



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
PROGRAMA DE MAESTRÍA Y DOCTORADO EN MÚSICA

FACULTAD DE MÚSICA  
CENTRO DE CIENCIAS APLICADAS Y DESARROLLO TECNOLÓGICO  
INSTITUTO DE INVESTIGACIONES ANTROPOLÓGICAS

ACTIVIDAD NEURONAL RELACIONADA CON EL CANTO DE INTERVALOS  
MUSICALES CONSONANTES Y DISONANTES.

TESIS  
QUE, PARA OPTAR POR EL GRADO DE  
MAESTRA EN MÚSICA (COGNICIÓN MUSICAL)

PRESENTA

MARTHA ANGÉLICA GONZÁLEZ MONTIEL

TUTORES PRINCIPALES

DRA. NADIA GONZÁLEZ GARCÍA (LABORATORIO DE NEUROCIENCIAS,  
HOSPITAL INFANTIL DE MÉXICO “FEDERICO GÓMEZ”).

DR. PABLO L. RENDÓN GARRIDO (CENTRO DE CIENCIAS APLICADAS Y  
DESARROLLO TECNOLÓGICO, UNAM).

CIUDAD DE MÉXICO, 2017.



Universidad Nacional  
Autónoma de México

Dirección General de Bibliotecas de la UNAM

**Biblioteca Central**



**UNAM – Dirección General de Bibliotecas**  
**Tesis Digitales**  
**Restricciones de uso**

**DERECHOS RESERVADOS ©**  
**PROHIBIDA SU REPRODUCCIÓN TOTAL O PARCIAL**

Todo el material contenido en esta tesis esta protegido por la Ley Federal del Derecho de Autor (LFDA) de los Estados Unidos Mexicanos (México).

El uso de imágenes, fragmentos de videos, y demás material que sea objeto de protección de los derechos de autor, será exclusivamente para fines educativos e informativos y deberá citar la fuente donde la obtuvo mencionando el autor o autores. Cualquier uso distinto como el lucro, reproducción, edición o modificación, será perseguido y sancionado por el respectivo titular de los Derechos de Autor.

Dedico este logro a mi abuela, Ma. Luz Rentería Acosta, por ser un ejemplo de determinación, humildad, fortaleza y entrega; y, a mi mejor amigo y hermano del alma, Adrián Ruiz Ramírez, por siempre apoyarme, amarme incondicionalmente, recordarme que la vida es un instante y enseñarme el significado de las palabras fortaleza, alegría, amor, amistad, lealtad y trascendencia.

Agradecimientos:

A la Universidad Nacional Autónoma de México y a su sistema de becas para estudios de posgrado, pues sin su apoyo no hubiera podido realizar mis estudios y desarrollar esta investigación.

Al personal del Posgrado en Música de la Facultad de Música (UNAM) y a los docentes que me apoyaron con su guía: Dr. Eduardo Castro Sierra †, Dra. Nadia González García y Dr. Pablo Rendón Garrido.

Al personal del Hospital Infantil de México “Federico Gómez” por el apoyo y por permitirme realizar las resonancias magnéticas en sus instalaciones.

Al personal del CCADET por apoyarnos con el equipo necesario para llevar a cabo la adquisición de señal sonora y las audiometrías.

A la Academia de Jazz de la Escuela Superior de Música por las facilidades brindadas.

A las estudiantes de canto de la Escuela Superior de Música por ser parte de este hermoso proyecto.

A Verónica Ituarte por su guía, amistad y apoyo incondicional.

A Khenia Yamel Bello Valverde por ser mi fuente de inspiración, por permitirme amarla y compartir mi vida con ella, por apoyarme y amarme bajo cualquier circunstancia.

A mi familia.

## Índice.

1. Introducción y antecedentes.
  - 1.1. Consonancia y disonancia.
  - 1.2. Anatomía y producción de la voz cantada.
  - 1.3. Correlatos neurales del canto.
  - 1.4. Cambios cerebrales (funcionales y estructurales) dependientes del entrenamiento musical.
2. Métodos.
  - 2.1. Metodología e identificación demográfica.
  - 2.2. Pruebas conductuales.
  - 2.3. Resonancia magnética funcional (fMRI).
3. Resultados.
  - 3.1. Pruebas conductuales.
  - 3.2. Tiempos de reacción.
  - 3.3. Canto durante la fMRI.
  - 3.4. Resonancia magnética funcional (fMRI).
4. Discusión.
  - 4.1. Resonancia magnética funcional y pruebas conductuales.
5. Alcances y limitaciones.
6. Bibliografía.
7. Anexos.
  - 7.1. Neural activity related to discrimination and vocal production of consonant and dissonant musical intervals.

## 1. Introducción y antecedentes.

La práctica musical es ambas, una actividad humana natural, presente en todas las sociedades y uno de los retos más complejos y demandantes cognitivamente que la mente humana puede realizar. Los mecanismos neurales involucrados en la producción y percepción musical proveen una rica fuente de preguntas para las neurociencias cognitivas (Zatorre et al., 2007).

Comprender la percepción auditiva y producción vocal de los intervalos musicales consonantes y disonantes, el automonitoreo de la emisión vocal por medio del cual los cantantes se aseguran de estar entonando correctamente el tono deseado, la retroalimentación somatosensorial y la identificación de los correlatos neurales involucrados, son fundamentales para entender la práctica musical del canto.

En la presente investigación nos centramos en identificar y comprender los mecanismos neurobiológicos que se encuentran involucrados en el canto de intervalos consonantes y disonantes. Es decir, el objetivo de este estudio es investigar los mecanismos neurales a través de los cuales, los cantantes entrenados producen intervalos consonantes y disonantes con su voz.

## 1.1 Consonancia y disonancia.

Aunque no existe una sola definición, la consonancia comúnmente es entendida como la sensación sonora estable (reposo) y placentera, mientras que la disonancia es la sensación sonora no placentera e inestable (movimiento) (Moncada, 2011).

La consonancia ha sido privilegiada en la teoría y práctica de la música occidental a lo largo de la historia, por lo que desde hace algunos siglos se ha tratado de describir y entender el complejo proceso de percepción de los sonidos y su relación con la práctica musical (como escucha, ejecutante y creador).

El cómo explicamos y entendemos la percepción de la consonancia y la disonancia se ha ido complejizando a través del tiempo. La teoría dominante para explicar las sensaciones que nos produce la escucha de intervalos musicales consonantes y disonantes, es atribuida a Pitágoras y sugiere que entre más simple es la proporción de frecuencias entre dos tonos, estos serán percibidos como más consonantes (Shapira and Stone, 2008).

Helmholtz estableció una base científica para explicar el fenómeno de la consonancia y la disonancia en 1877, a partir del número y fuerza de los componentes armónicos “batientes” en un par de tonos complejos simultáneos. Desde los tiempos de Von Helmholtz, se asumía que por alguna razón nuestro sistema auditivo no gusta de los batidos ni de la rugosidad (Roederer, 1975) (Hartmann, 1998).

El batido o batimiento es un efecto de primer orden que resulta de la superposición de dos o más sonidos, cuando estos se encuentran a una distancia aproximada menor a 15hz (banda crítica), ya que es un procesamiento mecánico que ocurre en el fluido coclear a lo largo de la membrana basilar. Los efectos de segundo orden son el resultado del procesamiento neural. Cuando se da el efecto del batido, tenemos la sensación de que se escucha un solo tono cuya amplitud varía lentamente. El batido se acelera cuando se acerca la diferencia de las frecuencias superpuestas a la banda crítica (15Hz), convirtiéndose en un efecto de rugosidad, hasta que finalmente desaparece cuando se sobrepasa la misma y se perciben los dos tonos claramente (Roederer, 1975).

Helmholtz argumentaba que en el caso de los dos sonidos componentes de un intervalo de unísono ( $f_1=P$ ,  $f_2=Q$ )( $P:Q=1:1$ ), o de octava ( $P:Q=1:2$ ), todos los armónicos del segundo tono se alinean y coinciden en frecuencia con el primero, y así se perciben como más consonantes que los intervalos cuyos armónicos no coinciden. Sin embargo, conforme la proporción de las frecuencias se va haciendo más compleja, ambos tonos tienen menos armónicos comunes o coincidentes, mientras que hay un incremento de los armónicos que no encajan ligeramente en la frecuencia o que no coinciden.

Lo anterior permite que la interacción de estas componentes armónicas provoque una sensación no placentera de batido y rugosidad que resulta en lo conocemos como disonancia.

Una vez que la diferencia de frecuencias entre dos tonos puros sobrepasa los 3 semitonos aproximadamente, la rugosidad desaparece (Plomp and Levelt, 1965). Sin embargo, más allá de la banda crítica, la evaluación de la consonancia puede variar considerablemente y cambiar de dirección cuando la diferencia entre las frecuencias, se incrementa. Estos cambios en la consonancia ocurren aún en ausencia de armónicos e incluso de batidos (Shapira and Stone, 2008).

Las secuencias de intervalos de tonos puros con una proporción simple de frecuencias son más fácilmente procesados por los escuchas (Schellenberg and Trehub, 1996). Esto sugiere un estado perceptual especial para los intervalos con proporción de frecuencias simples. En estos últimos casos, la teoría de Von Helmholtz ya no es aplicable pues no hay componentes armónicas que coincidan o que se desfasen.

La consonancia y la disonancia son sensaciones subjetivas asociadas con dos o más tonos que suenan simultáneamente, con la ausencia o presencia de batidos y rugosidad (Roederer, 1975). Aunque también existen estos fenómenos perceptuales cuando se tocan intervalos melódicamente con o sin un contexto armónico definido y cuando los sonidos se escuchan dicóticamente (un estímulo por oído). Los intervalos armónicos son aquellos que se aprecian simultáneamente y los melódicos son aquellos que se aprecian sucesivamente. Podemos hablar de



consonancia y disonancia ya sea que los intervalos musicales sean armónicos o melódicos (Danhauser, 2001).

En la teoría musical occidental los intervalos se dividen en consonantes y disonantes, y a su vez, los primeros en consonancias perfectas, imperfectas y para algunos autores, también en mixtas (Danhauser, 2001).

Los intervalos que corresponden a la categoría de consonancias perfectas son los justos, es decir, la primera justa, la cuarta justa, la quinta justa y la octava justa. Los que corresponden a las consonancias imperfectas son los intervalos de terceras y sextas mayores y menores (Fig. 1.1-1).

El intervalo que es considerado por algunos autores como consonancia mixta o como disonancia es el de cuarta justa por ser el armónico número 21 en la escala de los armónicos y porque no tiene el mismo efecto de reposo o conclusión que caracteriza a los intervalos de quinta y octava justa. Pero este intervalo también es considerado por algunos autores como consonancia perfecta por que aparece entre el tercer y el cuarto armónico en la escala de los armónicos y tiene una relación de 4/3 entre estos (Danhauser, 2001) (Moncada, 2011) (Hernández, 1989).

I N T E R V A L O S

<i>Consonancias</i>	<i>Disonancias</i>
<p><i>Consonancias perfectas:</i>  <i>1<sup>as</sup> Justas, 4<sup>as</sup> Justas</i>  <i>5<sup>as</sup> Justas y 8<sup>as</sup> Justas</i></p>	<p><i>2<sup>as</sup> y 7<sup>as</sup> mayores y menores y todos los aumentados (+) y todos los disminuidos (-)</i></p>
<p><i>Consonancias imperfectas;</i>  <i>3<sup>as</sup> y 6<sup>as</sup> mayores y menores</i></p>	

Fig. 1.1-1. Tabla de intervalos consonantes y disonantes. (Tomado de Hernández, 1989)

Los armónicos son sonidos concomitantes que son producidos por la resonancia de otro sonido generador. Los distintos tipos de consonancias se establecen en función de la cercanía de los armónicos al sonido fundamental, es decir, las consonancias perfectas son las más cercanas a la tónica, le siguen las imperfectas y por último las mixtas (Hernández, 1989) (Moncada, 2011) (Fig.1.1-2).

<i>Armónico:</i>	<i>Intervalo:</i>
1	1 <sup>a</sup> Justa o perfecta
3	5 <sup>a</sup> Justa o perfecta
5	3 <sup>a</sup> Mayor
7	7 <sup>a</sup> Menor
9	2 <sup>a</sup> Mayor
11	4 <sup>a</sup> Aumentada
13	6 <sup>a</sup> Menor o (5 <sup>a</sup> aumentada)
15	7 <sup>a</sup> Mayor
17	2 <sup>a</sup> Menor
19	3 <sup>a</sup> Menor
21	4 <sup>a</sup> Justa o perfecta
23	1/4 de Tono entre 4 <sup>a</sup> aumentada y 5 <sup>a</sup> Justa
25	1/4 de Tono entre 5 <sup>a</sup> justa y 6 <sup>a</sup> Menor
27	6 <sup>a</sup> Mayor

Fig. 1.1-2. Orden de los intervalos y su categoría. (Tomado de Hernández, 1989)

Algunas evaluaciones sobre la estructura de las escalas musicales muestran que en distintas culturas musicales humanas se le da una gran importancia a los intervalos considerados como consonancias perfectas (quintas, octavas y cuartas). (Schellenberg and Trehub, 1994) La simplicidad y complejidad de las proporciones de frecuencias ha jugado un rol central en las teorías musicales que se han creado para explicar los fenómenos de la consonancia y la disonancia de los intervalos (Helmholtz, 1877).

Más recientemente se relacionó a la consonancia con el término “armonicidad”, el cual se refiere a un espectro agregado, compuesto por la superposición de los componentes consonantes de los acordes. (McDermott, 2010) Esto se refiere sobre todo a la percepción de la consonancia en un determinado contexto musical.

Hay quienes explican la percepción de la consonancia y la disonancia dándole un gran peso a la exposición cultural a determinados tipos de música y otros que lo hacen a partir de predisposiciones biológicas. El fenómeno de la percepción auditiva de estímulos consonantes y disonantes es sumamente complejo, por lo tanto, debe explicarse y entenderse desde distintas perspectivas.

Otro modo de entender y explicar la percepción subjetiva de la consonancia y la disonancia es el modelo del oscilador acoplado de la percepción auditiva, donde una neurona o grupo de neuronas se consideran osciladores que tienen una determinada frecuencia de disparo en función del estímulo auditivo entrante y que tiende a acoplarse con otro oscilador más fácilmente si la proporción de frecuencias entre los tonos percibidos es más simple. Es decir, al parecer para el sistema nervioso es mucho más sencillo decodificar y codificar señales con frecuencias simples. Aunque las frecuencias de disparo de las neuronas, que se encuentran dentro del giro temporal superior (corteza auditiva), son versiones a baja escala de las frecuencias de los tonos que estimulan el sistema auditivo. Hay que tomar en cuenta que las neuronas tienen límites en su capacidad de emitir determinadas frecuencias, por lo que algunos estímulos incluso no podrán ser emulados por ellas (Shapira and Stone, 2008).

El fenómeno de la percepción auditiva de la consonancia y la disonancia es muy complejo y aún hay muchas preguntas en el aire. Varios grupos de investigadores han dedicado sus recursos y esfuerzos a entenderlo (Peretz et al., 2001) (Dalla Bella et al., 2001) (Perani et al., 2010) (Bidelman and Krishnan, 2009) (Itoh et al., 2010).

Escuchas ordinarios, incluidos los niños, distinguen más fácilmente combinaciones de tonos consonantes que de disonantes, considerando a las primeras como más placenteras. La preferencia a la consonancia sobre la disonancia, fue evaluada por Peretz y sus colaboradores en el 2001, en una paciente (I.R.) que sufría desórdenes de memoria y de percepción musical (amusia)

como resultado de una lesión bilateral en la corteza auditiva. Se le aplicaron dos pruebas conductuales, en la primera, I.R. escuchó extractos musicales con distintos grados de consonancia y disonancia y tuvo que valorar en una escala el nivel de agrado y desagrado de dichos extractos. En la segunda prueba, se le pidió a I.R. que valorara la valencia emocional feliz-triste de dichos extractos en una escala.

En el experimento 1, se encontró que I.R. era incapaz de distinguir entre las versiones de extractos musicales consonantes y disonantes tomados del repertorio clásico, ya que sus respuestas no establecían que la escucha de los mismos le provocaran distintos niveles de agrado y desagrado, como sí sucedió en los controles sanos. Incluso Peretz y colaboradores reportan que I.R. interrumpió la prueba varias veces indicando que no encontraba ninguna diferencia entre los distintos extractos musicales. Esta indiferencia a la disonancia no se debió a la pérdida de todas las respuestas afectivas a la música, ya que en el segundo experimento, I.R. clasificó los mismos extractos como felices o tristes, con el mismo desempeño que los controles sanos.

Un escaneo cerebral de I.R. se comparó con datos de activación obtenidos en las Tomografías por Emisión de Positrones (PET)<sup>1</sup> de voluntarios sanos y se observó que las regiones dañadas se traslapaban con las regiones identificadas que están involucradas en el análisis perceptual del estímulo musical, pero no con las regiones paralímbicas involucradas en las respuestas afectivas a éste. Estos resultados sugieren que la disonancia puede estar codificada bilateralmente en el giro temporal superior por mecanismos especiales anteriores a la interpretación emocional y que la consonancia-disonancia y el modo (mayor-menor) son fenómenos separables de la percepción del tono (frecuencia); lo que concuerda con reportes de que los niños menores de 5 años no son capaces de dar un valor emocional (feliz-triste), a extractos musicales, con base en la percepción del modo sino en el tempo (Dalla Bella et al., 2001).

Estos hallazgos apuntan a una arquitectura funcional subyacente a la interpretación emocional de la disonancia. Lo cual sugiere que la percepción y la emoción no tienen lugar en dos vías paralelas e independientes al menos en este

---

<sup>1</sup> Por sus siglas en inglés.

tipo de proceso afectivo. El que se haya encontrado especialización cerebral para la codificación de la disonancia es consistente con la visión nativista en la cual el cerebro puede estar pre-cableado para procesar los intervalos consonantes (Peretz et al., 2001).

En los adultos, sistemas neuronales específicos con ponderación hemisférica-derecha son necesarios para procesar la altura, la melodía y la armonía, así como la estructura y el significado emergentes de secuencias musicales. No se sabe hasta qué punto la especialización de estos sistemas es resultado de la exposición a la música o de límites neurobiológicos (Perani et al., 2010).

Para contestar a esta interrogante, Daniela Perani y colaboradores (2010) llevaron a cabo una investigación, con Resonancia Magnética Funcional en niños entre uno y tres días de nacidos mientras escuchaban música tonal occidental y versiones alteradas de estos mismos extractos.

La música evocó predominantemente activaciones en el hemisferio derecho en la corteza auditiva primaria. Durante la presentación de los extractos alterados, las respuestas hemodinámicas se redujeron significativamente en la corteza auditiva derecha y emergieron activaciones en la corteza frontal inferior izquierda y en estructuras límbicas. Estos resultados demuestran que el cerebro de los niños muestra una especialización hemisférica para el procesamiento de la música desde las primeras horas posnatales. Los resultados también indican que la arquitectura neural subyacente al procesamiento musical en recién nacidos es sensible a cambios en la tonalidad así como a diferencias entre consonancia y disonancia (Perani et al., 2010).

En un estudio llevado a cabo por Bidelman y Krishnan (2009), se midieron las respuestas de frecuencia del tallo cerebral en no músicos en respuesta a la presentación dicótica de 9 intervalos musicales que variaron en su grado de consonancia y disonancia. Consideraron que la actividad neural de fase-enganchada en el tallo cerebral podía preservar información relevante para atributos perceptuales tan importantes de la música como son las relaciones de consonancia y disonancia.

Las respuestas del tallo cerebral fueron más robustas para los intervalos consonantes y tuvieron una prominencia del tono neural más fuerte en comparación con los disonantes. Además, el orden de la prominencia del tono neural, a través de los intervalos musicales, siguió la disposición jerárquica del tono estipulada por la teoría de la música occidental. Finalmente, la prominencia del tono derivada de los datos neurales mostró una alta correspondencia con los juicios conductuales de consonancia (Bidelman and Krishnan, 2009).

Estos resultados sugieren que los mecanismos neurales del tallo cerebral que median el procesamiento del tono muestran una codificación preferencial de las relaciones musicales consonantes y, además, preservan las relaciones jerárquicas del tono que encontramos en la música, aún en individuos sin entrenamiento musical formal. Las relaciones básicas del tono que gobiernan la música pueden estar arraigadas en un procesamiento sensorial de bajo nivel y un esquema de codificación que favorece las relaciones de tono consonante puede ser una razón por la cual tales intervalos son preferidos conductualmente (Bidelman and Krishnan, 2009).

La consonancia de acordes individuales presentados fuera de un contexto musical, o la consonancia no contextual de acordes, es definida usualmente como la ausencia de rugosidad, que es una sensación que se percibe cuando frecuencias ligeramente desafinadas no se resuelven claramente en la cóclea (Itoh et al., 2010).

Itoh y su equipo de trabajo (2010), hicieron un estudio con potenciales evocados para demostrar que la ausencia de rugosidad no es suficiente para explicar por completo la percepción de la consonancia no contextual. Los estímulos fueron presentados en una secuencia aleatoria de varios intervalos de tonos puros (0-13 semitonos). Las activaciones cortico-cerebrales de los escuchas distinguieron estos estímulos de acuerdo con su consonancia no contextual, de una manera consistente con la práctica musical estándar, aun cuando los intervalos excedían la banda crítica (3 semitonos). Este efecto fue evidente sólo en los músicos, indicando plasticidad en los mecanismos neurales subyacentes. El modelo basado en la rugosidad, de la consonancia no contextual, no puede explicar estos resultados, ya que los intervalos

que sobrepasan los 3 semitonos, tienen niveles bajos e indistinguibles de rugosidad (Itoh et al., 2010).

Estos resultados permitieron que este grupo de investigadores propusiera un nuevo modelo de multi-componentes y multi-estados de la consonancia no contextual. Primero, mecanismos periféricos tempranos, diferencian los intervalos en términos de rugosidad independientemente del entrenamiento musical. Subsecuentemente, la rugosidad es procesada en el sistema auditivo central para dar lugar a la percepción subjetiva de la consonancia, que en este nivel de procesamiento puede estar afectada por la experiencia auditiva. Después, los mecanismos centrales, diferencian la consonancia perceptual de los intervalos basándose en otras características aparte de la rugosidad y de una manera que depende de la experiencia. Por último, la salida integrada de estos procesos neurales representa la consonancia no contextual del intervalo (Itoh et al., 2010).

