can be extremely high<sup>45</sup>, so the microphone must be placed as close as possible to the source. By asking participants to hum rather than openly sing notes, the movement in and around the mouth area can be minimized. By placing the microphone just below the larynx, covered by tape, it is possible to obtain a faithful recording of the singer's voice. The recording will naturally be very noisy – this cannot be avoided – but if researchers are mainly interested in pitch and not in the articulation or enunciation of words, a variety of software packages can be used to clean the signal enough for the detection of the fundamental frequency of each sung note. A standard method would be to use audio-editing software to filter the time signals through a Hamming window and then to use the autocorrelation algorithms built into certain speech and phonetics software packages to identify the sung fundamentals. Vocal accuracy can then be calculated for each participant. Potential applications of the data obtained from the recordings include correlating pitch or rhythmic accuracy with either degree of training or interval distances.

Functional images are acquired using a sparse sampling design so as to minimize BOLD or auditory masking due to scanning noise<sup>25,28,29,46,47</sup>. Every subject undergoes 3 experimental runs, each one lasting 10 min. During each run, subjects are first asked to lay in silence during 10 silent baseline trials, then to listen passively to a block of 10 intervals, and finally to listen to and sing back another block of 40 intervals. One purpose of keeping individual runs as short as possible is to avoid participant fatigue. Nonetheless, it has since been concluded that it might be better in the future to include the same number of listen-only and singing trials, which can then be presented in alternating blocks. This would have the effect of increasing statistical power. As an example, a run could consist of 2 blocks of 5 silent baseline trials, 4 blocks of 5 listen-only trials, and 4 blocks of singing trials. The blocks would then be presented to participants in alternation, with a total duration of 500 s per run.

The main reason for having participants listen passively inside the resonator is to have a means of subtracting auditory activity from motor activity. Thus, a favorable comparison of singing activations against the "singing network"<sup>25,27,28,29,36,37</sup> is indispensable for the proper validation of the study. Note that "singing network" activations are very robust and well-established and are usually detected by means of one-sample t-tests and a corrected cluster significance threshold of  $p = 0.05$ . Activations corresponding to the contrast between singing dissonant/consonant and consonant/dissonant intervals are typically identified by means of 2-way repeated factors analyses of variance (ANOVA) according to the significance voxel threshold  $p < 0.001$ , uncorrected for multiple comparisons<sup>28,29</sup>. It is expected that participants will find singing dissonant intervals more challenging than singing consonant intervals<sup>48,49</sup>, thus, different activations for each of the two contrasts described above are anticipated. Results indicate that singing dissonant intervals involves a reprogramming of the neural mechanisms recruited for the production of consonant intervals. During singing, produced sound is compared to intended sound, and any necessary adjustment is then achieved through the integration of external and internal feedback from auditory and somatosensory pathways. A detailed discussion of these results and the conclusions drawn from them is included in the article by González-García, González, and Rendón<sup>30</sup>.

This protocol provides a reasonably straightforward method for the study of the neural correlates of musical production and for monitoring the activity of both the motor and auditory systems. It can be used to track differences in brain activation between binary conditions, such as singing consonant or dissonant intervals, and singing narrow or wide intervals<sup>30</sup>. It is also well-suited to study the effect of training on a variety of tasks associated with singing specific frequencies. On the other hand, because of the very large amount of noise contained in recordings of the sung voice obtained during the scan, it would be difficult to employ this protocol to analyze tasks concerned with quality of tone or timbre, especially because these are qualities that cannot be gauged correctly while humming.

## Disclosures

The authors declare no conflicts of interest.

## Acknowledgements

The authors acknowledge financial support for this research from Secretaría de Salud de México (HIM/2011/058 SSA. 1009), CONACYT (SALUD-2012-01-182160), and DGAPA UNAM (PAPIIT IN109214).

## References

1. Burns, E. M. *Intervals, scales, and tuning. The psychology of music*. Deutsch, D., ed., London: Academic Press, 215-264 (1999).
2. Lundin, R. W. Toward a cultural theory of consonance. *J. Psychol.* **23**, 45-49 (1947).