La percepción auditiva de la consonancia y la disonancia es una de las funciones cognitivas complejas que se llevan a cabo cuando escuchamos y hacemos música, y que es susceptible de estudio. Sin embargo, lo que nos atañe en este estudio no es sólo la percepción, sino también la producción de dichos intervalos con la voz.

En esta investigación, utilizamos el canto de distintos intervalos musicales (consonantes y disonantes), para estudiar la cognición. El canto es una actividad multimodal sumamente compleja que requiere de la participación de distintos correlatos neurales implicados en varios sistemas cognitivos (audición, memoria, respuesta emocional, planeación y ejecución motora, auto-monitoreo, etcétera).

A pesar de que existe información suficiente y variada sobre la psicoacústica de la consonancia y la disonancia, hay poca o nula información sobre los mecanismos neurales a través de los cuales el ser humano es capaz de emitir, con la voz, intervalos musicales consonantes y disonantes. Tampoco se sabe si existe alguna diferencia en el nivel de exigencia de los mecanismos neurobiológicos que permiten llevar a cabo el canto de dichos intervalos, dependiendo de la clasificación (consonantes-disonantes) y de la longitud de los mismos.

Así mismo, también es de suma importancia conocer el papel de la retroalimentación auditiva, la planeación-ejecución motora y el entrenamiento

musical, en el perfeccionamiento de la técnica vocal y el ajuste requerido, de la emisión, al cantar intervalos musicales consonantes y disonantes de manera precisa.



## 1.2 Anatomía y producción de la voz cantada.

Es importante estudiar el canto pues la voz es producida por el cuerpo humano y no sólo cumple la función de crear los sonidos que conforman el lenguaje articulado, sino que también es un instrumento musical sumamente complejo, en el cual comúnmente están combinadas el habla y la producción de una determinada melodía que a su vez están mediadas por la retroalimentación auditiva, el contexto musical y socio-cultural en el que se desarrolla el individuo.

Los procesos de producción, retroalimentación auditiva y ajuste del tono emitido por la voz cantada tienen como base una maquinaria anatómica que permite la práctica y el desarrollo de habilidades musicales vocales. Es importante conocer su funcionamiento para entender cuáles son los desafíos que enfrenta el cerebro al coordinarla para emitir la voz cantada.

El canto es tan natural y antiguo en la historia del ser humano, que pareciera ser una práctica simple y sencilla, cuando en realidad no lo es, pues requiere del control de una maquinaria anatómica muy compleja, del acoplamiento de distintos correlatos neurales y de muchos años de práctica para perfeccionarse. Uno de los principales retos para el cantante es adquirir el conocimiento necesario para entender cómo desarrollar y mantener un set particular de conductas músico-culturales específicas usando un instrumento que no es visible y en donde sus componentes cambian físicamente a través de la vida (Parncutt and McPherson, 2002).

Los cantantes se enfrentan a grandes obstáculos para poder controlar su instrumento, es decir, su cuerpo. Un cuerpo que lleva a cabo procesos vitales sin que estos puedan ser controlados o concientizados del todo. El cuerpo del cantante tiene tres componentes fundamentales que están directamente relacionados con la producción de la voz hablada y cantada: el sistema respiratorio (Fig. 1.2-1), el aparato fonador (Fig. 1.2-2 y 1.2-3) y el sistema óseo y las cavidades, que funcionan como un sistema de resonancia (caja torácica, senos maxilares, senos nasales, senos frontales, cavidad bucal, faringe, etc.) (Fig. 1.2-4 y 1.2-5).

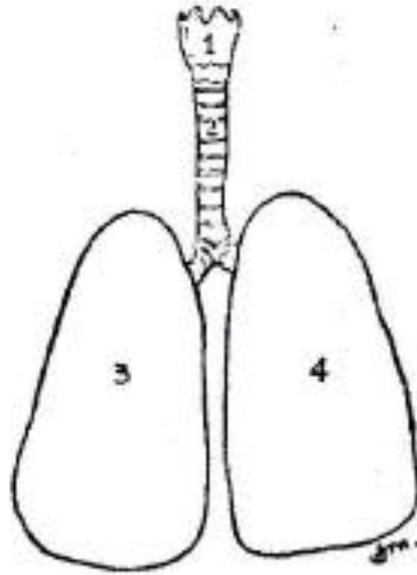


Fig. 1.2-1. Pulmones. 1) Laringe. 2) Tráquea. 3) Pulmón Derecho. 4) Pulmón Izquierdo. (Tomada de Mansion, 1947)

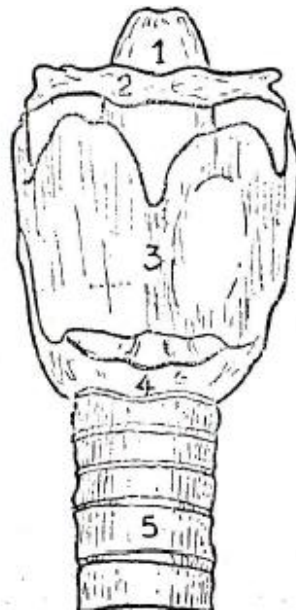
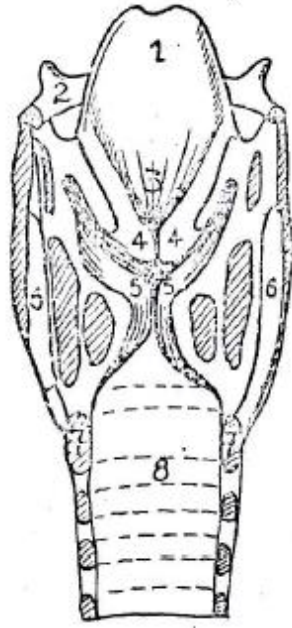
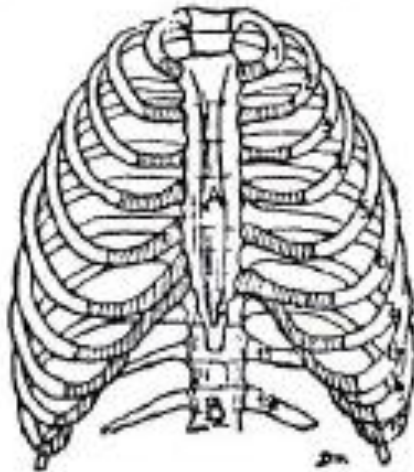


Fig. 1.2-2. Laringe (vista anterior). 1) Epiglotis. 2) Hueso Hioides. 3) Cartílago Tiroideo (Manzana de Adán). 4) Cartílago Cricoides. 5) Tráquea. (Tomada de Mansion, 1947)



*Fig. 1.2-3. Laringe (corte vertical, vista posterior). 1) Epíglotis. 2) Hueso Hioides. 3) Glotis. 4) Cuerdas Vocales Superiores o Falsas. 5) Cuerdas Vocales Inferiores. 6) Cartílago Tiroideo. 7) Cartílago Cricoides. 8) Tráquea. (Tomada de Mansion, 1947)*



*Fig. 1.2-4. Caja Torácica. A) Esternón. B) Columna Vertebral. (Tomada de Mansion, 1947)*

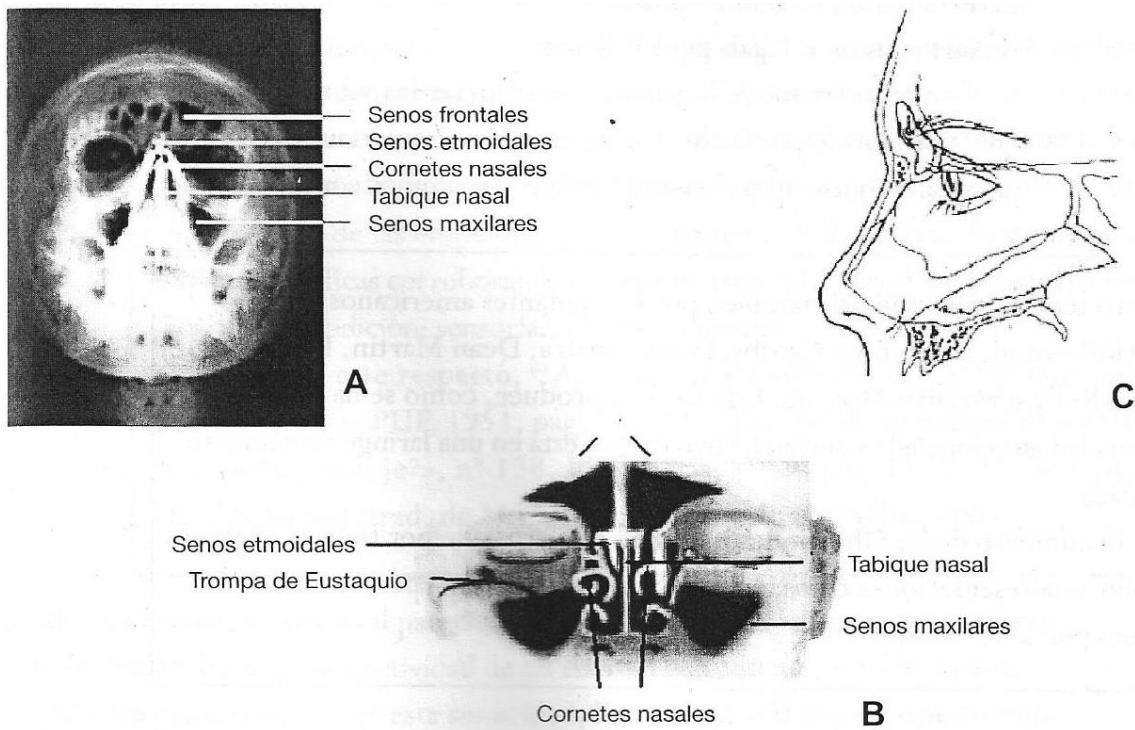


Fig. 1.2-5. Cavidades craneales. (Tomada de Ferrer, 2008)

El sistema respiratorio es responsable de las variaciones en el volumen; los cambios en el patrón y en la frecuencia de vibración de las cuerdas vocales son percibidos como variaciones en la entonación y en la calidad de la voz; y los cambios en la configuración del tracto vocal están vinculados con la resonancia. Estas tres características funcionales son susceptibles a través de la educación, de centrarse en el desarrollo y el control consciente de la voz (Parncutt and McPherson, 2002). Mas debemos recordar que este control consciente es limitado pues el cuerpo humano lleva a cabo muchas de sus funciones de manera inconsciente, por lo tanto, están fuera de nuestro control.

Todo el extremo inferior del tracto vocal es la laringe. Las dos cuerdas vocales están localizadas dentro de esta, y están hechas de músculos en forma de pliegues. (Fig. 1.2-3, 1.2-6, 1.2-7) Sus dimensiones horizontales y verticales cambian en función de la frecuencia del sonido que emiten. Las cuerdas vocales están cubiertas por una membrana mucosa, tienen una longitud aproximada de 3mm en los recién nacidos y crecen en una tasa promedio de 0.4mm por año en las mujeres y 0.7mm

por año en los hombres. Las cuerdas vocales de los hombres crecen desproporcionadamente durante la pubertad para alcanzar su máxima longitud en la edad adulta (15 a 22mm), mientras que las de una mujer adulta miden aproximadamente de 9 a 13mm (Tager-Flusberg, 1997).

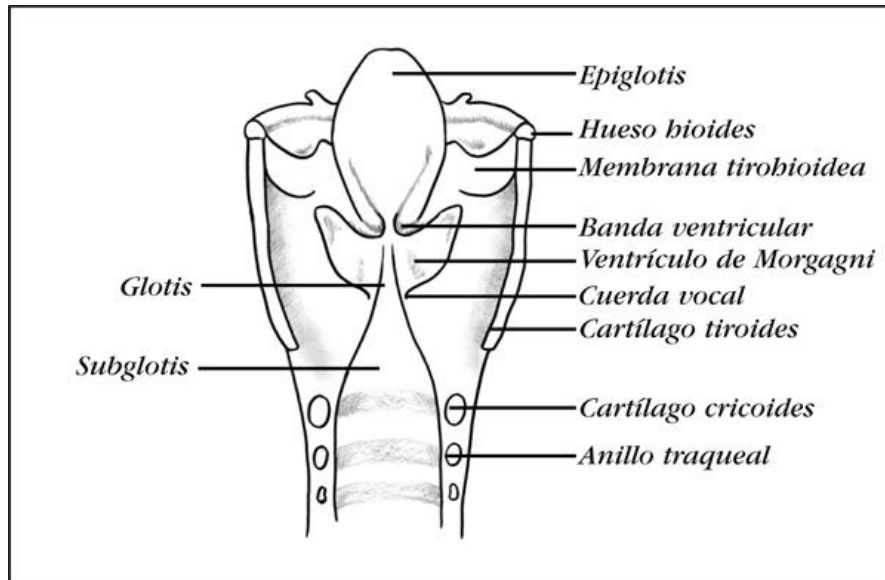


Fig. 1.2-6. Laringe seccionada en sentido cráneo-caudal. (Tomada de Arfelis, 2005)

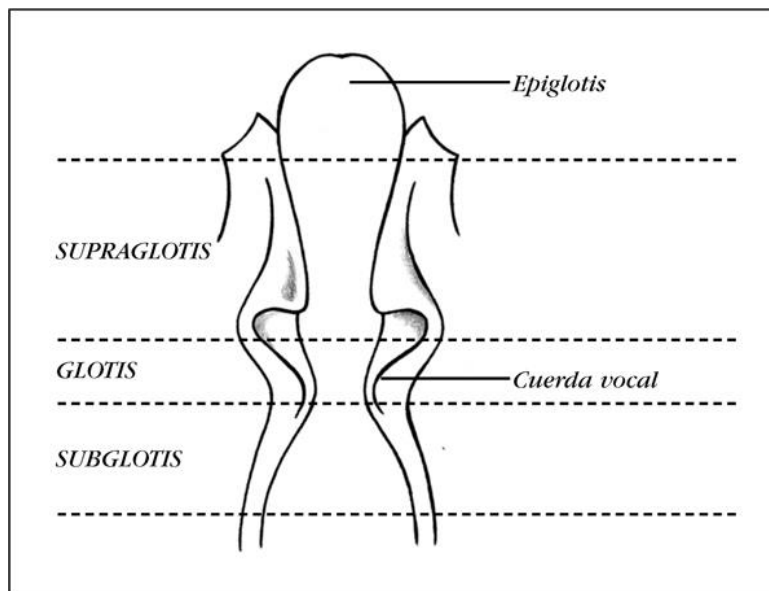


Fig. 1.2-7. Zonas de la laringe y cuerdas vocales. (Tomada de Arfelis, 2005)

La longitud y el volumen de las cuerdas vocales son cruciales para que se establezca el rango de entonación vocal (registro) de cada individuo: entre más largas y gruesas sean las cuerdas vocales, más bajo será el rango de entonación, es decir, se tendrá la capacidad de emitir sonidos más graves. Por el contrario, si las cuerdas vocales son más cortas y delgadas, el registro vocal será más agudo. Estas diferencias anatómicas establecerán la tesitura o clasificación vocal del cantante (Fig. 1.2-8).



FIG. 12. — Clasificación de las voces (aproximativa)

- 1) Soprano. 2) Soprano ligera. 3) Mezzo-soprano. 4) Contralto. 5) Tenor.  
6) Barítono. 7) Bajo.

Fig. 1.2-8. Clasificación vocal por tesitura. (Tomada de Mansion, 1947)

Las cuerdas vocales corren horizontalmente de adelante hacia atrás dentro del cartílago tiroideo y están unidas al frente de la parte interna del cartílago tiroideo y en la parte posterior de los cartílagos móviles aritenoides.

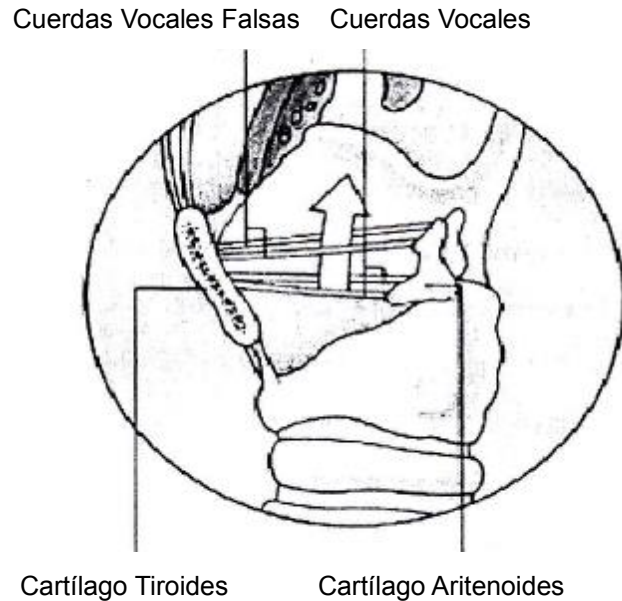


Fig. 1.2-9. Cuerdas vocales y cartílagos. (Tomada y modificada de Ferrer, 2008).

Cuando uno respira silenciosamente, estos cartílagos son separados (abducción) por el primer grupo de músculos (músculos posteriores cricoaritenoides), permitiendo que la corriente de aire fluya libremente (Fig. 1.2-11). El espacio entre las dos cuerdas es conocido como glotis.

Cuando uno está hablando o cantando, los sonidos expresados con la voz son producidos por los músculos cerradores cricoaritenoides e interaritenoides laterales que unen los cartílagos aritenoides, por lo tanto, se juntan los extremos posteriores de las cuerdas vocales (aducción) de modo que sus bordes vibran con la corriente de aire. Las cuerdas vocales son abducidas para los sonidos sordos (ciertas consonantes) y aducidas para los sonidos emitidos o expresados con la voz (las vocales) (Fig. 1.2-10, 1.2-11 y 1.2-12).

Los músculos responsables de abrir y cerrar la glotis están en su mayoría localizados en la parte posterior de las cuerdas vocales, cerca de la espina. Este conjunto de músculos que abren y cierran está también involucrado con los cambios de volumen de la voz (Parncutt and McPherson, 2002).

La presión del aire en los pulmones es la herramienta principal para variar el volumen vocal: cuando la presión es más alta, la voz es más potente. Una mayor presión del aire separa las cuerdas vocales con una mayor fuerza, lo cual da como

resultado que se cierran con mayor rapidez. Los músculos laríngeos alargadores y acortadores son esencialmente responsables del cambio en la frecuencia del tono emitido. Los músculos tiroaritenoides (acortadores) están localizados dentro de las cuerdas vocales, con una orientación horizontal (adelante hacia atrás). Cuando éstos se contraen, el resultado es un tejido vibrante voluminoso y corto que produce un tono más bajo o grave. En contraste, los músculos cricotiroides (alargadores) tienen una orientación más vertical y están cerca de la parte anterior de la laringe, estrechados entre los cartílagos tiroideos y cricoides. El efecto de la contracción de los músculos alargadores en los cartílagos es el de inclinar y deslizar el cartílago tiroideo hacia adelante y así extender (alargar) las cuerdas vocales (Fig. 1.2-10). De este modo, se plantea la frecuencia fundamental (Parncutt and McPherson, 2002).

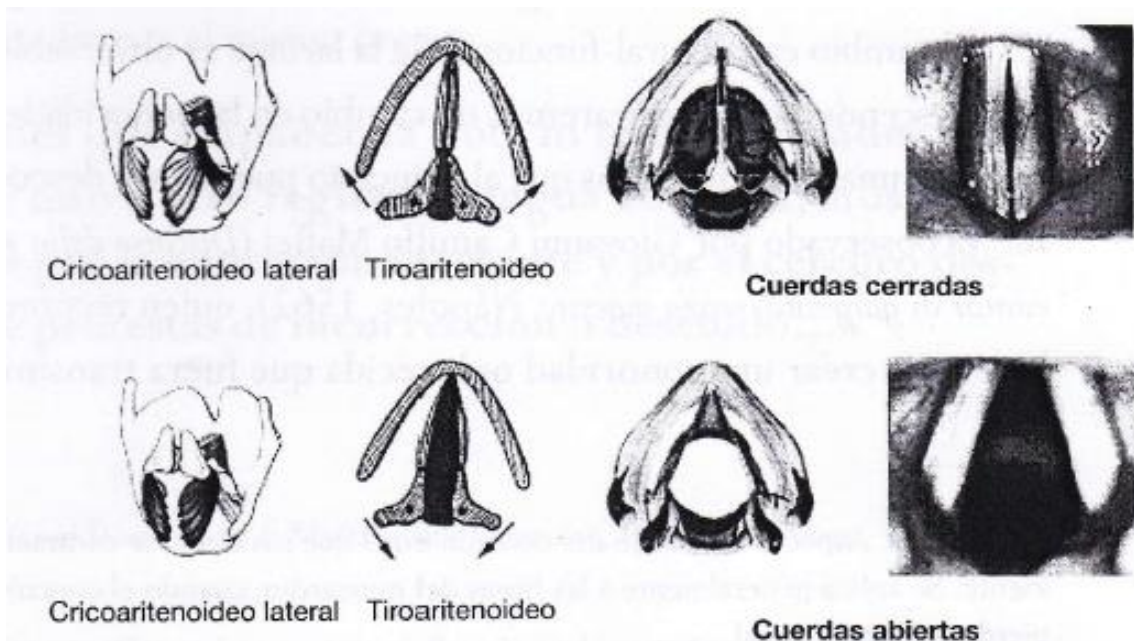


Fig. 1.2-10. Arriba: cuerdas cerradas. Abajo: cuerdas abiertas (Tomada de Ferrer, 2008).





Fig. 1.2-11. Tomografía de la laringe cuando las cuerdas vocales son abducidas al respirar (Tomada de Ferrer, 2008).

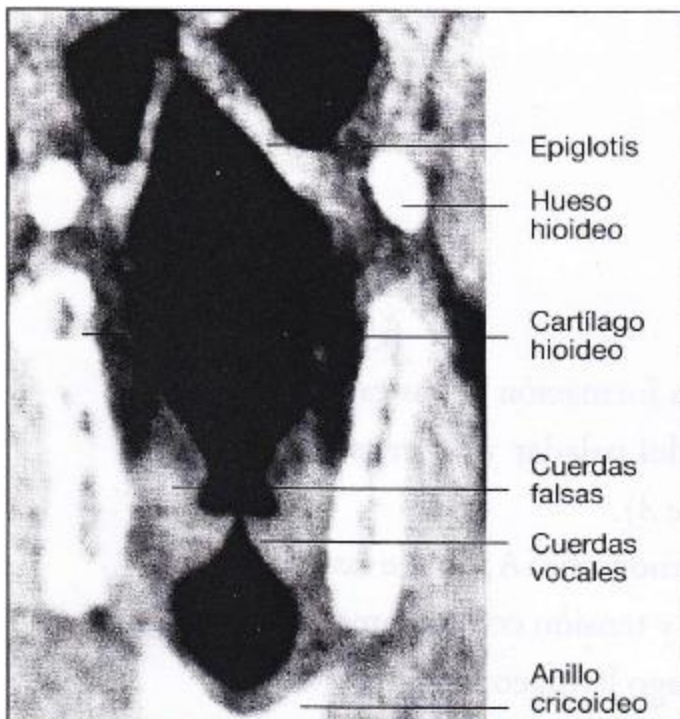
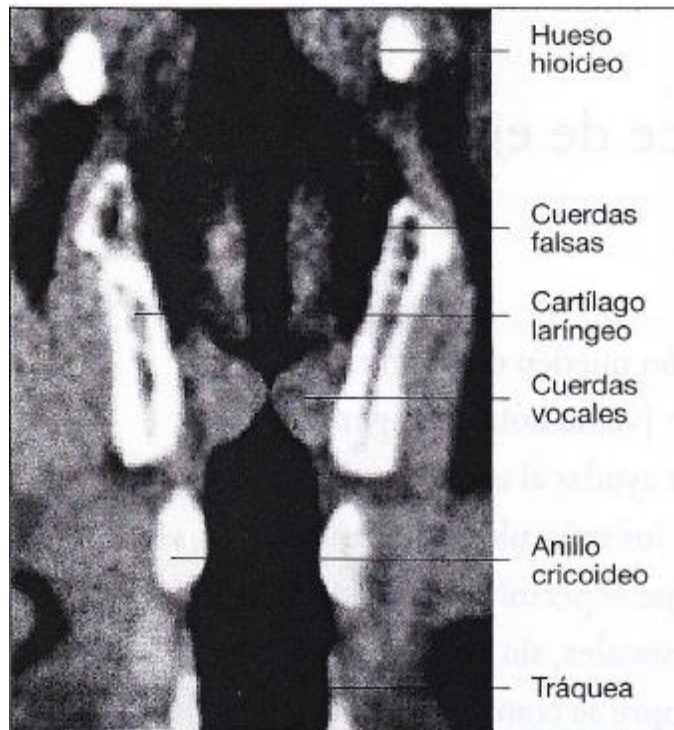


Fig. 1.2-12. Tomografía de la emisión sonora, técnicamente correcta, al cantar. Cuerdas vocales aducidas en forma de cuña, cuerdas vocales falsas inhibidas, paladar elevado, el ámbito laríngeo ensanchado y dilatado (laringe relajada) (Tomada de Ferrer, 2008).



*Fig. 1.2-13. Tomografía al emitir incorrectamente el sonido, es decir, con la laringe tensa (elevada y el paladar blando sin elevar). (Tomada de Ferrer, 2008)*

La acción coordinada efectiva de cada uno de estos músculos acortadores y alargadores es requerida para una expresión vocal experta. Cuando la voz se vuelve más fuerte, la activación de los músculos acortadores se incrementa y la de los músculos alargadores se reduce. La razón es que la corriente de aire incrementada crea un fuerte ajuste; en conjunto, las cuerdas vocales, los músculos alargadores y acortadores tienen que trabajar más para resistir la presión de los pulmones, ya sea para mantener o cambiar el tono y la intensidad (Titze, 1995). La corriente de aire fluye a través de la tráquea, que consta de un grupo de cámaras vinculadas que permiten producir timbres vocales particulares (Parncutt and McPherson, 2002).

La sobrepresión del aire en la tráquea genera una corriente a través de la glotis. La presión justo por debajo de la glotis, la presión subglótica, proporciona la fuerza impulsora de la voz y es la principal herramienta para controlar el volumen vocal: entre más alta sea la presión, más fuerte será la voz. Esta variación de presión en el canto necesita ser exacta, ya que la presión afecta el tono; cuando la presión no es la correcta puede tener como resultado errores en la frecuencia del tono emitido.

(cantar desafinado). Los tonos altos necesitan más presión que los tonos bajos, así que cada tono usualmente necesita una presión diferente. Los cantantes también usan las variaciones de la presión subglótica para demarcar los tonos en secuencias rápidas, como en el canto del legato coloratura. Entonces cada tono recibe su propio pulso de presión, de manera que un pulso de presión se produce para cada tono en sincronía con el patrón de la frecuencia fundamental. El tono percibido corresponde a la frecuencia de la vibración de las cuerdas vocales (Parncutt and McPherson, 2002).

Otro desafío que es importante mencionar es el de la entonación de notas que se encuentren en los extremos (graves y agudos) del registro del cantante. La voz hablada es la que utilizamos cotidianamente, ésta se caracteriza por utilizar un registro más pequeño que el que se utiliza al cantar, es decir, cuando hablamos, utilizamos en la mayoría de las ocasiones parte de nuestro registro medio y en muy pocas, las regiones más graves y agudas del mismo. Incluso no es común que utilicemos intervalos muy amplios o muy pequeños, por lo que la entonación de estos resulta ser, en la mayoría de las ocasiones, una actividad que requiere de mucha práctica para conseguir su perfeccionamiento.

La entonación de intervalos muy pequeños ( $1/4$  o  $1/2$  tono) o muy grandes (séptima, octava, novena, oncen... ) requiere de un oído entrenado capaz de otorgarnos una buena retroalimentación auditiva que nos permita ajustar el tono emitido en el caso de no estar afinados, buena memoria y conocimiento musical de los sonidos, y de un aparato fonador y sistema respiratorio entrenados para ajustarse rápidamente y de forma precisa a los requerimientos de llevar a cabo una actividad que no estamos acostumbrados a ejecutar cotidianamente.

La entonación de intervalos grandes, implica hacer ajustes precisos y rápidos de movimientos contrarios en el aparato fonador. Por un lado la nota grave del intervalo, requiere que las cuerdas vocales se acorten y engrosen y por el otro, la nota más aguda requiere lo opuesto (alargar y adelgazar las cuerdas vocales). También es necesario variar la presión subglótica para cada tono que se deba emitir, tener un registro unificado de preferencia (facilita el paso de las regiones grave y media a la aguda) y tener dominada la colocación de la voz, es decir, requiere de mayor control

motor del aparato fonador y del desarrollo de distintas habilidades vocales por medio del entrenamiento y comprensión de las técnicas del canto (Parncutt and McPherson, 2002) (Mansion, 1947) (Ferrer, 2008).

En el caso de los intervalos muy pequeños, el ajuste de las cuerdas vocales es mínimo y debe ejecutarse con mucha precisión. Incluso la escucha de este tipo de intervalos se dificulta en las primeras etapas del entrenamiento auditivo-musical, pues la tendencia en los estudiantes noveles es a no identificar fácilmente variaciones tan pequeñas entre dos frecuencias.

Los aspirantes a ser cantantes profesionales, en principio, deben aprender a “controlar” su respiración y a llevarla a cabo de manera relajada, profunda y concentrándose en que sea generalmente una respiración diafragmática o abdominal. Aunque se conocen también la respiración costoabdominal y la respiración pectoral, pero se considera más apropiada para el canto la abdominal porque favorece la relajación y el control de los músculos abdominales para variar el volumen y el color vocal (Mansion, 1947) (Ferrer, 2008)(Beltrán, 1949).

Dado que anatómicamente los pulmones no contienen masa muscular en su estructura, se requieren de músculos ajenos a está, para que ocurra la expansión y contracción de los mismos al inhalar y al exhalar. El diafragma es un músculo crucial para que esto se lleve a cabo. Este se encuentra situado por debajo de la caja torácica. Cuando se inhala, el diafragma hace movimientos hacia abajo y adelante, y al exhalar, regresa paulatinamente a su posición original empujando los pulmones y el contenido de la caja torácica hacia arriba (Fig. 1.2-14).

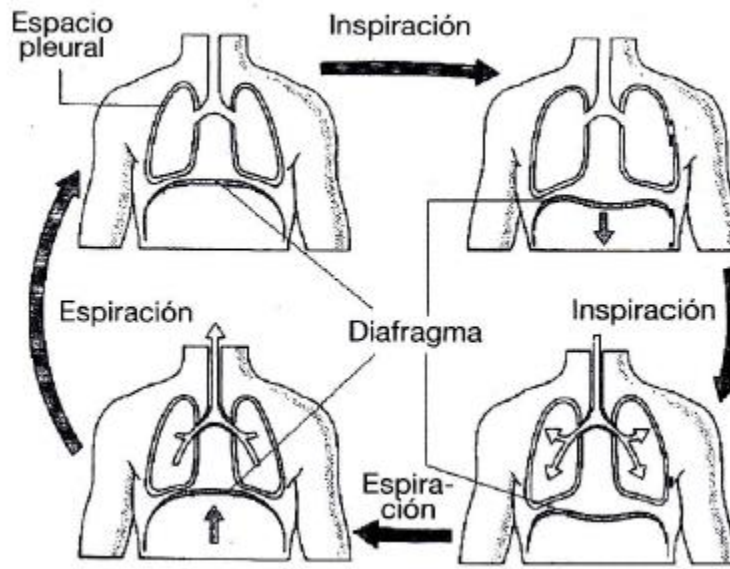


Fig. 1.2-14. Movimientos del diafragma al inhalar y exhalar. (Tomada de Ferrer, 2008)

Para controlar la presión subglótica del aire, es necesario mantener el diafragma “tenido” (apoyar la voz), es decir, mantenerlo en una posición baja y anterior la mayor cantidad de tiempo posible. Esto se logra con la ayuda del movimiento hacia adelante de la pelvis, la ligera flexión de las rodillas y de la contracción de los músculos abdominales al emitir la voz cantada. Cuando se inhala, la pelvis regresa a su posición original y los músculos abdominales se relajan.

El cantante debe aprender una técnica vocal que le permita mantener la laringe completamente relajada, una postura correcta (equilibrada) y libre de tensiones innecesarias al cantar<sup>2</sup> (Fig. 1.2-15), a apoyar su voz y a utilizar sus resonadores naturales para tener mayor control sobre la calidad, intensidad, duración y afinación de los sonidos emitidos.

<sup>2</sup> Columna vertebral recta, los brazos a los costados sin tensar los hombros, el cuello o la espalda; la mirada debe dirigirse hacia adelante, sin pegar la barbilla al pecho o levantar la cabeza como si se mirara hacia arriba en un ángulo de 45° o más. Debe haber una relación equilibrada entre el pecho y el maxilar inferior, es decir, que no se pegue la barbilla al pecho ni se aleje demasiado del mismo. Levantar la cabeza y tensar el cuello, tal vez, con la intención de “abrir la boca” sin bajar el maxilar inferior, es un error muy común en los principiantes, personas muy nerviosas y/o inseguras o con pánico escénico. Los pies deben estar alineados con la cadera, con el compás ligeramente abierto y evitar que las puntas de los mismos apunten hacia los costados.

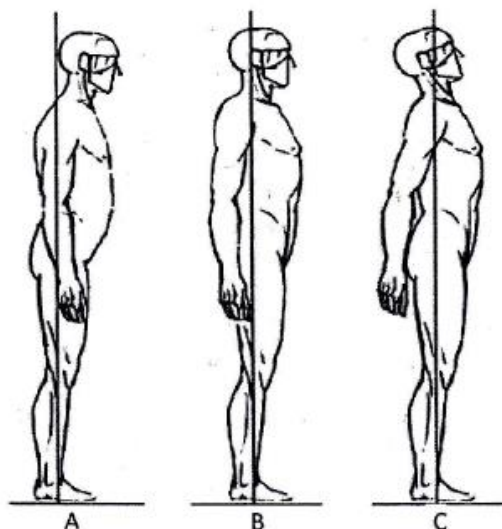
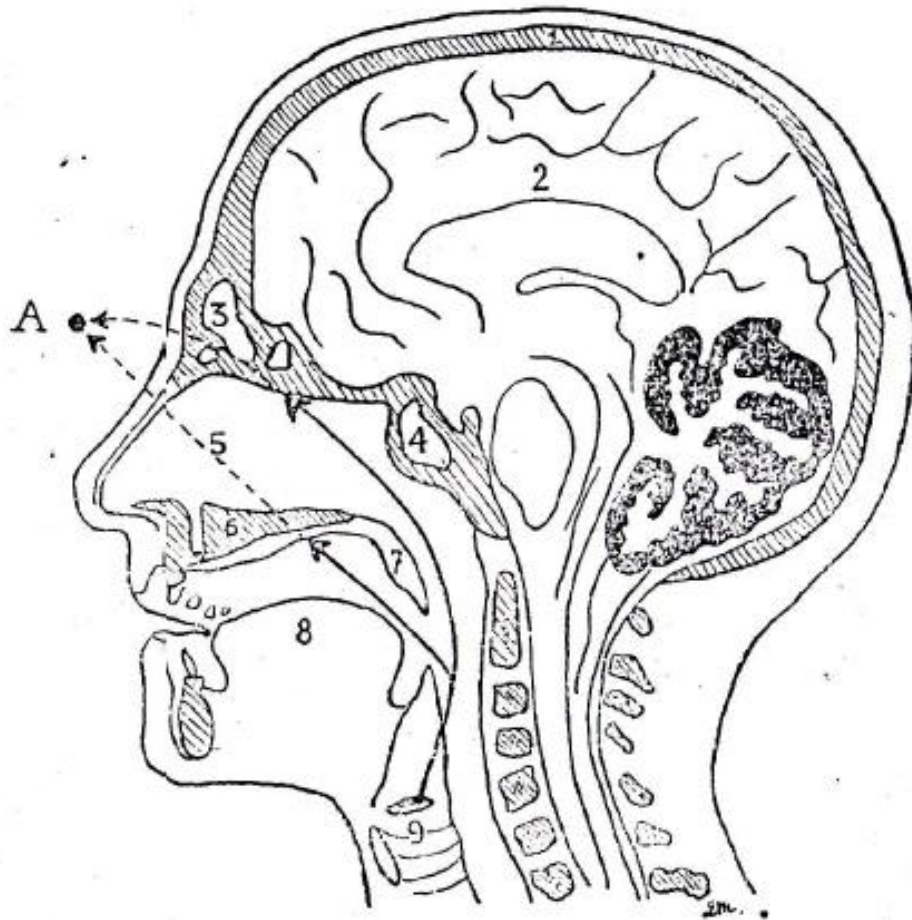


Fig. 1.2-15. A) Espaldas caídas, vientre salido. B) Equilibrado. C) Tenso, envarado. (Tomada de Ferrer, 2008)

Una vez llevada a cabo la respiración diafragmático-abdominal, el apoyo y la producción relajada de la voz (por el aparato fonador), es necesario dirigir el sonido hacia los resonadores (colocar la voz) para que los componentes armónicos del sonido resuenen. Cuando la voz se dirige hacia el pecho, se privilegian los armónicos más graves, perdiéndose el brillo y flexibilidad de esta, obteniéndose como consecuencia un sonido más oscuro e incrementándose la dificultad para pasar de los sonidos graves a los sonidos agudos y viceversa. De esta manera se corta el registro vocal, como si el cantante tuviera dos voces distintas, una para los sonidos graves y otra para los sonidos agudos. Esto debe evitarse pues es un error muy común de la gente, con poco o nulo entrenamiento vocal o en estudiantes de canto poco experimentados (Mansion, 1947) (Ferrer, 2008).

Cuando la voz se dirige hacia la “máscara” (resonadores de la cabeza), se enfatizan los armónicos más agudos, obteniéndose una voz flexible y brillante, y la unificación del registro vocal. Cuando se producen sonidos graves, el pecho (esternón) vibra de manera natural, pero conforme se incrementa la frecuencia del sonido o sonidos emitidos, vibran más los resonadores de la cabeza. (Mansion, 1947) (Ferrer, 2008) No es sencillo aprender a “colocar” la voz, toma años de entrenamiento el hacerlo de manera correcta y esto se consigue con varios

métodos, ya sea con visualizaciones diversas (dirigir la voz hacia un punto imaginario que está frente a los ojos) (Fig. 1.2-16) o con ejercicios específicos de vocalización, como por ejemplo: cantar las primeras 5 notas de la escala mayor con las letras “m” o “n”, entre otros.



*Fig. 1.2-16. Llegada del sonido a los resonadores faciales. 1) Cráneo. 2) Cerebro. 3) Senos Frontales. 4) Seno Esfenoidal. 5) Fosas Nasales. 6) Paladar. 7) Velo del Paladar. 8) Lengua. 9) Cuerdas Vocales (Laringe). A) Punto al que debe tenerse la impresión de enviar el sonido. (Tomada de Mansion, 1947)*

Todo lo anterior se refiere solamente al proceso de aprendizaje y entrenamiento en relación a adquirir una correcta técnica vocal al cantar. Mas también es sumamente importante el entrenamiento auditivo, rítmico, melódico, armónico y cultural, y el conocimiento teórico para que el cantante se especialice en el género o géneros musicales deseados, tenga una nutrida retroalimentación auditiva para incrementar su capacidad de corrección de errores al emitir la voz cantada e incluso

pueda componer e improvisar melodías dentro de un contexto armónico determinado o sin él.

Una diferencia importante entre la producción de la voz hablada y la voz cantada es que en la segunda, los sonidos tienden a ligarse y se alarga la entonación de las vocales, en comparación con algunas consonantes que se consideran sonoras, como “l”, “m”, o “n”.

Otras características del canto son la búsqueda de la emisión relajada sin escape de aire; la dicción, pronunciación, entonación y afinación casi perfectas; una excelente interpretación emocional e incluso la construcción de personajes congruentes con la letra de las canciones que permitan otorgarle mayor sentido y fuerza a cada una de las frases que la componen (Mansion, 1947) (Ferrer, 2008)(Beltrán, 1949).



### 1.3 Correlatos neurales del canto.

Además del control de la maquinaria anatómica, previamente descrita, para producir la voz hablada y la voz cantada, es necesario un acoplamiento muy complejo de diferentes zonas del cerebro por lo cual en los últimos años se han realizado investigaciones para estudiar y entender este fenómeno.

Perry y colaboradores (1999) se interesaron en la localización de la actividad cerebral durante el canto simple, ya que el canto es una habilidad humana altamente evolucionada, que involucra control voluntario de la frecuencia fundamental vocal. En el canto (sin palabras), este control está crucialmente dirigido hacia los movimientos de las cuerdas vocales responsables del control de la frecuencia fundamental (Titze, 1994).

El patrón de la variación de la frecuencia fundamental en las expresiones cantadas está típicamente organizado alrededor de series de segmentos que están relativamente en estado estacionario (o notas), cada uno de los cuales expresa un tono musical único. En contraste con otros primates, el control lateral de la vocalización es mucho más prominente en los humanos (Perry et al., 1999).

Este estudio consistió en aplicar Tomografía por Emisión de Positrones (PET) a 13 voluntarios diestros sin entrenamiento substancial en la ejecución musical (un año en promedio) o entrenamiento formal vocal y en teoría musical, que incluyeron a seis hombres y siete mujeres. Las condiciones de escaneo fueron dos, la primera (canto) consistió en vocalizar un tono único repetidamente; la segunda, en una condición auditiva base (percepción), donde escuchaban tonos complejos presentados en una tasa similar y transportados a un rango de frecuencia igual (Perry et al., 1999).

Los resultados de la medición del flujo sanguíneo cerebral (CBF)<sup>3</sup> al cantar un tono único y una vocal en contraste con la escucha pasiva de tonos complejos, fueron: el CBF incrementó en las áreas corticales relacionadas con el control motor, fue visto en el área suplementaria motora, la corteza del cíngulo anterior, el giro precentral, la ínsula anterior (y la cara interior del opérculo precentral) y el cerebelo,

---

3 Por sus siglas en inglés “Cerebral Blood Flow”.

regiones que se activan comúnmente durante el habla. Los incrementos en la corteza auditiva fueron vistos dentro del giro derecho de Heschl, y en el plano temporal superior posterior (y en la corteza parietal suprayacente inmediata). Dado que la corteza cercana al giro de Heschl derecho fue relacionada con la percepción de tonos complejos, su activación asimétrica puede relacionarse con el análisis de la frecuencia fundamental de la propia voz para una modulación guiada por la retroalimentación (Perry et al., 1999).

Zarate y Zatorre (2008) llevaron a cabo una investigación donde se plantea la importancia de la retroalimentación auditiva en el control del sistema motor vocal al cantar. Dichos investigadores observaron que el canto apropiado requiere de la integración de los mecanismos de retroalimentación auditiva con el sistema motor vocal, para que la entonación vocal pueda ser controlada de manera precisa.

Para determinar los sustratos neurales involucrados en la integración audio-vocal, cantantes experimentados y no músicos se sometieron a una exploración de resonancia magnética funcional (fMRI) mientras cantaban un tono único con retroalimentación auditiva en su estado natural o con el tono modificado. En los ensayos con tono modificado, se instruyó a los sujetos para que ignoraran el tono modificado y continuaran cantando el primer tono que escucharon, o para que ajustaran su entonación al tono modificado.

La fMRI es una técnica que ha ido adquiriendo importancia en la investigación cognitiva, dado que no es invasiva, tiene mejor resolución temporal y espacial que otras técnicas (Armony, 2012); y, nos permite relacionar la actividad cerebral con los procesos mentales y cognitivos.