3. Fritz, T., Jentschke, S., *et al.* Universal recognition of three basic emotions in music. *Curr. Biol.* **19**, 573-576 (2009).
4. Schellenberg, E. G., & Trehub, S. E. Frequency ratios and the discrimination of pure tone sequences. *Percept. Psychophys.* **56**, 472-478 (1994).
5. Trainor, L. J., & Heinmiller, B. M. The development of evaluative responses to music: *Infant Behav. Dev.* **21**, 77-88 (1998).
6. Zentner, M. R., & Kagan, J. Infants' perception of consonance and dissonance in music. *Infant Behav. Dev.* **21**, 483-492 (1998).
7. Terhardt, E. Pitch, consonance, and harmony. *J. Acoust. Soc. America.* **55**, 1061 (1974).
8. Minati, L., *et al.* Functional MRI/event-related potential study of sensory consonance and dissonance in musicians and nonmusicians. *Neuroreport.* **20**, 87-92 (2009).
9. Helmholtz, H. L. F. *On the sensations of tone.* New York: Dover (1954).
10. McDermott, J. H., Lehr, A. J., & Oxenham, A. J. Individual differences reveal the basis of consonance. *Curr. Biol.* **20**, 1035-1041 (2010).
11. Cousineau, M., McDermott, J. H., & Peretz, I. The basis of musical consonance as revealed by congenital amusia. *Proc. Natl. Acad. Sci. USA.* **109**, 19858-19863 (2012).
12. Plomp, R., & Levelt, W. J. M. Tonal Consonance and Critical Bandwidth. *J. Acoust. Soc. Am.* **38**, 548-560 (1965).
13. Kameoka, A., & Kuriyagawa, M. Consonance theory part I: Consonance of dyads. *J. Acoust. Soc. Am.* **45**, 1451-1459 (1969).
14. Tramo, M. J., Bharucha, J. J., & Musiek, F. E. Music perception and cognition following bilateral lesions of auditory cortex. *J. Cogn. Neurosci.* **2**, 195-212 (1990).
15. Schellenberg, E. G., & Trehub, S. E. Children's discrimination of melodic intervals. *Dev. Psychol.* **32**, 1039-1050 (1996).
16. Peretz, I., Blood, A. J., Penhune, V., & Zatorre, R. J. Cortical deafness to dissonance. *Brain.* **124**, 928-940 (2001).
17. McDermott, J. H., Schultz, A. F., Undurraga, E. A., & Godoy, R. A. Indifference to dissonance in native Amazonians reveals cultural variation in music perception. *Nature.* **535**, 547-550 (2016).
18. Blood, A. J., Zatorre, R. J., Bermudez, P., & Evans, A. C. Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nat. Neurosci.* **2**, 382-387 (1999).
19. Pallesen, K. J., *et al.* Emotion processing of major, minor, and dissonant chords: A functional magnetic resonance imaging study. *Ann. N. Y. Acad. Sci.* **1060**, 450-453 (2005).
20. Foss, A. H., Altschuler, E. L., & James, K. H. Neural correlates of the Pythagorean ratio rules. *Neuroreport.* **18**, 1521-1525 (2007).
21. Limb, C. J., & Braun, A. R. Neural substrates of spontaneous musical performance: An fMRI study of jazz improvisation. *PLoS ONE.* **3**, (2008).
22. Zarate, J. M. The neural control of singing. *Front. Hum. Neurosci.* **7**, 237 (2013).
23. Larson, C. R., Altman, K. W., Liu, H., & Hain, T. C. Interactions between auditory and somatosensory feedback for voice F0 control. *Exp. Brain Res.* **187**, 613-621 (2008).
24. Belyk, M., Pfordresher, P. Q., Liotti, M., & Brown, S. The neural basis of vocal pitch imitation in humans. *J. Cogn. Neurosci.* **28**, 621-635 (2016).
25. Kleber, B., Veit, R., Birbaumer, N., Gruzeliel, J., & Lotze, M. The brain of opera singers: Experience-dependent changes in functional activation. *Cereb. Cortex.* **20**, 1144-1152 (2010).
26. Jürgens, U. Neural pathways underlying vocal control. *Neurosci. Biobehav. Rev.* **26**, 235-258 (2002).