La obtención de imágenes cerebrales anatómicas y funcionales por medio de esta técnica, es posible, dado que al llevarse a cabo una determinada tarea, hay cambios en la actividad de distintas zonas cerebrales. Lo cual hace que las neuronas requieran de mayor energía. Dicha energía les es proveída por la hemoglobina oxigenada a través del sistema vascular, causando así que disminuya la cantidad de desoxihemoglobina<sup>4</sup> e incremente la de oxihemoglobina (Armony, 2012).

---

<sup>4</sup> Hemoglobina sin oxígeno.

Gracias a que la oxihemoglobina y la desoxihemoglobina tienen distintas propiedades magnéticas<sup>5</sup>, el resonador capta las diferencias que se dan en la actividad cerebral cuando se está en reposo o cuando se lleva a cabo una determinada tarea. La diferencia puede tomarse como un índice indirecto del grado de actividad neuronal en esa área. Esta señal se conoce como contraste BOLD y cuando un cambio se da gracias a un cambio breve en la actividad neuronal, este se conoce como Respuesta Hemodinámica.

La hipótesis de Zárate y Zatorre (2008) era que la corteza del cíngulo anterior (ACC), el giro temporal superior (STG), y la ínsula anterior podrían estar involucradas en la integración audio-vocal, gracias a sus roles funcionales durante el canto y a su conectividad anatómica. Aunque los cantantes fueron más exactos en el canto simple que los no músicos, ambos grupos reclutaron redes funcionales similares (cortezas auditivas bilaterales, cortezas motoras primarias bilaterales, área suplementaria motora (SMA), ACC, ínsula y cerebelo). Los cantantes ignoraron la retroalimentación auditiva modificada mejor que los no músicos, y ambos grupos incluso mostraron distintos patrones de actividad neural para esta tarea: los cantantes reclutaron las áreas auditivas bilaterales y el putamen izquierdo, mientras que los no músicos reclutaron el giro supramarginal izquierdo y la corteza motora primaria (Zarate and Zatorre, 2008).

Mientras que no hubo diferencias significativas entre los grupos en el desarrollo de las tareas de compensación, los cantantes mostraron actividad incrementada en ACC, el surco temporal superior y el putamen, mientras que los no músicos exhibieron actividad incrementada en la corteza premotora dorsal, una región involucrada con las interacciones sensoriomotoras. Por lo cual propusieron dos sustratos neurales para la integración audio-vocal: la corteza premotora dorsal puede actuar como una interface básica, pero con entrenamiento vocal y práctica, la ACC, las cortezas auditivas, y el putamen pueden ser reclutados cada vez más cuando la gente aprende a monitorear su retroalimentación auditiva y a ajustar su “salida” vocal, en consecuencia (Zarate and Zatorre, 2008).

---

<sup>5</sup> Las propiedades magnéticas de la molécula de la hemoglobina dependen de su unión con el oxígeno. La oxihemoglobina es ligeramente diamagnética y la hemoglobina sin oxígeno es paramagnética.

Kleber y colaboradores (2007) realizaron una primera investigación sobre el canto imaginado y abierto de un aria italiana por cantantes profesionales de ópera. En ésta se obtuvieron imágenes cerebrales, por medio de resonancia magnética funcional (fMRI) y un método de muestreo escaso (Gaab et al., 2003), al llevarse a cabo una tarea de canto imaginado y una de canto abierto, por 16 cantantes clásicos diestros (4 coristas de ópera, 4 solistas y 8 estudiantes de canto).

Para poder obtener dichas imágenes, tuvieron que enfrentarse a las limitantes que se tienen en cuanto a la adquisición de imágenes cerebrales mediante la técnica de resonancia magnética funcional. El espacio es reducido, se tienen que limitar los movimientos de la cabeza y la mandíbula, y controlar los cambios en los espacios de aire durante la vocalización y la articulación (Smith and Zelaznik, 2004).

Las tareas de interés fueron las de canto imaginado y abierto, y las tareas base fueron periodos de respiración profunda. Las primeras consistieron en pedir a los sujetos de estudio que cantaran imaginariamente las primeras 6 frases de “Caro mio ben”, y la segunda, en cantar dichas frases (canto abierto).

Dichos investigadores reportaron que el canto abierto involucró las cortezas sensoriomotoras y auditivas bilaterales primaria y secundaria, pero también áreas asociadas con el habla y la producción del lenguaje (Broca, Wernicke y sus homólogos). La magnitud de activación en el giro de Heschl (A1) fue comparable en ambos hemisferios. El cerebelo y las áreas subcorticales motoras (tálamo, médula y ganglios basales) se activaron también. Las áreas asociadas con el procesamiento emocional (corteza del cíngulo anterior, ínsula anterior) mostraron una leve activación (Kleber et al., 2007).

Los sitios de activación cerebral durante el canto imaginado se centraron en las áreas fronto-parietales e involucraron a las áreas sensoriomotoras primaria (M1/S1) y secundaria (SII) en ambos hemisferios. Las áreas que procesan las emociones mostraron una intensa activación (ACC y la ínsula bilateral, el hipocampo y los polos temporales anteriores, la amígdala bilateral). La creación de imágenes no mostró una activación significativa en A1. El contraste entre canto abierto menos el imaginado reveló una activación incrementada en las áreas corticales (motora primaria bilateral: M1) y subcorticales (hemisferio derecho cerebelar, médula espinal

superior o mesencéfalo) así como en áreas sensoriales (corteza primaria somatosensorial, A1 bilateral) (Kleber et al., 2007).

El contraste entre canto imaginado menos el abierto mostró una activación incrementada en el área medial 6 de Broadmann, en la corteza prefrontal (PFC) ventro lateral y medial, la corteza del cíngulo anterior y el lóbulo parietal inferior. Adicionalmente, las áreas de Wernicke y Brocca incrementaron su activación durante la creación de imágenes mentales. El canto imaginado y abierto, involucran en parte, diferentes sistemas cerebrales en cantantes profesionales con mayor activación prefrontal y límbica, y una cadena más larga de las funciones de alto orden asociativo durante la creación de imágenes mentales (Kleber et al., 2007).

En una segunda investigación de Kleber y colaboradores (2009) se obtuvieron las áreas del cerebro que se encuentran directamente relacionadas con el canto de un aria (Caro mio ben), y se centró la atención en los cambios, dependientes de la experiencia, que se dan en la morfología cerebral y la reorganización de las cadenas neurales. Es decir, se observaron las diferencias entre lo que sucede en el cerebro de personas sin entrenamiento vocal (canto), estudiantes de canto y cantantes de ópera profesionales de amplia trayectoria cuando cantan un aria italiana. 49 sujetos diestros, sin antecedentes de enfermedad psiquiátrica o neurológica, participaron en el estudio. 10 cantantes profesionales de ópera, 21 estudiantes de canto y 18 personas sin entrenamiento.

Fue importante llevar a cabo dicho estudio por la complejidad de los procesos implicados en el canto. La producción vocal en el habla y el canto requiere de un involucramiento de más de 100 músculos; el sistema motor vocal es diferente en naturaleza del sistema motor periférico; todo sonido es producido y amplificado enteramente dentro del cuerpo en ausencia de control visual sobre el movimiento; las fibras musculares de la laringe y el sistema orofacial son distintos de los sistemas musculares esqueléticos (Herdener et al., 2010). Estos músculos son heterogéneos y únicamente diseñados para la rápida así como variable contracción y resistencia a la fatiga, que habilita a las cuerdas vocales a abrirse y cerrarse varios cientos de veces con una interacción coordinada precisa de la actividad de la laringe, la respiración y la articulación (Dejonckere and Lebacq, 1981).

La actividad motora vocal es rutinariamente ejecutada en una forma discretamente más rápida que cualquier otra conducta humana, la cual requiere de un sistema de control que principalmente dependa de un sistema flexor intrínseco (Abo-el-Enein, 1966). Estas complejas operaciones se piensa que son automáticas en el habla cuando el proceso de desarrollo se completa (Stager et al., 2003).

Los cantantes deben construir sus habilidades en un sistema ya altamente involucrado con respecto a la función motora en el canto, mientras que los instrumentistas desarrollan enteramente nuevos patrones motores. Sin embargo, los cantantes deben entrenar intensamente por años antes de poder ajustar su sistema vocal lo suficiente para conseguir las típicas características sonoras y el nivel de control vocal requerido en el canto clásico (Stager et al., 2003).

Kleber y colaboradores (2009) utilizaron el método de muestreo escaso y la resonancia magnética funcional. Se tuvieron dos tareas, una de canto y una tarea control, compuesta por periodos de inhalación y exhalación. Los resultados reportados son la primera evidencia de que el entrenamiento de las capacidades vocales está acompañado por una activación incrementada de la corteza somatosensorial primaria bilateral, representando a los articuladores y la laringe.

Los cantantes de ópera mostraron una activación adicional en la corteza sensoriomotora primaria derecha. La activación adicional relacionada con el entrenamiento comprendió el lóbulo parietal inferior y la corteza prefrontal dorsolateral bilateral. A nivel subcortical, los cantantes expertos mostraron una activación incrementada en los ganglios basales, el tálamo y el cerebelo.

Un análisis de regresión de la activación funcional con acumulación de la práctica del canto confirmó que el entrenamiento de las capacidades vocales está correlacionada con una actividad incrementada de la cadena cortical para mejorar el control kinestésico motor y la guía sensoriomotora junto con un involucramiento incrementado implícito de las áreas de memoria motora en los niveles subcortical (ganglios basales) y cerebelar. Estos descubrimientos pueden tener ramificaciones para la rehabilitación de la voz y la práctica deliberada de otras capacidades motoras que requieran interocepción (Kleber et al., 2009).

Las aportaciones de estas investigaciones han permitido establecer el “circuito del canto”, es decir, las zonas de cerebro humano que se espera se recluten al momento en que se lleva a cabo una vocalización humana (canto y habla).

Las zonas implicadas en el control motor vocal son la corteza motora primaria (M1), la corteza cingulada anterior (ACC), los ganglios basales, el tálamo, y el cerebelo; para procesar la retroalimentación somatosensorial, se reclutan, las cortezas sensoriales primaria y secundaria (S1 y S2); para el procesamiento de la retroalimentación auditiva, las regiones corticales auditivas bilaterales (corteza auditiva primaria (A1) y una región sensible al tono dentro del giro de Heschl, varias porciones del giro temporal superior (STG) y el surco temporal superior (STS)); y, la ínsula durante el procesamiento multimodal de la retroalimentación sensorial y la integración audio-motora, además de áreas premotoras como la corteza premotora dorsal (PMd) y parietales como el surco intraparietal (IPS) (Zarate 2013) (Kleber et al., 2007) (Kleber et al., 2009) (Zarate and Zatorre, 2008) (Perry et al., 1999).

Jean Marie Zarate (2013), propone que la retroalimentación sensorial y la retroalimentación auditiva están integradas con el control motor vocal y que existe un circuito de control audio-motor para el canto.

Para controlar la calidad y afinación de la emisión vocal al cantar, la retroalimentación sensorial y la retroalimentación auditiva son cruciales. Los seres humanos somos capaces de tener respuestas de emisión vocal a un cambio de tono en la retroalimentación auditiva. Dichas respuestas son de dos tipos: tempranas (100-150ms después del cambio de tono) y tardías (250-600ms) (Zarate 2013).

Las respuestas tempranas a un cambio de tono se considera que se dan de manera inconsciente, pero las tardías son aquellas que pueden producirse bajo el control voluntario del escucha (cambiar el tono emitido hacia la dirección opuesta del cambio de tono escuchado o seguir la misma dirección del cambio de tono escuchado, con la voz) (Zarate 2013).

Se requiere de varios circuitos neuronales para controlar de forma precisa los músculos necesarios para llevar a cabo la fonación. El primer circuito está formado por la corteza cingulada anterior (ACC) y el gris periacueductal del mesencéfalo (PAG); el segundo, por la corteza primaria motora (M1) y dos cadenas subcorticales

(putamen, globo pálido, gris pontino y el cerebelo). Estas cadenas subcorticales modulan los comandos de control motor vocal desde M1 y mandan programas motores modificados por medio del tálamo ventrolateral de vuelta hacia M1 (Zarate 2013).

La formación reticular del puente y el mesencéfalo tienen conexión directa con las motoneuronas de los músculos implicados en la fonación y reciben entradas excitatorias de los dos circuitos neuronales mencionados (Zarate 2013).

Zarate explica que estos circuitos conforman una jerarquía tripartita del control motor vocal. El primer nivel está conformado por la formación reticular, en donde se generan los patrones vocales completos; el segundo, por ACC y PAG, a los cuales se les atribuye la iniciación voluntaria y el control emocional-motivacional de la vocalización; y, el tercero, por M1 y algunas regiones modulatorias del cerebro con las cuales está relacionada la corteza motora primaria (putamen, globo pálido, gris pontino y el cerebelo) (Zarate 2013).

Zarate considera que existe un circuito del control audio-motor dependiente de la experiencia que está conformado por el pSTS (corteza auditiva), el surco intraparietal (IPS), la ínsula anterior y la corteza cingulada anterior (ACC). Este circuito se recluta en mayor o menor medida, dependiendo del nivel de práctica y entrenamiento vocal (Zarate 2013).

Basándose en todo lo anterior, Zarate propone un “circuito general para la producción del habla y el canto” conformado por el pSTS e IPS para el procesamiento auditivo y la transformación de una salida motora, S1 para el procesamiento somatosensorial, la ínsula anterior para la integración audio-motora y para posibilitar la retroalimentación somatosensorial; y, PMd y ACC, regiones premotoras para la preparación motora-vocal y la iniciación de una respuesta (Zarate 2013).



#### 1.4 Cambios cerebrales (funcionales y estructurales) dependientes del entrenamiento musical.

Halwani y colaboradores (2011) mencionan que la estructura y la función del cerebro humano son afectadas por el entrenamiento musical en ambos dominios, el musical y el lingüístico. Los individuos con entrenamiento musical vocal proveen un modelo útil para la investigación de adaptaciones neurales del aprendizaje en el dominio motor-vocal y esto puede ser comparado con el entrenamiento en un dominio musical más general (Halwani et al., 2011).

Estos investigadores llevaron a cabo un estudio donde confirmaron diferencias generales en la macroestructura (volumen del tracto) y la microestructura (anisotropía fraccional (FA), del fascículo arqueado (AF)), en un tracto prominente de materia blanca que conecta las regiones cerebrales frontal y temporal, entre cantantes, instrumentistas y no músicos. Ambos grupos de músicos difirieron de los no músicos, teniendo un mayor volumen del tracto y valores más altos de FA del AF izquierdo y derecho, también mostraron mayor volúmenes del tracto en los tractos dorsal y ventral derechos comparados con los no músicos, pero no mostraron diferencias significativas entre ellos. Además, en los cantantes la FA en la rama dorsal izquierda del AF se correlacionó inversamente con el número de años de entrenamiento vocal de los participantes (Halwani et al., 2011).

Estos descubrimientos sugieren que el entrenamiento vocal motor de largo plazo podría conducir a un incremento en el volumen y la organización microestructural de tractos específicos de materia blanca que conectan regiones que son fundamentales para la percepción del sonido, la producción, y el control de la alimentación hacia adelante y la retroalimentación que puede ser diferenciado de un efecto del entrenamiento musical, más general (Halwani et al., 2011).

También se ha observado que el entrenamiento de la voz cantada puede ayudar a disminuir problemas de lenguaje que caracterizan a algunos padecimientos neurológicos como el autismo, la enfermedad de Parkinson y el tartamudeo, entre otras (Wan et al., 2010) (Chan et al., 2009) (Wiens et al., 1999) (Prasse and Kikano,

2008) (Healey et al., 1976) (Andrews et al., 1982) (Colcord and Adams, 1979) (Davidow et al., 2009) (Soltysik and Hyde, 2006) (Williams et al., 2006).

La creación musical es una actividad multimodal que involucra la integración de procesos sensoriomotores. La habilidad de cantar en los humanos es evidente desde la infancia y no depende del entrenamiento vocal formal, pero puede mejorarse con el mismo. Dadas las similitudes conductuales entre el canto y el habla, así como los correlatos neurales distintos y compartidos de ambos, los investigadores han empezado a examinar si el canto puede ser utilizado para tratar algunas anormalidades motoras del habla asociadas con varias condiciones neurológicas (Wan et al., 2010).

Muchos estudios han mostrado que escuchar música puede incrementar el funcionamiento emocional y cognitivo de pacientes afectados por varias condiciones neurológicas (Chan et al., 2009).

Al igual que el hecho de escuchar música, la creación musical coloca demandas adicionales en el sistema nervioso, dando lugar a un fuerte enganche de percepción y acción; procesos que están mediados por regiones motoras, sensoriales e integrativas multimodales distribuidas a lo largo del cerebro. Esta red fronto-temporo-parietal se superpone con componentes del sistema putativo de neuronas espejo, que es importante para la percepción y la ejecución de acciones motoras. Las deficiencias en la ejecución de acciones motoras y articuladoras son síntomas de muchos desórdenes neurológicos (Wan et al., 2010).

El canto, o el acto de producir sonidos musicales con la voz, tiene el potencial para tratar anormalidades del habla porque estimula directamente la musculatura asociada con la respiración, la fonación, la articulación y la resonancia. El acto de cantar involucra inspiraciones relativamente fuertes y rápidas, seguidas por expiraciones extendidas y reguladas. El canto requiere que la respiración sea regulada para sostener las notas. Éste también resulta en una intensidad (Thurman and Welch, 2000) y control vocal incrementados (Natke et al., 2003) en comparación con el habla. Además, se ha sugerido que el canto incrementa la fuerza de la musculatura respiratoria (Wiens et al., 1999).

El tartamudeo es una condición en gran medida de la etapa de desarrollo que afecta la fluidez del habla. Se caracteriza por la repetición de palabras o partes de palabras, así como también de prolongaciones de sonidos del habla, resultando en interrupciones en el flujo normal del habla. Esta condición ocurre de manera más común en los niños pequeños, mientras están desarrollando sus habilidades del habla y del lenguaje. El tartamudeo puede persistir en la edad adulta: aproximadamente el 1% de los adultos continua estando afectado por esta condición (Prasse and Kikano, 2008).

Muchos tratamientos existentes se han enfocado en enseñar a los individuos tartamudos vías para producir un habla más fluida, utilizando condiciones de "incremento de la fluidez". Se ha identificado que el canto, en particular, tiene un potencial terapéutico importante, y la investigación ha proveído de evidencia en favor de esta aproximación para incrementar la fluidez entre individuos que tartamudean (Wan et al., 2010).

Healey, Mallard y Adams examinaron si el canto podría reducir el tartamudeo y si fuera así, si la familiaridad con las letras de canciones conocidas jugaría un rol en producir tal efecto. Se les pidió a los participantes que leyeran o cantaran las letras de canciones conocidas con texto original o alterado. La reducción del tartamudeo fue mayor en la condición de canto que en la de lectura. La mayor reducción se observó cuando letras familiares fueron cantadas. Así, la duración incrementada de la fonación, la entonación, así como también la familiaridad pueden contribuir al efecto de fluidez incrementada del canto (Healey et al., 1976).

Andrews, Howier, Dozsa, y Guitar (1982) examinaron los efectos de 15 diferentes métodos de mejoramiento de la fluidez (incluyendo el canto) en un número de mediciones de tartamudeo. En la condición de canto, se les pidió a los participantes que cantaran canciones de su elección por 10 minutos. Los resultados mostraron que el canto reduce la frecuencia del tartamudeo aproximadamente un 90%, presumiblemente debido al incremento de la duración de la fonación.

Los resultados reportados por el estudio de neuroimagen llevado a cabo por Stager, Jeffries y Braun donde utilizaron tomografía por emisión de positrones (PET) para examinar las activaciones cerebrales asociadas con tareas que inducen fluidez

(canto y habla con ritmo, utilizando un metrónomo) en contraste con tareas que típicamente inducen el tartamudeo (construcción de enunciados). Las regiones cerebrales que estuvieron significativamente más activas durante las tareas que inducen fluidez comparadas con las que inducen el tartamudeo incluyeron a las áreas auditivas que procesan el habla y reciben la retroalimentación sensorial así como también las regiones motoras y premotoras que están involucradas en las acciones motoras articuladoras. Esto sugiere que un mecanismo auditivo-motor común podría ser la base de la fluidez en los tartamudos, el cual permite a los individuos participar en el automonitoreo del habla, resultando en un control más efectivo de los articuladores orales (Stager et al., 2003) (Soltysik and Hyde, 2006).

Otra condición en la que se ha reportado que el canto puede ayudar a disminuir los síntomas es el autismo. Se ha estimado que afecta a aproximadamente el 1% de la población. (Williams et al., 2006) El autismo se caracteriza por deficiencias en el lenguaje expresivo y la comunicación, con algunos individuos afectados completamente carentes de habla funcional (Sundberg, 1988).

Los individuos con autismo tienen habilidades superiores de procesamiento auditivo (Heaton et al., 1998) (Heaton, 2003) y comúnmente exhiben un fuerte interés en aprender y hacer música (Hairston, 1990) (Tonkinson, 1994).

Una intervención conocida como entrenamiento de mapeo audio-motor (AMMT), especialmente diseñada para ayudar a los niños con autismo a desarrollar el lenguaje expresivo, involucra tres componentes principales –canto, actividad motora e imitación- que recluta un presumible sistema disfuncional humano de neuronas espejo que se cree es la base de los déficits de comunicación en el autismo (Wan et al., 1989) (Wan et al., 2009) (Wan et al., 2010).

El canto recluta una cadena fronto-temporal bilateral más prominente que el habla, y esta cadena contiene algunos componentes del sistema de neuronas espejo (Brown et al., 2004) (Ozdemir et al., 2006). La actividad motora a través de tocar un instrumento de percusión no sólo captura el interés del niño, sino que también recluta una cadena sensoriomotora que controla los movimientos orofaciales y articuladores. (Meister et al., 2003) (Meister et al., 2009) Además, el sonido producido por un instrumento de percusión puede facilitar el mapeo audio-

motor que es crítico para la significativa comunicación vocal (Lahab et al., 2007). Finalmente, la imitación a través del entrenamiento repetitivo facilita el aprendizaje y altera las respuestas en el sistema de neuronas espejo (Catmur et al., 2007). Dado que el AMMT incrementa las interacciones entre los sistemas auditivo y motor, puede representar una estrategia terapéutica efectiva a través de la cual los individuos con autismo pueden desarrollar sus habilidades de comunicación (Wan et al., 2010) .

Se ha encontrado en algunas investigaciones que el entrenamiento musical formal sin interrupciones, al parecer, modifica estructural y funcionalmente el cerebro, disminuyendo así el deterioro cognitivo relacionado con la vejez.

Mathilde Groussard y colaboradores (2010), usando fMRI examinaron el impacto del entrenamiento musical en las funciones cerebrales durante la recuperación de la memoria de largo plazo, una facultad particularmente importante en la música, durante una tarea de familiaridad musical (por ejemplo, la memoria semántica para la música). La experiencia musical indujo activaciones suplementarias en el hipocampo, el giro medial frontal, y las áreas superiores temporales en ambos lados, sugiriendo una interacción constante entre la memoria episódica y semántica durante estas tareas en músicos.

Además, una investigación de morfometría basada en voxel (VBM) se llevó a cabo en estas áreas y reveló que la densidad de la materia gris del hipocampo era mayor en los músicos que en los no músicos. Indicando que la experiencia musical modifica de manera crítica los procesos de la memoria de largo plazo e induce plasticidad estructural y funcional en el hipocampo (Groussard et al., 2010). Lo cual es una evidencia sumamente importante, pues el entrenamiento musical es capaz de generar cambios estructurales y funcionales en el cerebro.

Groussard y su equipo de trabajo identificaron diferencias funcionales entre músicos y no músicos durante una tarea de memoria de largo plazo. Este patrón de activación aparece ampliamente para combinar las redes neurales involucradas en la memoria episódica y semántica. Estas diferencias cerebrales funcionales indican que cuando los músicos escuchan melodías familiares, más detalles perceptuales

y contextuales les vienen a la mente, vinculando memorias autobiográficas episódicas con recolección subjetiva (Groussard et al., 2010).

Además, estos investigadores encontraron diferencias estructurales en el hipocampo, un área clásicamente dedicada a los procesos de memoria. Estos descubrimientos apoyan la idea de que el entrenamiento musical puede estar asociado con el desarrollo de habilidades específicas de la memoria, que podrían contribuir a una mayor reserva cognitiva, la cual podría reducir el declive de la memoria relacionado con la edad (Groussard et al., 2010).

Marcus Herdener y colaboradores (2010) reportaron que el entrenamiento puede cambiar la organización funcional y estructural del cerebro, y que los modelos animales demuestran que la formación del hipocampo es particularmente susceptible a la neuroplasticidad relacionada con el entrenamiento.

En su estudio, Herdener usó los cerebros de músicos como un modelo para examinar las capacidades plásticas del hipocampo humano adulto. Por medio del uso de imágenes de resonancia magnética funcional optimizadas para la investigación del procesamiento auditivo, examinaron las respuestas cerebrales inducidas por la novedad temporal en patrones de sonido isócrono en músicos y en personas sin entrenamiento musical, ya que ha sido sugerido previamente que el hipocampo está crucialmente involucrado de varias maneras en la detección de la novedad (Herdener et al., 2010).

En el primer experimento transversal, identificaron respuestas neurales incrementadas a la novedad temporal en el hipocampo anterior izquierdo de músicos profesionales, apuntando a diferencias relacionadas con la experiencia en el procesamiento de los estímulos auditivos con novedad temporal en el hipocampo. En el segundo experimento, evaluaron las respuestas neurales a la novedad acústica temporal en un acercamiento longitudinal para identificar los cambios relacionados con el entrenamiento para factores predisponentes. Para este propósito examinaron una muestra independiente de estudiantes de academia de música antes y después de dos semestres de entrenamiento intensivo de habilidades auditivas. Después de este periodo de entrenamiento, las respuestas hipocámpicas a la novedad temporal en los sonidos se incrementó en los

estudiantes de música, y los análisis estadísticos de interacción de los cambios de la actividad cerebral en el tiempo sugieren al entrenamiento en lugar de efectos de predisposición. Por lo tanto, estos resultados proveen evidencia directa de cambios funcionales del hipocampo adulto en humanos relacionados con el entrenamiento musical (Herdener et al., 2010).

Karen Johanne Pallesen y colaboradores (2010) expresan que la competencia musical puede conferir ventajas cognitivas que van más allá del procesamiento de sonidos musicales familiares. La evidencia conductual indica un incremento general en ambas, la memoria de trabajo y la atención en los músicos. Es posible que los músicos, debido a su entrenamiento sean más aptos para mantenerse concentrados en un estímulo de tarea relevante, una habilidad que es crucial para la memoria de trabajo (Pallesen et al., 2010).

Estos investigadores midieron la señal de activación dependiente del nivel de oxigenación de la sangre (BOLD) en músicos y no músicos durante la memoria de trabajo de sonidos musicales para determinar la relación entre la ejecución, la competencia musical y la cognición generalmente aumentada.

Se ve reflejado en los tiempos de reacción y en las tasas de error que los músicos se desempeñaron mejor. Los músicos incluso tuvieron respuestas de BOLD más grandes en comparación con los no músicos en redes neuronales que mantienen la atención y el control cognitivo, incluyendo regiones de la corteza prefrontal lateral, la corteza parietal lateral, la ínsula y el putamen en el hemisferio derecho, y bilateralmente en la corteza prefrontal dorsal posterior y el giro cingulado anterior. La relación entre la tarea de desempeño y la magnitud de la respuesta BOLD fue más positiva en los músicos que en los no músicos, particularmente durante la tarea más difícil de memoria de trabajo. Los resultados incluso sugieren que la ejecución superior de una tarea de memoria de trabajo en músicos cuenta con una habilidad incrementada para ejercer un control cognitivo sostenido. Este beneficio cognitivo en los músicos es tal vez una consecuencia del entrenamiento musical enfocado (Pallesen et al., 2010).

Parbery-Clark y colaboradores (2009) dicen que los músicos tienen experiencia a lo largo de la vida analizando melodías dentro de armonías de fondo, y que esto

puede ser considerado un proceso análogo a la percepción del habla en ruido. Para investigar el efecto de la experiencia musical en la representación neural del habla-en-sonido, compararon respuestas subcorticales neuropsicológicas al habla en silencio y al habla en ruido en un grupo de músicos altamente entrenados y en no músicos como controles.

En los músicos se encontró que tuvieron una representación subcortical más robusta del estímulo acústico en presencia de sonido. Específicamente, los músicos mostraron una secuenciación neural más rápida, una representación incrementada de los armónicos del habla, y una respuesta morfológica menos degradada en el sonido. Las mediciones neurales se asociaron con un mejor desempeño conductual en el Test de Escuchar en el Sonido (HINT) para el cual los músicos superaron a los controles no músicos. Estos descubrimientos sugieren que la experiencia musical limita los efectos negativos del sonido de fondo, proveyendo así la primera evidencia biológica de la ventaja perceptual de los músicos para el habla-en-ruido (Parbery-Clark et al., 2009).

Los músicos tuvieron una codificación temporal y espectral más robusta del estímulo hablado provocador, compensando así los nocivos efectos del sonido de fondo. La secuenciación neural rápida y la codificación armónica incrementada en los músicos sugieren que la experiencia musical confiere una ventaja resultando en una sincronía neural más precisa en el sistema auditivo (Parbery-Clark et al., 2009).

Las investigaciones citadas en los puntos 1 al 1.4, en el presente documento, nos ayudan a tener mayor conocimiento y comprensión de cómo funciona nuestro cerebro cuando escuchamos y hacemos música e incluso de cómo la práctica musical puede modificar funcional y estructuralmente nuestro cerebro.

A pesar de que se ha estudiado ampliamente el fenómeno psicoacústico de la percepción de la disonancia y la consonancia, que se han identificado los correlatos neurales del canto, e incluso los cambios cerebrales dependientes de la práctica musical vocal e instrumental, poco sabemos de lo que sucede en nuestros cerebros cuando producimos intervalos musicales melódicos (consonantes y disonantes) con nuestra voz, ya que ninguna investigación se ha centrado en este fenómeno.



## 2. Métodos

El objetivo de esta investigación es estudiar los procesos neurobiológicos que permiten el canto y que ocurren al momento de ejecutarlo. En específico, correlacionar la actividad neuronal con el canto de intervalos musicales consonantes y disonantes (amplios y angostos), por estudiantes de jazz-canto de la Escuela Superior de Música.

### Hipótesis:

Será mayor el grado de ajuste requerido de la emisión vocal cuando se canten intervalos musicales disonantes, en contraste con los consonantes, provocando un incremento en la actividad de áreas relacionadas con la integración somatosensorio-motora durante el canto voluntario de los mismos.

## 2.1 Metodología e identificación demográfica.

El estudio se aplicó a 11 mujeres diestras entre 16 y 22 años de edad, sin ningún trastorno neurológico, psicológico y auditivo, y cuya lengua materna es el español.<sup>6</sup> Para verificar que todas fueran diestras, se les aplicó el test de lateralidad manual “Edinburgh Handedness Inventory”. Se llevaron a cabo audiometrías para descartar algún problema auditivo y que ninguna de las participantes tuviera alguna pérdida auditiva importante en algún rango de frecuencias. Se comprobó que todas las participantes escuchaban normalmente al momento de llevarse a cabo el estudio.

Para tener información detallada sobre el entrenamiento musical de las cantantes a las que se les practicó el estudio, se les pidió que contestaran el cuestionario sobre educación musical de BRAMS.

Las participantes comenzaron sus estudios musicales a una edad promedio de  $12.36 \pm 3.82$  años. Todas tienen entrenamiento musical formal ( $5.36 \pm 1.28$  años en canto y  $6.81 \pm 3.09$  en general), practican en promedio  $9.9 \pm 4.09$  horas a la semana de canto,  $6.18 \pm 5.7$  horas de otro instrumento musical (piano, violín, guitarra, saxofón, trombón, clarinete y percusiones), y en el momento de llevar a cabo el estudio, estaban cursando sus estudios en canto-scat, 5 a nivel medio superior y 6 a nivel superior, en la Academia de Jazz de la Escuela Superior de Música del Instituto Nacional de Bellas Artes (INBA).

Todas las participantes eran “sopranos”, es decir, con la misma tesitura vocal y para homogeneizar las pruebas de canto y que el grado del esfuerzo requerido al emitir los distintos intervalos musicales, fuera el mismo las pruebas conductuales y tareas de canto se programaron en un registro comprendido entre el sol 4 (G4) y do 6 (C6).

Cada participante contestó dos pruebas conductuales de discriminación de intervalos musicales y se les tomaron dos sesiones de Resonancia Magnética Funcional (fMRI) mientras ejecutaban una tarea de canto y una control (reposo-escucha).

---

<sup>6</sup> Se decidió que los sujetos de estudio fueran sólo mujeres con características similares, para que la muestra fuera lo más homogénea posible.

Para llevar a cabo las distintas pruebas conductuales y las tareas de canto dentro del resonador, se eligieron 4 intervalos musicales distintos: la octava justa=6 tonos, la séptima mayor=5 1/2 tono, la quinta justa=3 1/2 tonos y el tritono= 3 tonos. Como se explicó previamente en el punto 1.1 de este documento, el tritono es uno de los intervalos musicales considerados más polémicos y disonantes en la historia de la teoría musical. Otros intervalos que son conocidos por su alto grado de disonancia son la segunda menor y la séptima mayor. Los intervalos que se consideran consonancias perfectas son la quinta justa y la octava justa. Elegimos dos intervalos musicales considerados consonancias perfectas y dos de los más disonantes.

Se utilizaron dos intervalos amplios (uno consonante y otro disonante con una diferencia entre ellos de medio tono) y dos estrechos (uno consonante y otro disonante con una diferencia entre ellos de medio tono), ya que las cantantes generalmente, desde el principio de su entrenamiento, se habitúan a escuchar y a cantar intervalos consonantes con mayor frecuencia que los disonantes, por lo que las variables consonancia/disonancia y familiaridad/desconocimiento pueden llegar a confundirse. Se escogieron dos intervalos amplios (8J y 7M) y dos angostos (tritono y 5J) pues se considera que es más demandante técnicamente hablando, entonar correctamente intervalos amplios (independientemente de su grado de consonancia y disonancia).

Con el uso de estos cuatro intervalos, esperamos formar las combinaciones necesarias para distinguir entre las diferencias, si las hay, correlacionadas con el grado de consonancia/disonancia y las correlacionadas con la dificultad técnica por cantar un intervalo angosto o amplio.

Tanto las pruebas conductuales como la adquisición de imágenes cerebrales por medio de fMRI contaron con el consentimiento escrito de todas las participantes, en el caso de las menores de edad, con el de los padres o tutores legales. A las participantes y a los tutores legales se les explicó previamente, a la firma del consentimiento, el funcionamiento de la fMRI y la audiometría.

La adquisición de fMRI se llevó a cabo en las instalaciones del Hospital Infantil “Federico Gómez”, bajo la supervisión de personal especializado. El proyecto de investigación fue revisado y aprobado por las Comisiones de Investigación, de Ética en Investigación y de Bioseguridad, del Hospital Infantil de México “Federico Gómez”.

## 2.2 Pruebas Conductuales

Se llevaron a cabo varias pruebas conductuales, la primera consistió en que las 11 participantes cantaran intervalos musicales consonantes y disonantes (8J, 7M, 5J y tritono), en un registro comprendido entre el sol 4 y el do 6, después de escucharlos en un teclado. Esto se hizo para establecer un registro vocal unificado para las tareas de canto que se llevarían a cabo al hacer la resonancia magnética funcional. Posteriormente se grabó el canto, que ejecutaron las participantes, de cada uno de los cuatro intervalos antes mencionados con un micrófono Shure PG58 conectado a una laptop Hp Pavilion dv5 para hacer la grabación y el análisis de frecuencia con el software de edición de audio Cool Edit Pro 2.1. El análisis se hizo con una Fast Fourier Transform (FFT) de 65536 puntos y una ventana Hamming. Esta prueba ayudó a confirmar que ninguna de las participantes tuviera algún problema grave de afinación que les impidiera llevar a cabo las tareas de canto dentro del resonador.

La segunda y la tercera prueba fueron programadas en el software E-prime para que se escucharan (por medio de unos audífonos SONY Dynamic Stereo headphones MDR-CD555) aleatoriamente 100 intervalos musicales, 25 de cada uno de los 4 antes mencionados y que se guardaran los tiempos de reacción y respuestas de las participantes.

La segunda consistió en que las participantes indicaran si los intervalos escuchados eran consonantes o disonantes, esto se hizo presionando las teclas 1, del teclado de la computadora, para los consonantes y 2, para los disonantes. A esta prueba le llamaremos "CoDi".

En la tercera se indicó que identificaran cada intervalo presionando la tecla 8 si correspondían a una octava, la 7 a una séptima mayor, la 5 a una quinta justa y la 4 a un tritono. A esta prueba la identificaremos como "Intervalos".

### 2.3 Resonancia magnética funcional (fMRI).

Se utilizó la técnica de Resonancia Magnética Funcional (fMRI) para identificar con precisión las distintas zonas cerebrales que se activan cuando se entonan intervalos musicales consonantes y disonantes, y si existen diferencias entre estas.

Gracias al método de muestreo escaso<sup>7</sup> (Fig. 2.3.1) y a la utilización de micrófonos no-ferromagnéticos acoplados a distintas interfaces mediante cables debidamente aislados, (Cho et al., 1997) se pudo llevar a cabo la adquisición de la señal sonora sin problemas, en el presente estudio.

La adquisición de las imágenes de fMRI se realizó en un equipo General electric HDx de 1.5 Teslas con una antena para cabeza estándar, usando secuencias de gradiente EPI. Se estabilizó la cabeza de los participantes con relleno de hule espuma. Durante la adquisición de los datos los sujetos realizaron diferentes tareas que involucran el canto de distintos intervalos musicales consonantes y disonantes (octava justa, séptima mayor, quinta justa y tritono).

Se escucharon 50 intervalos musicales con la indicación de que no se cantaran los primeros 10 y que se cantaran los siguientes 40 dentro de un registro vocal cómodo para todas las participantes (Sol 4 a Do 6). Se le solicitó a las participantes que no cantaran los primeros 10 intervalos musicales escuchados, para que esas tomas nos sirvieran como base (reposo/escucha pasiva), es decir, para restar la actividad neural de la escucha de los distintos intervalos a la del canto de éstos y así establecer la actividad neural al momento de ejecutar la emisión vocal de los intervalos musicales consonantes y disonantes.

El resonador se programó para adquirir imágenes cerebrales durante 3 segundos cada 10 segundos y así poder dar 7 segundos de silencio para que se escuchara un intervalo y que, en su caso, fuera cantado por los sujetos de estudio y que el movimiento de la mandíbula no afectara las imágenes. Los intervalos que se utilizaron durante la fMRI también se programaron en el software E-prime para que fueran aleatorios. Cada intervalo musical tuvo una duración de 2 segundos. El tiempo entre cada tono fue de 10 segundos (Fig. 2.3.1). El sonido fue liberado por

---

<sup>7</sup> Basado en el método propuesto en Gaab et al., 2003.

audífonos binaurales, y se adquirieron 2 corridas por sujeto. Cada corrida duró 490 segundos=8.16 minutos.

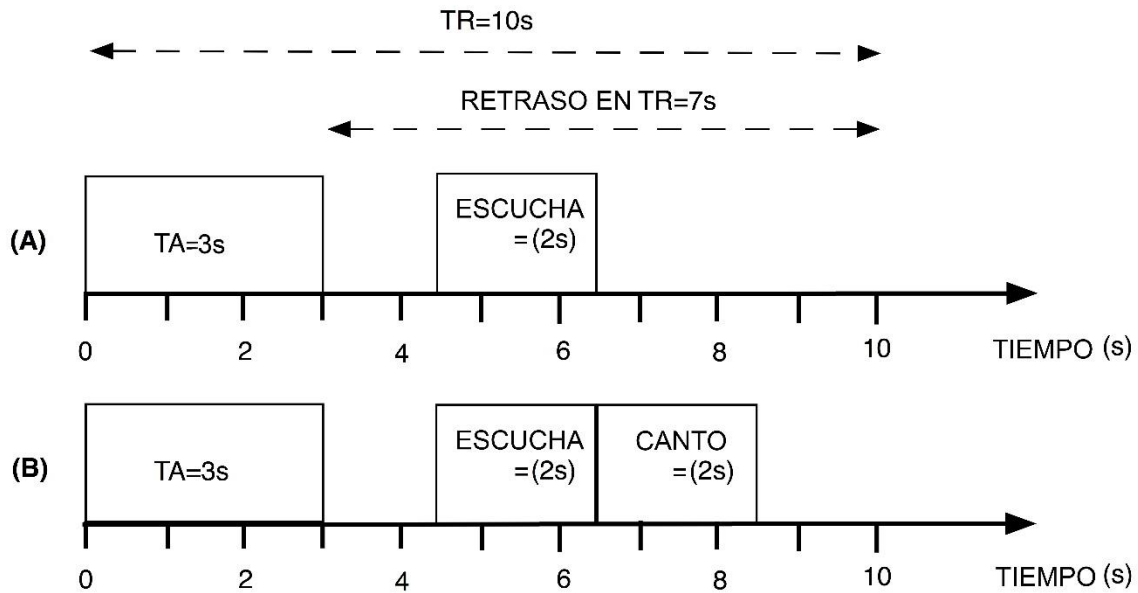


Fig. 2.3.1. Diseño de Muestreo Escaso. A) Cronología de eventos en una adquisición de fMRI, involucrando solamente la escucha pasiva de un intervalo musical melódico (2s) sin la reproducción vocal del mismo. B) Cronología de eventos en una adquisición de fMRI, involucrando la escucha de un intervalo musical melódico (2s) y su posterior reproducción vocal (2s).

Al mismo tiempo en que se obtenían las imágenes cerebrales, al llevar a cabo las tareas de canto, se grabó la señal sonora por medio de un micrófono Shure SM93 conectado, a través de un cable blindado trenzado triple de 10 metros<sup>8</sup>, a una interfaz de audio/midi Tascam US-144MKII USB 2.0 a su vez conectada a una laptop hp con sistema operativo Windows Xp y con el software de edición de audio Cool Edit Pro 2.1 (FFT= 65536 y ventana Hamming para el análisis de frecuencia).

Los datos obtenidos de la adquisición funcional se pre-procesaron usando SPM8 (Wellcome Department of Cognitive Neurology, London, UK; see <http://www.fil.ion.ucl.ac.uk/spm>). Las imágenes se realinearon con el primer

<sup>8</sup> Fue necesario blindar el cable para que actuara como una jaula de Faraday (anulando el campo electromagnético) y así evitar la aparición de ruidos o artefactos de imagen, que pudieran sesgar los resultados, debidos a la presencia del mismo.

volumen, posteriormente normalizadas espacialmente (voxel 2x2x2mm<sup>3</sup>) al espacio estereotáxico de Talairach y Tournoux usando el modelo del MNI y por último se suavizaron a 8mm isotrópicos FWHM Gaussian Kernel. Los datos se analizaron usando un modelo lineal general con una función de respuesta hemodinámica sintetizada con el software especializado SPM8.

Se hicieron distintos análisis con el método de sustracción de la fMRI, canto-reposo/escucha pasiva, canto de octava-escucha pasiva de octava, canto de séptima-escucha pasiva de séptima, canto de quinta-escucha pasiva de quinta y canto de tritono-escucha pasiva de tritono. Utilizando estos contrastes se comparó la actividad cerebral presente en el canto de intervalos angostos con amplios y la de los intervalos disonantes con los consonantes.



### 3. Resultados.

#### 3.1 Pruebas conductuales.

Se observó en las pruebas conductuales que el promedio de aciertos en la prueba “Intervalos” ( $77.42 \pm 12.36$ ) fue menor que en la prueba “CoDi” ( $81.57 \pm 10.45$ ) (Tabla 3.1-1) (Fig. 3.1-1). Hubo mayor efectividad, en la prueba “CoDi”, en donde sólo se les solicitó a las participantes que indicaran si el intervalo escuchado era consonante o disonante e incluso la desviación estándar fue menor en comparación con la obtenida en la prueba “Intervalos”, donde debían indicar cuál era el intervalo escuchado (octava justa, séptima mayor, quinta justa y tritono o quinta disminuida).