27. Kleber, B., Birbaumer, N., Veit, R., Trevorrow, T., & Lotze, M. Overt and imagined singing of an Italian aria. *Neuroimage.* **36**, 889-900 (2007).
28. Kleber, B., Zeitouni, A. G., Friberg, A., & Zatorre, R. J. Experience-dependent modulation of feedback integration during singing: role of the right anterior insula. *J. Neurosci.* **33**, 6070-6080 (2013).
29. Zarate, J. M., & Zatorre, R. J. Experience-dependent neural substrates involved in vocal pitch regulation during singing. *Neuroimage.* **40**, 1871-1887 (2008).
30. González-García, N., González, M. A., & Rendón, P. L. Neural activity related to discrimination and vocal production of consonant and dissonant musical intervals. *Brain Res.* **1643**, 59-69 (2016).
31. Oldfield, R. C. The assessment and analysis of handedness: The Edinburgh inventory. *Neuropsychologia.* **9**, 97-113 (1971).
32. Samuels, M. L., Witmer, J. A., Schaffner, A. *Statistics for the Life Sciences.* Harlow: Pearson (2015).
33. Eickhoff, S. B., *et al.* A new SPM toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *NeuroImage.* **25**, 1325-1335 (2005).
34. Evans, A. C., Kamber, M., Collins, D. L., & MacDonald, D. An MRI-based probabilistic atlas of neuroanatomy. In: *Magnetic Resonance Scanning and Epilepsy.* Shorvon, S.D., Fish, D.R., Andermann, F., Bydder, G. M., Stefan, H., eds., **264**, 263-274 (1994).
35. Ashburner, J., *et al.* *SPM8 Manual.* London: Wellcome Trust Centre for Neuroimaging (2013).
36. Özdemir, E., Norton, A., & Schlaug, G. Shared and distinct neural correlates of singing and speaking. *Neuroimage.* **33**, 628-635 (2006).
37. Brown, S., Ngan, E., & Liotti, M. A larynx area in the human motor cortex. *Cereb. Cortex.* **18**, 837-845 (2008).
38. Worsley, K. J. Statistical analysis of activation images. In: *Functional MRI: An introduction to methods.* Jezzard, P., Matthews, P. M., Smith, S. M., eds., Oxford: Oxford University Press, 251-270 (2001).
39. *FSL Atlases.* <<https://fsl.fmrib.ox.ac.uk/fsl/fslwiki/Atlases>> 12/05/2016 (2015).
40. Bidelman, G. M., & Krishnan, A. Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *J. Neurosci.* **29**, 13165-13171 (2009).
41. McLachlan, N., Marco, D., Light, M., & Wilson, S. Consonance and pitch. *J. Exp. Psychol. - Gen.* **142**, 1142-1158 (2013).
42. Thompson, W. F. Intervals and scales. In: *The psychology of music.* Deutsch, D., ed., London: Academic Press, 107-140 (1999).
43. Hurwitz, R., Lane, S. R., Bell, R. A., & Brant-Zawadzki, M. N. Acoustic analysis of gradient-coil noise in MR imaging. *Radiology.* **173**, 545-548 (1989).
44. Ravicz, M. E., Melcher, J. R., & Kiang, N. Y.-S. Acoustic noise during functional magnetic resonance imaging. *J. Acoust. Soc. Am.* **108**, 1683-1696 (2000).
45. Cho, Z. H., *et al.* Analysis of acoustic noise in MRI. *Magn. Reson. Imaging.* **15**, 815-822 (1997).
46. Belin, P., Zatorre, R. J., Hoge, R., Evans, A. C., & Pike, B. Event-related fMRI of the auditory cortex. *Neuroimage.* **429**, 417-429 (1999).
47. Hall, D. A., *et al.* "Sparse" temporal sampling in auditory fMRI. *Hum. Brain Mapp.* **7**, 213-223 (1999).
48. Ternström, S., & Sundberg, J. Acoustical factors related to pitch precision in choir singing. *Speech Music Hear. Q. Prog. Status Rep.* **23**, 76-90 (1982).
49. Ternström, S., & Sundberg, J. Intonation precision of choir singers. *J. Acoust. Soc. Am.* **84**, 59-69 (1988).