Tabla 3.1-1

Prueba Conductual	Identificación de Intervalos (Intervalos)	de Consonantes-disonantes (CoDi)
Aciertos (Media y Desviación).	$77.42 \pm 12.36$	$81.57 \pm 10.45$

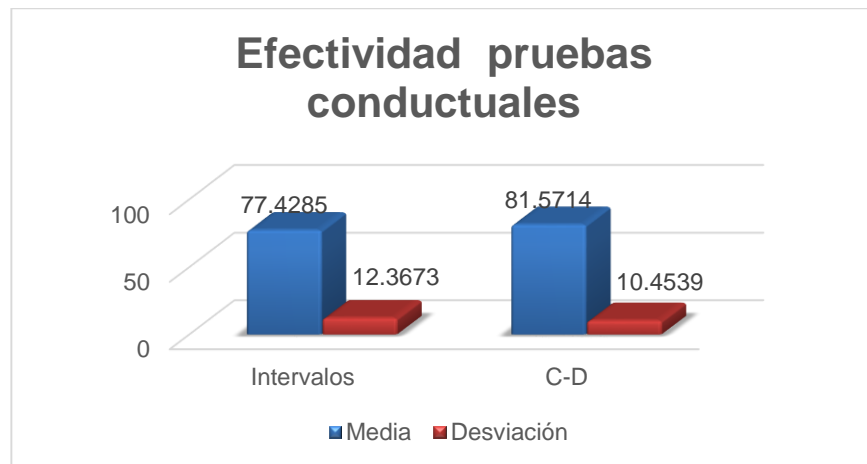


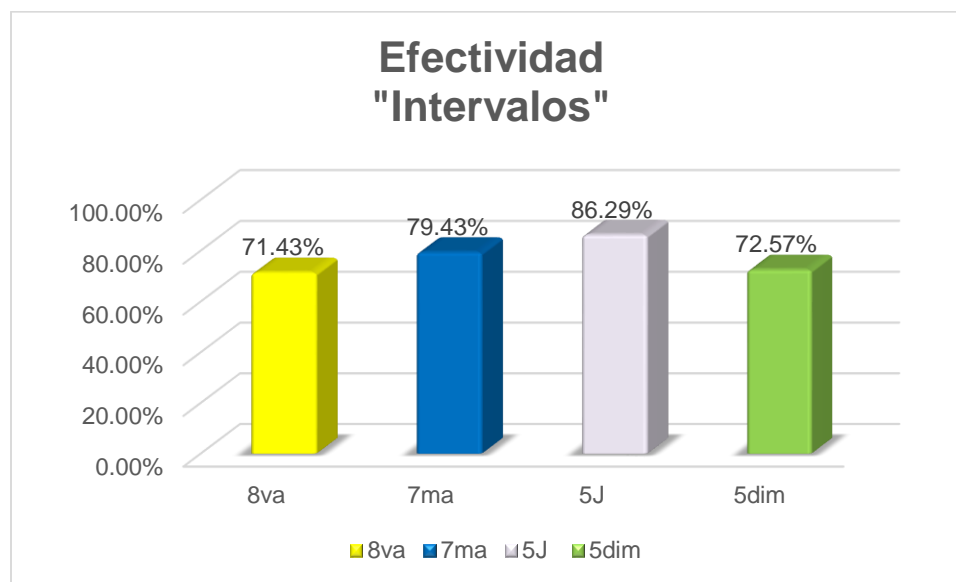
Fig. 3.1-1. Media y desviación estándar de aciertos en las pruebas conductuales “Intervalos” y “CoDi” (C-D).

En la prueba “Intervalos”, donde debían indicar el tipo de intervalo que escucharon, el promedio de aciertos para la octava fue  $17.8571 \pm 3.93 = 71.43\%$ , para la séptima  $19.8571 \pm 3.93 = 79.43\%$ , para la quinta justa  $21.5714 \pm 2.07 = 86.29\%$  y para el tritono  $18.1428 \pm 4.81 = 72.57\%$  (Tabla 3.1-2) (Fig. 3.1-2 y Fig. 3.1-3).

El intervalo con mayor porcentaje de aciertos fue la quinta justa, después le siguen en orden descendente la séptima mayor, la quinta disminuida o tritono y por último la octava justa.

Tabla 3.1-2

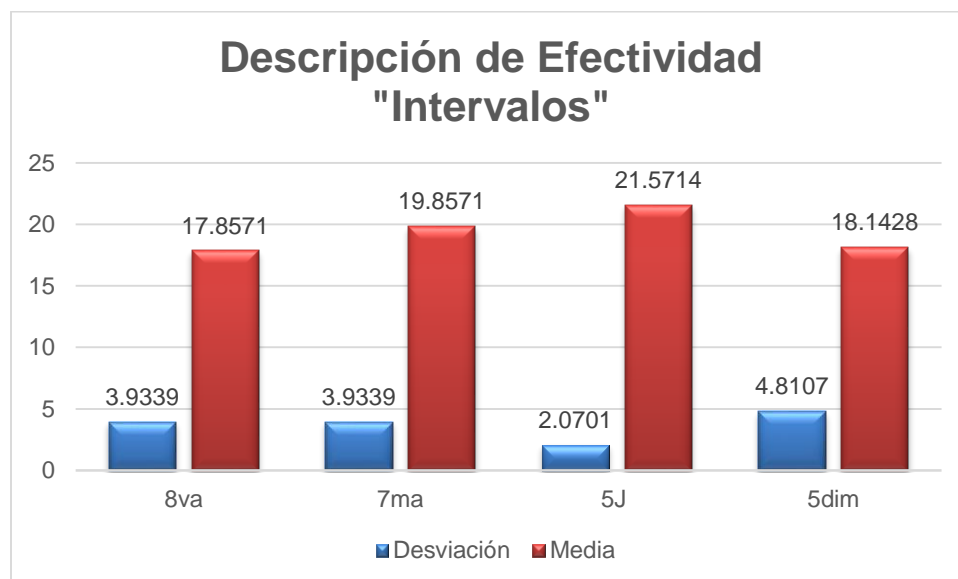
Intervalos	Promedio de Efectividad, Desviación y %
Tritono	$17.8571 \pm 3.93 = 71.43\%$
Quinta justa	$21.5714 \pm 2.07 = 86.29\%$
Séptima	$19.8571 \pm 3.93 = 79.43\%$
Octava	$18.1428 \pm 4.81 = 72.57\%$



En la figura 3.1-2. Porcentaje de aciertos por intervalo (octava justa=8va, séptima mayor=7ma, quinta justa=5J y tritono o quinta disminuida=5dim) en la prueba conductual “Intervalos”.

Todas las participantes escucharon 100 intervalos aleatoriamente, de los cuales eran 25 de cada uno (octavas justas, séptimas mayores, quintas justas y quintas disminuidas). (Figura 3.1-3) Por ello es que el eje de las Y's tiene un valor total de 25. De las 25 quintas justas que escucharon, en promedio acertaron en 21.57; de las 25 séptimas mayores escuchadas, acertaron en promedio en 19.85; de los 25 tritonos escuchados, acertaron en 18.14; y, de las octavas escuchadas, en 17.85.

La desviación estándar fue mayor en la identificación de los tritonos, después le siguieron las octavas y las séptimas donde la desviación fue idéntica, y por último las quintas justas. Perteneciendo la desviación más pequeña al intervalo de quinta justa, que también obtuvo la mayor cantidad de aciertos. Y la desviación más grande la tiene el intervalo más disonante, sin embargo no es el que tiene una menor cantidad de aciertos, aunque se encuentra entre los dos últimos (Figura 3.1-3).



La figura 3.1-3. Medias y desviaciones estándar de aciertos que hubo, por intervalo, prueba conductual "Intervalos".

Confundieron el 60% de las octavas con séptimas, el 52.5% con quintas y el 12.5% con tritonos. En la segunda prueba conductual las participantes confundieron en promedio los intervalos de octava, en orden descendente, con la séptima mayor, la quinta justa y por último con el tritono. Es decir, el intervalo con el que más

confundieron la octava justa fue el más cercano a pesar de este ser disonante, en segundo lugar está la quinta justa que es un intervalo consonante al igual que la octava y por último el intervalo de tritono o quinta disminuida (Tabla 3.1-3A) (Fig. 3.1-4A).

El 58.33% de las séptimas las confundieron con tritonos, el 19.44% con octavas, y de igual manera, el 19.44% con quintas. En la figura 3.1-4B podemos observar un fenómeno muy interesante en el que, a diferencia de la figura 3.1-4A, el intervalo escuchado (séptima mayor), no se confunde en primer lugar con el intervalo más cercano. Se confunde en primer lugar con el que también es disonante (tritono o quinta disminuida) aunque en cuanto a longitud, sea muy diferente al de séptima. (Tabla 3.1-3 B) El tritono tiene una distancia de tres tonos entre las notas que lo componen y el de séptima mayor, de cinco tonos y medio.

El 45.83% de las quintas justas las confundieron con tritonos, el 29.16% con octavas y el 25% con séptimas (Tabla 3.1-3C) (Fig. 3.1-4C). En primer lugar confundieron el intervalo de quinta justa con el tritono, en segundo con la octava y por último con la séptima mayor. Repitiéndose así el fenómeno descrito en la figura 3.1-4A, en donde en principio el intervalo escuchado se confunde con el intervalo más cercano, después con el intervalo de la misma naturaleza y por último con el que tiene menos similitudes.

El 75% de los tritonos los confundieron con quintas, el 18.75% con séptimas y el 6.25% con octavas (Tabla 3.1-3D) (Fig. 3.1-4D). En Primer lugar lo confundieron con la quinta justa, el intervalo más cercano; en segundo lugar, con el de séptima mayor que es también un intervalo disonante y por último con el de octava justa, con el que tiene mayores diferencias.

Tabla 3.1-3

A)

Intervalo	Séptima	Quinta	Tritono
Porcentaje de confusión de la octava.	60%	52.5%	12.5%

B)

Intervalo	Octava	Quinta	Tritono
Porcentaje de confusión de la séptima.	19.44%	19.44%	58.33%

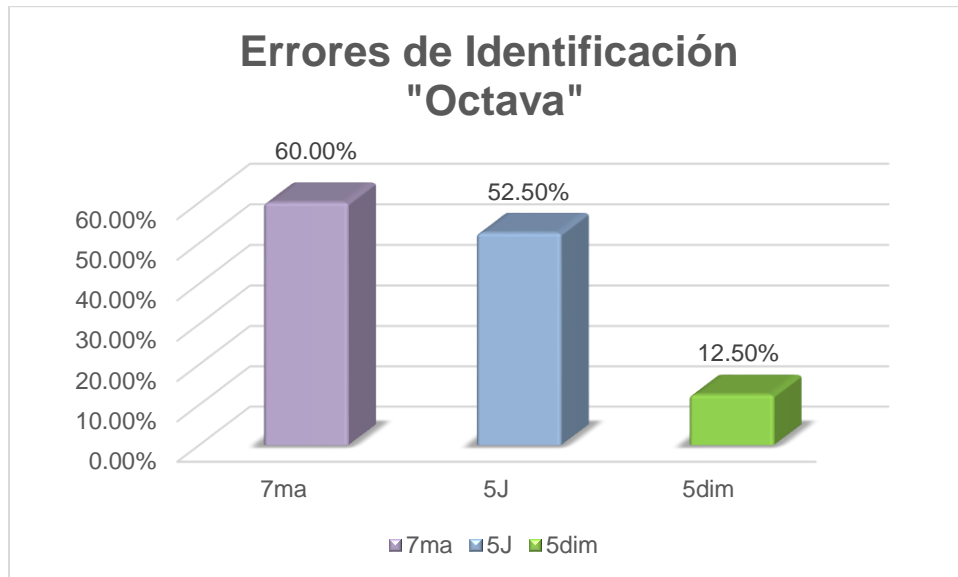
C)

Intervalo	Octava	Séptima	Tritono
Porcentaje de confusión de la quinta.	29.16%	25%	45.83%

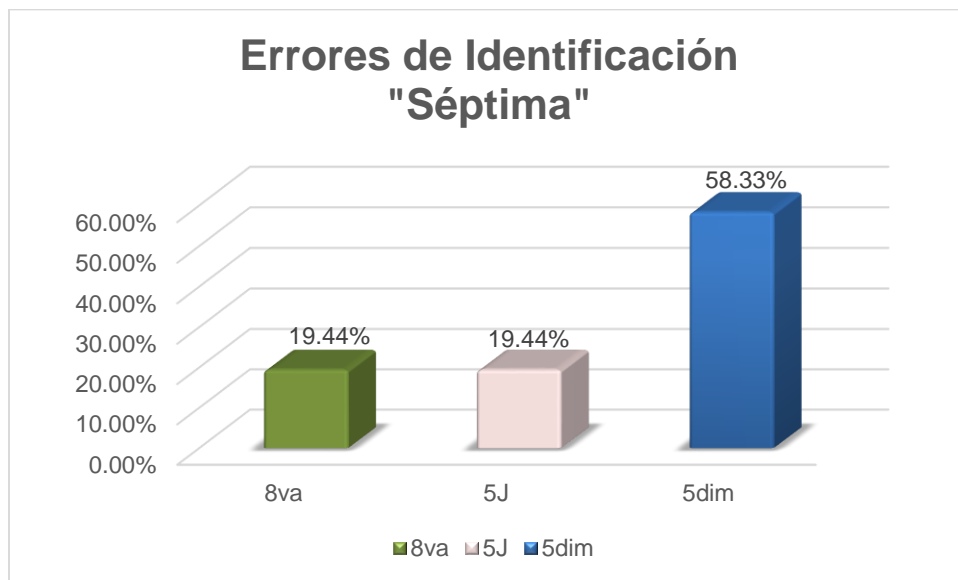
D)

Intervalo	Octava	Séptima	Quinta
Porcentaje de confusión del tritono.	6.25%	18.75%	75%

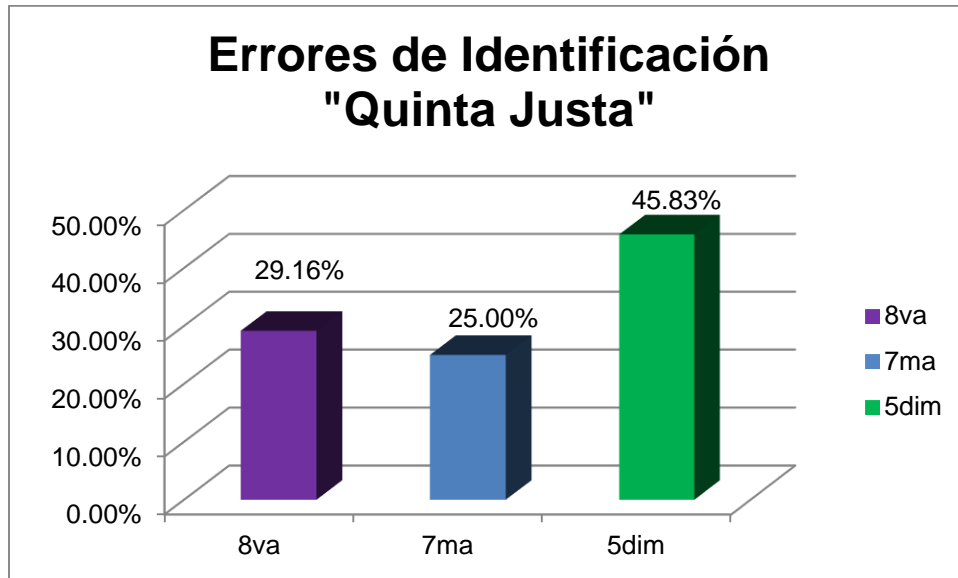
Figura 3.1-4



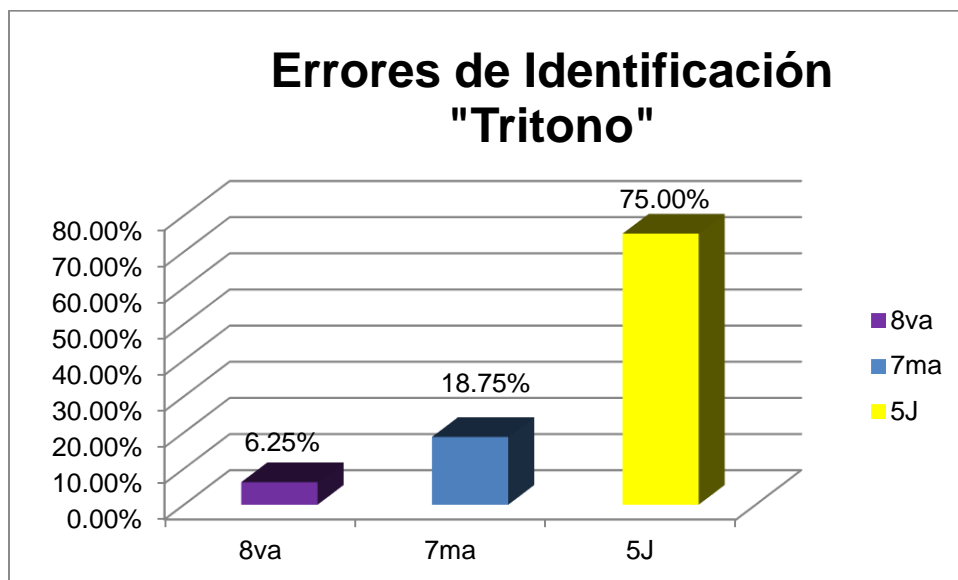
A) *Porcentaje de errores que tuvieron las participantes al identificar el intervalo de octava justa en la segunda prueba conductual.*



B) *Porcentajes de los intervalos (octava justa=8va, quinta justa=5J y tritono o quinta disminuida=5dim) con los que se confundió la séptima mayor.*



C) La figura expresa en valor porcentual con qué intervalos (octava justa=8va, séptima mayor=7ma y tritono o quinta disminuida=5dim) confundieron los sujetos de estudio al intervalo consonante de quinta justa.



D) La figura nos permite observar en porcentajes con qué intervalos (octava justa=8va, séptima mayor=7ma y quinta justa=5J) confundieron las participantes al intervalo disonante de tritono.

### 3.2 Tiempos de reacción.

En la prueba conductual “CoDi”, el tiempo que se tardaron en seleccionar los intervalos disonantes ( $3623.1971 \pm 1693.88$  ms) fue mayor que el que se tomaron para los consonantes ( $2936.47 \pm 558.54$  ms) (Tabla 3.2-1) (Fig. 3.2-1), y de igual manera en la prueba conductual “Intervalos”, el tiempo fue menor para identificar los intervalos consonantes ( $8va=4803.06286 \pm 1609.18$ ms,  $5J=4946.70857 \pm 1961.76$  ms) que los disonantes ( $7ma=5478.70857 \pm 2429.30$  ms,  $5dim=5800.994286 \pm 3087.70$  ms) (Tabla 3.2-2) (Fig. 3.2-2).

Tabla 3.2-1

Intervalos	Consonantes	Disonantes
Tiempo de Reacción (Promedio y Desviación).	$2936.47 \pm 558.54$ ms	$3623.1971 \pm 1693.88$ ms

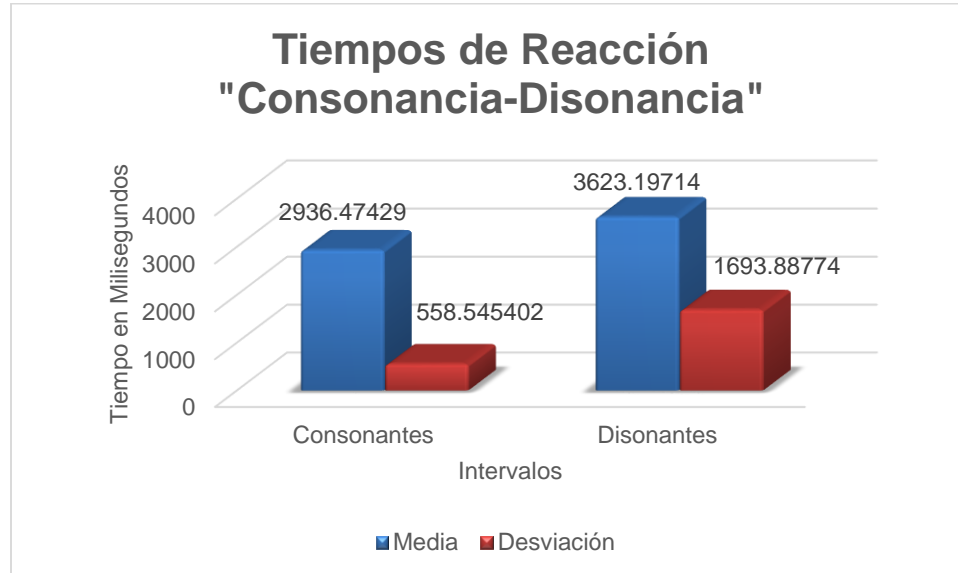


Fig. 3.2-1. Tiempos de reacción (milisegundos), en la prueba conductual “CoDi”, al identificar intervalos consonantes o disonantes.



Tabla 3.2-2

Intervalos	Tiempo de Reacción (Media y Desviación)
Tritono	5800.994286±3087.70 ms
Quinta justa	4946.70857±1961.76 ms
Séptima	5478.70857±2429.30 ms
Octava	4803.06286±1609.18 ms

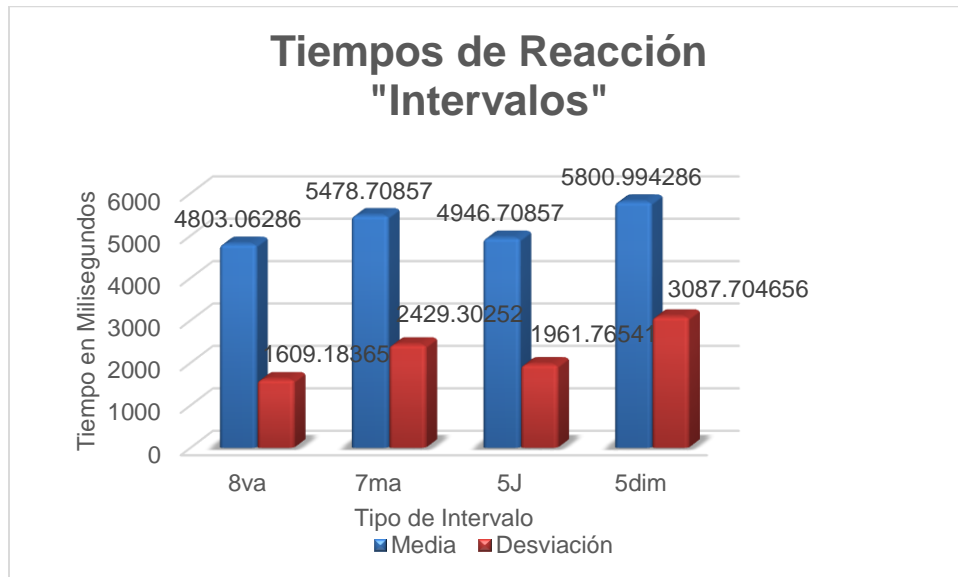


Fig. 3.2-2. Tiempos de reacción (milisegundos) por intervalo, en la prueba conductual "Intervalos".

### 3.3 Canto durante la fMRI

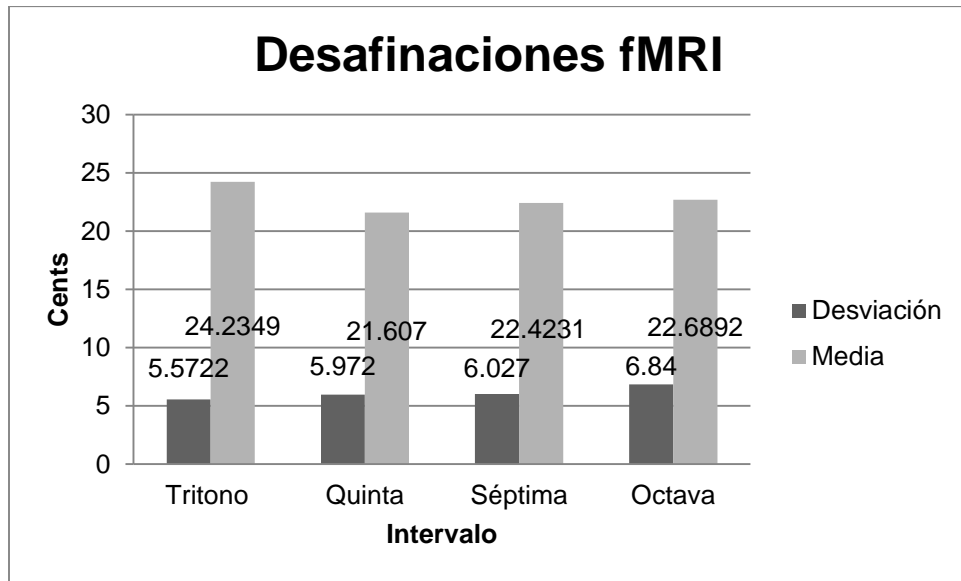
El análisis de la grabación del canto de los distintos intervalos musicales al momento en que se adquirían imágenes cerebrales por medio de Resonancia Magnética Funcional mostró que, aunque las participantes cantaban la mayoría de las veces el intervalo musical que escuchaban, hubo una tendencia a tener varios cents de diferencia en valor absoluto a la entonación perfecta de los mismos.

En promedio estuvieron desafinadas al cantar octavas  $22.6892818 \pm 6.84$  cents, al cantar séptimas mayores  $22.4231818 \pm 6.027$  cents, al cantar quintas justas  $21.6070545 \pm 5.972$  cents y al cantar tritonos  $24.2349545 \pm 5.5722$  cents (Tabla 3.3-1) (Fig. 3.3-1). Tomando en cuenta la convención del Dr. Felipe Orduña en la que al estar alejados de la afinación ideal entre 17 y 24 cents en valor absoluto, esto se considera una fuerte desviación de la afinación que puede provocar fuertes efectos melódicos y armónicos, entonces el canto de todas las participantes, estarían en este rango, al entonar octavas, séptimas mayores y quintas justas. Cuando la desviación es mayor a 24 cents esto se considera una desviación extrema de la afinación con efectos melódicos y armónicos, drásticos, comúnmente indeseables (Orduña-Bustamante y Ruíz, 1998).

El canto de los tritonos estuvo dentro esta última clasificación pues estuvo ligeramente más alejado de la afinación ideal, que el de los demás intervalos, sin embargo, no se observó una diferencia significativa, en la afinación, al entonar los distintos intervalos musicales, mientras se obtenían las imágenes cerebrales dentro del resonador (Fig. 3.3-1).

Tabla 3.3-1

Intervalo	Octava	Séptima	Quinta Justa	Tritono
Media	22.6892818 cents	22.4231818 cents	21.6070545 cents	24.2349545 cents
Desviación	6.084 cents	6.027 cents	5.972 cents	5.5722 cents



*Fig. 3.3-1. Media y desviación estándar de las desafinaciones, por intervalo, al llevarse a cabo la tarea de canto dentro del resonador.*

### 3.4. Resonancia Magnética Funcional (fMRI)

Las zonas cerebrales que exhibieron actividad incrementada relacionada con el canto, al llevar a cabo el análisis Canto vs. Reposo (Escucha Pasiva), que incluye a todos los intervalos musicales, son: las sensoriomotoras (corteza somatosensorial primaria (S1) y corteza somatosensorial secundaria (S2) bilaterales, motora primaria (M1) izquierda, área suplementaria motora (SMA), corteza premotora (PM) derecha, área de Broadman 44 (BA 44) bilateral); las auditivas (corteza auditiva primaria (PAC), giro temporal superior (STG), polo temporal); las multimodales (ínsula, lóbulos I-IV del Cerebelo); y, los ganglios basales (putamen y tálamo) (Tabla 3.4-1A).

Nuestros resultados son consistentes con los correlatos neurales del canto que han sido reportados en investigaciones previas (“el circuito del canto”). (Figura 3.4-1A)

Anteriormente fueron relacionadas con el canto, las cortezas auditivas bilaterales, las cortezas motoras primarias bilaterales (M1), el área suplementaria motora (SMA), la corteza del cíngulo anterior (ACC), la ínsula anterior, el cerebelo (Zarate and Zatorre, 2008) (Perry et al., 1999) (Kleber et al., 2007) el giro derecho de Heschl (Kleber, 2007) (Perry et al., 1999), el giro precentral, el plano temporal superior posterior (Perry et al., 1999), las cortezas sensoriomotoras y auditivas bilaterales secundarias, el área de Broca, el área de Wernicke y sus homólogos, el tálamo, el mesencéfalo y los ganglios basales (Kleber et al., 2007).

El análisis de las resonancias magnéticas mostró algunas diferencias de actividad cerebral cuando se cantan intervalos consonantes y disonantes. En el análisis del canto de intervalos disonantes menos el del canto de los consonantes, se encontró un incremento en la actividad de la corteza auditiva primaria, S1, la corteza orbitofrontal (OFC), la amígdala, el putamen izquierdo y la ínsula derecha (Tabla 3.4-1B) (Figura 3.4-1B).

También encontramos que al cantar intervalos amplios (octava y séptima mayor) en comparación con los más angostos (quinta justa y tritono), se incrementa la actividad en la ínsula anterior derecha (Tabla 3.4-1C) (Figura 3.4-1C).

Al cantar intervalos angostos en comparación con los amplios, se incrementó la actividad en la corteza premotora dorsal, el área suplementaria motora y el surco temporal superior (Tabla 3.4-1D) (Figura 3.4-1D).

Tabla 3.4-1

A) Regiones de actividad máxima neuronal asociada con el canto menos la escucha pasiva. Se presentan los valores Z-estadísticos de una imagen corregida en grupo ( $p=0.05$ ) por método del valor umbral ( $z > 3.00$ ) en x, y, z. La medida del grupo se reporta en voxels (CE). Las coordenadas se reportan en el espacio MNI.

Canto-Reposo (Escucha Pasiva)						
Región						
Sensorimotoras	Hemisferio	x	y	z	t	CE
S1	Izquierdo	-52	-15	36	7.06	615
	Derecho	58	-10	22	8.62	458
S2	Izquierdo	-62	-14	14	7.36	615
	Derecho	50	-12	20	6.30	1096
M1	Izquierdo	-46	-8	36	5.77	135
SMA		0	-2	58	7.82	216
PM	Derecho	54	4	34	11.94	458
BA 44	Izquierdo	-56	6	20	7.00	615
	Derecho	56	8	14	12.88	458
Auditivas						
PAC	Izquierdo	-46	-12	4	5.87	615
	Derecho	50	-8	4	6.33	1096
STG	Izquierdo	-62	-16	6	6.08	615
	Derecho	52	6	-12	5.69	1096
Polo Temporal	Izquierdo	-44	4	-18	5.49	615
	Derecho	48	14	-16	6.78	1096
Multimodales						
Ínsula	Derecho	42	4	-10	8.35	1096
Cerebelo						
Lóbulos I-IV	Izquierdo	-2	-54	0	6.22	148
Ganglios Basales						
Putamen	Izquierdo	-26	0	0	7.22	604
Tálamo	Izquierdo	-4	-28	0	6.22	148

B) Regiones de actividad máxima neuronal asociada con el canto de intervalos disonantes (comparado con el canto de intervalos consonantes). Se presentan los valores Z-estadísticos de una imagen segmentada por el método del valor umbral ( $z > 3.00$ ,  $p=0.05$  no corregida) en x, y, z. La medida del grupo se reporta en voxels (CE). Las coordenadas se reportan en el espacio MNI.

Canto de Intervalos Disonantes > Canto de Intervalos Consonantes						
Región		x	y	z	F	CE
Sensorimotoras						
S1	Derecho	58	-4	24	24.00	44
Auditivas						
PAC	Derecho	42	-30	10	25.50	262
Emocionales						
Amígdala	Izquierdo	-24	0	-22	17.29	54
OFC	Derecho	16	10	-12	15.68	25
Mesencéfalo						
Mesencéfalo	Izquierdo	-4	-12	-32	21.34	36
Multimodales						
Ínsula Posterior	Derecho	36	-18	-6	25.22	262
Subcorticales						
Putamen	Izquierdo	-24	-2	4	22.52	62

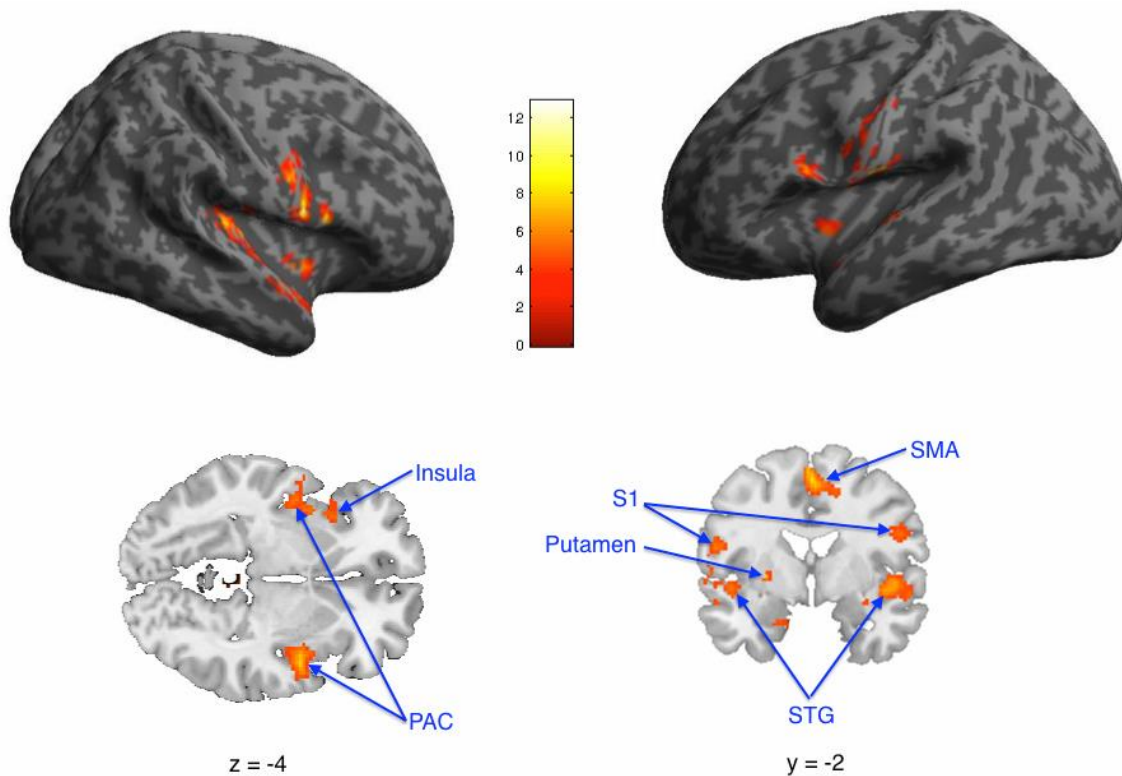
C) Regiones de actividad máxima neuronal asociada con el canto de intervalos angostos (comparado con el canto de intervalos amplios). Se presentan los valores Z-estadísticos de una imagen segmentada por el método del valor umbral ( $z > 3.00$ ,  $p=0.05$  no corregida) en x, y, z. La medida del grupo se reporta en voxels (CE). Las coordenadas se reportan en el espacio MNI.

Canto de Intervalos Angostos > Canto de Intervalos Amplios						
Región	Hemisferios	x	y	z	F	CE
Motoras						
PMd	Derecho	24	-10	54	21.61	149
SMA	Izquierdo	-6	-6	56	18.56	59

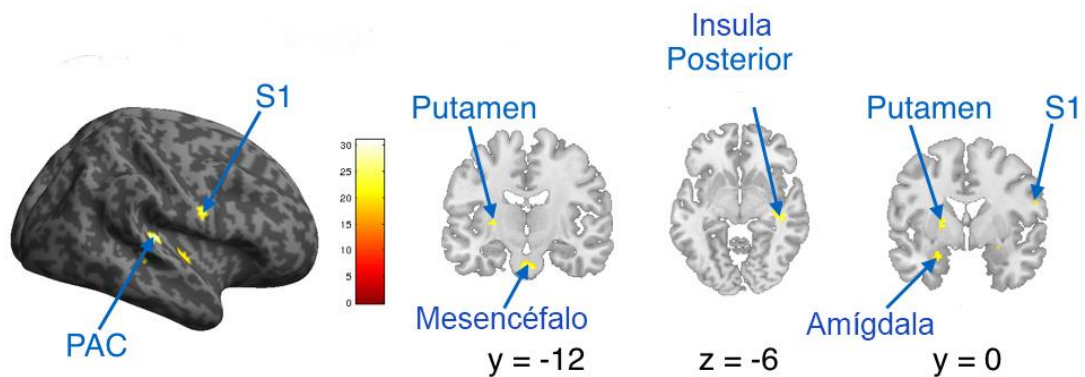
D) Regiones de actividad máxima neuronal asociada con el canto de intervalos angostos (comparado con el canto de intervalos amplios). Se presentan los valores Z-estadísticos de una imagen segmentada por el método del valor umbral ( $z > 3.00$ ,  $p=0.05$  no corregida) en x, y, z. La medida del grupo se reporta en voxels (CE). Las coordenadas se reportan en el espacio MNI.

Canto de Intervalos Amplios > Canto de Intervalos Angostos						
Región	Hemisferios	x	y	z	F	CE
Multimodales						
Ínsula Anterior	Derecho	36	-2	14	16.00	25

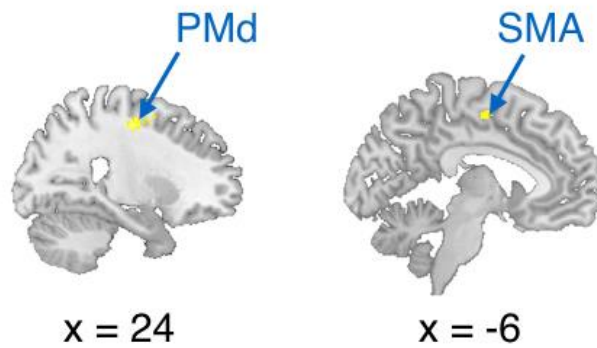
Fig. 3.4.1.



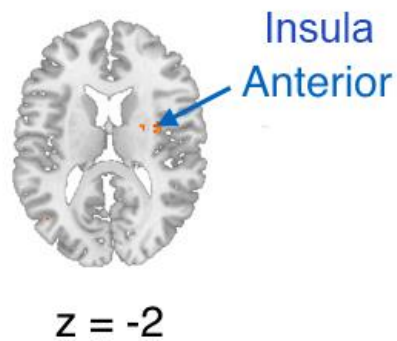
A) *Circuito del canto. Activación del cerebro para el canto, incluyendo todos los intervalos. Activación en regiones que constituyen el “circuito del canto”. Los mapas de activación fueron superpuestos usando la plantilla cortex\_20484.surf en el software SPM8, umbral significativo de grupo de  $p=0.05$  (FWE corregido), y Software MRICron.*



B) *Contraste entre el canto de intervalos disonantes vs. el canto de intervalos consonantes, mostrando activación BOLD en S1 derecha, PAC derecha, el putamen izquierdo, el mesencéfalo, la amígdala izquierda y la ínsula posterior.*



- C) *Contraste entre el canto de intervalos estrechos vs. el canto de intervalos amplios, mostrando activación incrementada en el Área Suplementaria Motora izquierda, la Corteza Auditiva Primaria y la Corteza Premotora Dorsal derecha.*



- D) *Contraste entre el canto de intervalos amplios vs. el canto de intervalos estrechos, mostrando actividad incrementada en la Insula Anterior derecha y el IPL izquierdo.*

*Exhibición visual en el umbral de  $p = 0,0001$  y el alcance del umbral de 20 voxeles.*



#### 4. Discusión.

En las pruebas conductuales no encontramos diferencias significativas en el desempeño de las participantes al identificar si los intervalos musicales escuchados eran disonantes o consonantes ni cuando tenían que indicar si el intervalo escuchado era una octava justa, séptima mayor, quinta justa o tritono. Los resultados, a pesar de no presentar diferencias significativas, nos indicaron que es posible que a las participantes les haya sido más complicado identificar eficazmente los intervalos disonantes, en comparación con los consonantes. Lo cual es apoyado por los resultados arrojados por la fMRI.

También es posible que para identificar los intervalos en las pruebas “CoDi” e “Intervalos”, las participantes no sólo utilicen las sensaciones subjetivas de placer o desagrado que se relacionan con la psicoacústica de la consonancia y la disonancia, sino que la longitud del intervalo también es crucial, ya que al confundir un intervalo con otro, en general lo hicieron con el que tiene características similares, es decir, o que tiene una longitud parecida o es de la misma naturaleza (consonante o disonante).

Un dato que se observó durante ambas pruebas conductuales es que la mayoría de las participantes entonaban el intervalo escuchado para asegurarse de que la respuesta que iban a seleccionar sería la correcta, desde su perspectiva.

Aparentemente la escucha, entonación posterior del intervalo musical y la retroalimentación auditiva y la somatosensorial, son cruciales, en personas que están en entrenamiento músico-vocal, para identificar y clasificar los intervalos musicales escuchados. Esto puede deberse al nivel de entrenamiento músico-vocal de las participantes y a que la entonación vocal con su respectiva retroalimentación auditiva y somatosensorial, son una herramienta de la cual pueden asirse para agilizar el acceso a su memoria y poder relacionar el intervalo entonado con una escala o melodía que les permita asegurarse que la respuesta motora que elegirán, es la correcta.

Es muy común que el entrenamiento musical de la identificación y entonación de intervalos se lleve a cabo a partir del aprendizaje de la escala mayor y de las escalas menores (natural, armónica y melódica) y la relación de los distintos grados de las mismas con los distintos intervalos musicales (Moncada, 1997). También es frecuente que los docentes de solfeo, canto y entrenamiento auditivo hagan que los alumnos relacionen algunas frases de melodías conocidas con algún intervalo musical para que sea entonado con mayor precisión y facilidad e identificado auditivamente de forma eficaz.

La interacción entre los sistemas sensorial y motor es indispensable para que podamos relacionarnos con nuestro entorno e interactuar en él, (Chen et al., 2009) en este caso es crucial para llevar a cabo el canto de intervalos consonantes, disonantes, amplios y angostos. Como se explicó previamente en este documento, es de suma importancia para todo músico poder utilizar la retroalimentación auditiva como una herramienta que permite saber si el tempo, ritmo, frecuencia e intensidad de la interpretación musical (instrumental y/o vocal) es la correcta. De no ser así, la interacción audio-motora permite, por medio de la retroalimentación auditiva y la somatosensorial, hacer los ajustes motores necesarios para corregir dicha interpretación.

Podemos entender la interacción audio-motora como un proceso que involucra sistemas cerebrales (auditivo y motor) que median sonidos y movimientos (Chen et al., 2009). Es decir, esta interacción es crucial para que el ser humano pueda reproducir, por medio de una respuesta motora, los sonidos percibidos en un estímulo auditivo determinado y saber que dicha reproducción o entonación es acertada.

Se ha demostrado que pacientes con lesiones en la corteza auditiva tienen problemas para procesar ritmos y reproducirlos con una acción motora a pesar de tener las cortezas motoras intactas. Esto sugiere que el deterioro en la integración audio-motora puede estar situado en el procesamiento auditivo y que sin un tiempo preciso o la transmisión, hacia el sistema motor, de información secuenciada, la tarea no puede llevarse a cabo (Chen et al, 2009).

La corteza auditiva primaria es primordial para llevar a cabo la decodificación de la información sensorial auditiva: procesamiento de los ritmos, tempo, frecuencias, intensidad, timbre, distancias entre los sonidos, etcétera.

Llevar a cabo ciertas tareas relacionadas con la práctica musical, implican, necesariamente la interacción audio-motora. Esto no sólo sucede con características de los estímulos musicales como el ritmo, sino también con el tempo y la frecuencia e intensidad de cada uno de los sonidos que componen una determinada melodía o acompañamiento musical (armonía).<sup>9</sup>

Es indispensable controlar la secuenciación, la sincronización y la organización espacial de los movimientos necesarios para hacer música (instrumental y vocal). Es sumamente importante recordar que los movimientos son llevados a cabo en un tiempo y espacio determinados, es decir, requieren de una planeación muy compleja para obtener una respuesta motora ante un estímulo auditivo que debe reproducirse con la voz (Zatorre et al., 2007).

Los músicos (instrumentistas y cantantes) perfeccionan, con muchos años de entrenamiento, la capacidad de ejecutar de manera casi perfecta moléculas rítmicas de alta complejidad, mantener una ejecución afinada de cada nota interpretada o ajustarla gracias a la retroalimentación auditiva y a la somatosensorial, no acelerar o disminuir la velocidad de la ejecución cuando no es apropiado hacerlo, etcétera (Zarate and Zatorre, 2008).

La corteza auditiva es clave para que se lleve a cabo la integración audio-motora, pero no es la única estructura cerebral implicada. Otros correlatos neurales del canto, que posiblemente están relacionados con la necesidad de reprogramar los mecanismos utilizados en la entonación de los intervalos musicales consonantes, para cantar correctamente los disonantes, son el putamen izquierdo y la corteza premotora dorsal (PMd).

La corteza premotora dorsal ha sido considerada una interface básica de la integración audio-vocal junto con la ACC, mientras que las cortezas auditivas, y el putamen se han relacionado con el automonitoreo auditivo y el ajuste de la emisión vocal (Zarate and Zatorre, 2008).

---

<sup>9</sup> Tanto las melodías como los acordes están conformados por distintos intervalos musicales.

La corteza premotora ha sido relacionada con la integración entre la información sensorial y la acción motora durante el habla, el canto y la reproducción-sincronización de ritmos (con escucha previa), así como en el desarrollo y elección de planes motores basados en señales sensoriales (Zarate and Zatorre, 2008) (Chen et al., 2009). También se ha considerado que la corteza premotora incrementa su actividad cuando se ha desarrollado la habilidad de coordinar de forma precisa, respuestas motoras en un determinado tiempo, con señales auditivas; o, cuando se escucha un determinado ritmo o melodía y se imagina-planea su reproducción. Esto sugiere un acoplamiento entre los sistemas auditivo y motor durante la percepción y la reproducción (Chen et al., 2009) (Kleber et al., 2007).

La corteza premotora es la única área cortical motora que está directamente conectada con el giro temporal superior posterior y con la corteza motora primaria (M1). Hay evidencia de presencia de tractos de fibras de sustancia blanca entre el giro temporal superior posterior y la corteza premotora dorsal rostral (Chen et al., 2009). También ha sido involucrada, la PMd, en aspectos de alto orden de la organización de los movimientos. Las neuronas dentro de la PMd recuperan e integran la información sensorial con instrucciones motoras para llevar a cabo un plan de acción. Estas neuronas están involucradas en la planeación motora, así como también en la preparación y selección de parámetros de movimiento (dirección y amplitud) en respuesta a una señal sensorial. Hay evidencia de que la actividad en la PMd se incrementa cuando se requiere de mayor control motor y hay una mayor dependencia de la retroalimentación auditiva para seleccionar movimientos precisamente cronometrados. Se sugiere que la PMd también tiene un papel importante en el aprendizaje de trayectorias espaciales (Zatorre et al., 2007) (Chen et al., 2009).

Recordemos que el putamen izquierdo mostró mayor actividad cuando, las participantes del estudio, cantaron intervalos musicales disonantes en comparación con la entonación de los consonantes.

El putamen tiene un incremento de actividad (dependiente de la experiencia) junto con las áreas auditivas bilaterales cuando, cantantes profesionales, necesitan llevar a cabo la corrección de la emisión vocal por medio de un programa motor bien aprendido. Esto sucede cuando ellos ignoran un estímulo auditivo distinto al tono cantado (Zarate and Zatorre, 2008).

Los ganglios basales y el cerebelo incrementan su actividad con secuencias motoras complejas (Zatorre et al., 2007). Los ganglios basales y SMA han sido involucrados también con el control de alto nivel de la ejecución de secuencias motoras, mientras que el cerebelo se relaciona más con la corrección fina de movimientos individuales (Zatorre et al., 2007).

El cerebelo, los ganglios basales y el área suplementaria motora han sido relacionadas con la sincronización del movimiento. En donde los ganglios basales son más importantes en escalas de tiempo superiores a un segundo y el cerebelo en escalas de tiempo en milisegundos (Zatorre et al., 2007). Esto puede significar que los ganglios basales están implicados en la sincronización voluntaria de movimientos, ya que, como se mencionó anteriormente, las respuestas motoras tardías suelen darse conscientemente (Zarate, 2013).

Al, las participantes, entonar intervalos musicales disonantes, también se incrementó la actividad en la corteza auditiva primaria, la corteza sensoriomotora primaria, la corteza orbitofrontal, la amígdala y la ínsula posterior derecha.

Se considera que la ínsula posterior tiene un rol en la percepción musical, la estabilidad al cantar y en las emociones evocadas por la música (Schirmer et al., 2012) (Terao et al., 2006) (Koelsch, 2014). Al observarse también un incremento en la actividad de la corteza sensoriomotora primaria, podemos deducir, que tal vez es necesario un mayor grado de integración de la información proveniente de la retroalimentación somatosensorial responsable de los movimientos requeridos para articular los sonidos, al cantar intervalos disonantes, que al cantar intervalos consonantes.

La amígdala se ha relacionado con la respuesta emocional ante estímulos auditivos no placenteros (Blood et al., 1999), junto con la corteza orbitofrontal (Stuss, 2002), aunque también se ha reportado incremento de actividad en la

corteza orbitofrontal derecha con la percepción de la consonancia musical (Blood et al., 1999) (Limb, 2006).

El que la corteza orbitofrontal se encuentre conectada con la región superior temporal rostral por medio del fascículo uncinado, puede indicar que es crucial para la evaluación y regulación de la respuesta a un estímulo auditivo con carga afectiva. Esta corteza tiene una localización anatómica que permite su interacción con áreas prefrontales laterales que subyacen el procesamiento cognitivo de alto orden, así como con las cortezas entorrinal y peririnal por medio de la vía límbica dorsal, lo cual puede indicar que es un componente importante del circuito cerebral relacionado con el procesamiento de nueva información (Stuss, 2002) .

La evidencia muestra que los intervalos disonantes son entonados mediante el ajuste de los mecanismos neurales usados para la producción de los intervalos consonantes.

En cuanto a los resultados obtenidos al comparar la actividad neural al cantar intervalos amplios y angostos, los resultados muestran un incremento de actividad en la ínsula anterior derecha cuando se cantan los primeros.

A la ínsula anterior se le ha relacionado con el ajuste motor necesario para entonar conscientemente una nota durante un tiempo determinado, sin que cambie la afinación de la frecuencia fundamental de la misma o bien, para llevar a cabo la entonación y ajuste de un cambio en la frecuencia de dicha nota (Zarate and Zatorre, 2008). Así mismo, se considera importante a la ínsula anterior para que se lleve a cabo el control voluntario de la respiración durante la vocalización, y es crucial en la integración audio-motora y para que se posibilite la retroalimentación somatosensorial (Zarate, 2013).

Cantar intervalos disonantes o amplios, requiere de un mayor involucramiento de los mecanismos neurales asociados con la integración de la retroalimentación interna y externa de los sistemas auditivo y sensoriomotor que cuando se cantan intervalos consonantes e intervalos angostos. Los circuitos neuronales involucrados en el canto de intervalos disonantes y amplios, sin embargo, son bastante diferentes e incluso cantar este tipo de intervalos es difícil por diferentes razones.El

canto de los intervalos amplios parece que requiere la integración de los sistemas auditivo y sensoriomotor, mientras que el canto de los intervalos estrechos no.

Tomando como base los resultados obtenidos en la presente investigación, se puede concluir que a las participantes del estudio se les dificultó más la identificación y el canto de los intervalos disonantes, así como el canto de intervalos amplios. Al parecer, se requiere de mayor control motor para llevar a cabo la emisión vocal de las frecuencias componentes de los estímulos auditivos de intervalos musicales disonantes y amplios.

El conocimiento obtenido en este estudio se puede utilizar como base para llevar a cabo investigaciones posteriores en el ámbito muy particular de la voz cantada, y es posible también que esta línea de investigación ayude en un futuro al desarrollo de procesos terapéuticos para solucionar problemas de fonación, lenguaje, atención y memoria; así como para comprender mejor los procesos cerebrales implicados en la entonación-reproducción de tonos, intervalos y melodías (con o sin letra), así como en la composición e improvisación músico-vocal.<sup>10</sup>

---

<sup>10</sup> Para análisis y resultados más detallados, consultar los anexos.

## 5. Alcances y limitaciones.

Aunque la fMRI, es una técnica con grandes ventajas y alcances, también tiene limitaciones y retos por vencer. Un problema al que nos enfrentamos al llevar a cabo la investigación fue obtener una correlación entre actividad cerebral y los intervalos musicales entonados por los sujetos de estudio (estudiantes de jazz-canto de la Escuela Superior de Música), a pesar de la gran intensidad del ruido producido por el equipo de Resonancia Magnética Funcional.

La adquisición de señal sonora en el interior de un sistema de resonancia magnética presenta dos dificultades principales. La primera tiene que ver con la adquisición misma, que se dificulta debido a la gran intensidad del campo magnético generado por el sistema, por otra parte está que el nivel de presión sonora entre las bobinas puede rebasar los 100 dB (Ravicz et al., 2000) . Varios autores, sin embargo, han logrado llevar a cabo medidas acústicas en el interior del equipo durante una Resonancia Magnética Funcional.

El ruido que produce el resonador puede alterar significativamente los resultados, dado que es un estímulo auditivo no deseado que provoca una activación de determinadas zonas cerebrales que no están propiamente relacionadas con los correlatos neurales específicos de una tarea de interés determinada. Para eliminar dicho estímulo indeseable Kleber y colaboradores, en el 2007, hicieron su versión del método de muestreo escaso utilizado previamente por Nadine Gaab (Gaab et al., 2003).

En el presente estudio se utilizó una adaptación de este método, para poder eliminar los inconvenientes que produce el ruido del resonador. Por ello fue fundamental diseñar una técnica estandarizada de acoplamiento temporal entre la adquisición de señal por Resonancia Magnética Funcional y la producción de la voz cantada. Esto nos permitió correlacionar la actividad neuronal con el canto de distintos intervalos musicales, y extraer información conductual.

También fue de suma importancia utilizar un cable blindado para eliminar los artefactos, en las imágenes cerebrales, que la presencia del mismo dentro del resonador, pudiera provocar.



## 6. Bibliografía.

- 1) Abo-el-Enein M. A. 1966. Laryngeal myotatic reflexes. *Nature*. 209: 682-686.
- 2) Andrews, G., Howie, P. M., Dozsa, M., Guitar, B.E. 1982. Stuttering: Speech pattern characteristics under fluencyinducing conditions. *Journal of Speech and Hearing Research* 25:208–216.
- 3) Arfelis, C. T. 2005. *Cantar y Hablar*. Barcelona, España.: Paidotribo.
- 4) Armony, J. L.-M. 2012. Resonancia Magnética Funcional (RMf): Principios y Aplicaciones en Neuropsicología y Neurociencias Cognitivas. *Neuropsicología Latinoamericana*. 4(2):36-50.
- 5) Beltrán, A. C. “Arte del canto”. 1949. México. Secretaría de Educación Pública.
- 6) Bidelman, G. and Krishnan, A. 2009. Neural correlates of consonance, dissonance and hierarchy of musical pitch in the human brainstem. *Neuroscience* 29 (42):13165-13171.
- 7) Blood, A. J., Zatorre, R. J., Bermudez, P., Evans, A. C. 1999. Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nat. Neurosci.* 2:382-387.
- 8) Brown, S., Martinez, M. J., Hodges, D. A., Fox, P. T., Parsons, L. M. 2004. The song system of the human brain. *Brain Research Cognitive Brain Research*. 20:363–375.
- 9) Catmur, C., Walsh, V., Heyes, C. 2007. Sensorimotor learning configures the human mirror system. *Current Biology* 17:1527–1531.
- 10) Chan, M.F., Chan, E. A., Mok, E., Kwan, Tse, F. Y. 2009. Effect of music on depression levels and physiological responses in community-based older adults. *International Journal of Mental Health Nursing* 18:285–294.
- 11) Chen, J. L., Penhune, V. B., Zatorre, R. J. 2009. The role of auditory and premotor cortex in sensorimotor transformations. *The Neuroscience and Musica III-Disorders and Plasticity: Ann N. Y. Acad. Sci.* 1169:15-39.

- 12) Cho, Z.H., Park, S. H., Kim, J.H., Chung, S.C., Chung, S.T., Chung, J.Y., Moon, C.W., Yi, J.H., Sin, C.H. and Wong, E.K. 1997. Analysis of Acoustic Noise in MRI. *Magnetic Resonance Imaging* 15(7): 815-822.
- 13) Colcord, R. D. and Adams, M. R. 1979. Voicing duration and vocal SPL changes associated with stuttering reduction during singing. *Journal of Speech and Hearing Research* 22:468–479.
- 14) Dalla Bella, S., Peretz, I., Rousseau, L., Gosselin, N. 2001. The role of tempo and mode in conveying music emotions: evidence from development. *Cognition*. In press.
- 15) Danhauser, A. 2001. *Teoría de la Música*. México. Editapsol.
- 16) Davidow, J. H., Bothe, A. K., Andreatta, R. D., Ye, J. 2009. Measurement of phonated intervals during for fluencyinducing conditions. *Journal Speech and Language Hearing Research* 52:188–205.
- 17) Dejonckere, P. and Lebacq, J. 1981. Mechanism of initiation of oscillatory motion in human glottis. *Arch Int Physiol Biochim*. 89:127--136.
- 18) Ferrer, J. S. 2008. *Teoría, Anatomía y Práctica del Canto*. Barcelona, España: Herder.
- 19) Gaab, N., Gaser, C., Zaehle, T., Jaencke, L., Schlaug, G. 2003. Functional anatomy of pitch memory — an fMRI study with sparse temporal sampling. *NeuroImage* 19: 1417–1426.
- 20) Groussard, M., La Joie, R., Rauchs, G., Landeau, B., Chételat, G., Viader, F., Desgranges, B., Eustache, F., Platel, H. 2010. When Music and Long-Term Memory Interact: Effects of Musical Expertise on Functional and Structural Plasticity in the Hippocampus. *PLoS ONE* 5(10): e13225. doi:10.1371/journal.pone.0013225.
- 21) Hairston M. 1990. Analyses of responses of mentally retarded autistic and mentally retarded nonautistic children to art therapy and music therapy. *Journal of Music Therapy* 27:137–150.
- 22) Halwani, Gus, F., Psyche, L., Rüber, T., and Schlaug, G. 2011. Effects of practice and experience on the arcuate fasciculus: comparing singers,

- instrumentalists, and non-musicians. *Frontiers in Psychology: Auditory Cognitive Neuroscience*. Vol. 2. Article 156.
- 23) Hartmann, W. M. 1998. *Signals, sound and sensation*. New York, NY. Springer.
- 24) Healey, E. C., Mallard, A. R. III, Adams, M. R. 1976. Factors contributing to the reduction of stuttering during singing. *Journal of Speech and Hearing Research* 19:475–480.
- 25) Heaton, P., Hermelin, B., Pring, L. 1998. Autism and pitch processing: A precursor for savant musical ability? *Music Perception* 15:291–305.
- 26) Heaton, P. 2003. Pitch memory, labelling and disembedding in autism. *Journal of Child Psychology and Psychiatry and Allied Disciplines* 44:543–551.
- 27) Helmholtz, H. 1877. *On the sensations of tone as a physiological basis for the theory of music*. New York, NY. Dover Publications.
- 28) Herdener, M., Esposito, F., Di Salle, F., Boller, C., Hilti, C.C., Habermeyer, B., Scheffler, K., Wetzel, S., Seifritz, E., Cattapan-Ludewig, K. 2010. Musical Training Induces Functional Plasticity in Human Hippocampus. *The Journal of Neuroscience* 30(4): 1377-1384.
- 29) Hernández, Heber Z. *Teoría y práctica de la armonía*. Libro 1. 1989. México. Ediciones las américas.
- 30) Itoh, K., Suwazono, S., Nakada, T. 2010. Central Auditory Processing of Noncontextual Consonance in Music: An Evoked Potential in Music. *J. Acoust. Soc. Am.* 128 (6) 3781-3787.
- 31) Kleber, B., Birbaumer, N., Veit, R., Trevorrow, T., Lotze, M. 2007. Overt and imagined singing of an italian aria. *Neuroimage* 36: 889-900.
- 32) Kleber, B., Veit, R., Birbaumer, N., Gruzelier, J., Lotze, M. 2009. The brain of opera singers: experience-dependent changes in functional activation. *Cerebral Cortex* 20: 1144-1152.
- 33) Koelsch, S. 2014. Brain correlates of music-evoked emotions. *Nat. Rev. Neuroci.* 15:170-180.

- 34) Lahab, A., Saltzman, E., Schlaug, G. 2007. Action representation of sound: Audiomotor recognition network while listening to newly acquired actions. *Journal of Neuroscience* 27:208–214.
- 35) Limb, Ch. J. 2006. Structural and functional neural correlates of music perception. *The Anatomical Record Part 288A*:435-446.
- 36) Mansion, Madeleine. *El estudio del canto, técnica de la voz hablada y cantada*. 1947. Buenos Aires, Argentina. Ricordi Americana.
- 37) Meister, I.G., Borojerdi B., Foltys, H., Sparing, R., Hube, W., Topper, R. 2003. Motor cortex hand area and speech: Implications for the development of language. *Neuropsychology* 41:401–406.
- 38) Meister, I. G., Buelte, D., Staedtgen, M., Boroierdi, B., Sparing, R. 2009. The dorsal premotor cortex orchestrates concurrent speech and fingertapping movements. *The European Journal of Neuroscience* 29:2074–2084.
- 39) Moncada, Francisco. 2011. *Teoría de la música*. México. Musical Iberoamericana. Ediciones Framong.
- 40) Natke, U., Donath, T. M, Kalveram, K. T. 2003. Control of voice fundamental frequency in speaking vs singing. *Journal of the Acoustical Society of America* 113:1587–1593.
- 41) Orduña-Bustamante, F. y Ruíz Boullosa, R. 1998. “Escalas musicales y sus temperamentos”. *Revista Mexicana de Física*, 44 (2). Pp. 205-210. ISSN:0035-001x
- 42) Ozdemir, E., Norton, A., Schlaug, G. (2006) Shared and distinct neural correlates of singing and speaking. *NeuroImage* 33:628–635.
- 43) Pallesen, K.J., Brattico, E., Bailey, C.J., Korvenoja, A., Koivisto, J., Gjedde, A., Carlson, S. 2010. Cognitive Control in Auditory Working Memory Is Enhanced in Musicians. *PLoS ONE* 5(6): e11120. doi:10.1371/journal.pone.0011120.
- 44) Parbery-Clark, A., Skoe, E., Kraus, N. 2009. Musical Experience Limits the Degradative Effects of Background Noise on the Neural Processing of Sound. *The Journal of Neuroscience*. 29: 14100-14107.

- 45) Parncutt, R. and McPherson, G. E. 2002. *The Science and psychology of music performance: creative strategies for teaching and learning*. Oxford: Oxford University Press.
- 46) Perani, D., Saccuman, M. C., Scifo, P., Spada, D., Andreolli, G., Rovelli, R., Baldoli, C., Koelsch, F. 2010. Functional Specializations for Music Processing in the Human New Born Brain. *PNAS* 107 (10): 4758-4763.
- 47) Peretz, I., Blood, A. J., Penhune, V., Zatorre, R. (2001) Cortical Deafness to Dissonance. *Brain* 124: 928-940.
- 48) Perry, D.W., Zatorre, R.J., Petrides, M., Alivisatos, B., Meyer, E., Evans, A.C. 1999. Localization of cerebral activity during simple singing. *NeuroReport* 10: 3979–3984.
- 49) Plomp, R. and Levelt, W. J. W. 1965. *Consonance and critical bandwidth*. Netherlands. Institute for Perception RVO-TNO.
- 50) Prasse, J. E. and Kikano, G. E. 2008. Stuttering: An overview. *American Family Physician* 77:1271–1276.
- 51) Ravicz, M.E., Melcher, J.R., Kiang, N.Y.-S. 2000. Acoustic Noise During Functional Magnetic Resonance Imaging. *J. Acoust. Soc. Am.* 108: 1683-1696.
- 52) Ravicz, M.E. and Melcher, J.R. 2001. Isolating the Auditory System From Acoustic Noise During Functional Magnetic Resonance Imaging: Examination of Noise Conduction Through the Ear Canal, Head, and Body. *J. Acoust. Soc. Am.* 109: 216-231.
- 53) Roederer J.G. 1975. *Introduction to physics and psychophysics of music*. New York, NY. Springer.
- 54) Schellenberg, E. G. and Trehub, S. E. 1994. Frequency ratios and the discrimination of pure tone sequences. *Percept. Psychophys.* 56:472-478.
- 55) Schellenberg E. G. and Trehub, S. E. 1996. Children's discrimination of melodic intervals. *Dev. Psychol.* 32:1039-1050.
- 56) Schirmer, A., Mickle Fox, P., Grandjean, D. 2012. On the spatial organization of sound processing in the human temporal lobe: a meta-analysis. *Neuroimage.* 63:137-147.

- 57) Shapira, L., I. and Stone, L. 2008. Perception of Musical Consonance and Dissonance: an Outcome of Neural Synchronization. *Interface* 5:1429-1434.
- 58) Smith, A. and Zelaznik, H. N. 2004. Development of functional synergies for speech motor coordination in childhood and adolescence. *Dev Psychobiol.* 45:22--33.
- 59) Soltysik, D.A. and Hyde, J.S. 2006. Strategies for block-design fMRI experiments during task-related motion of structures of the oral cavity. *NeuroImage* 29, 1260–1271.
- 60) Stager, S. V., Jeffries, K. J., Braun, A. R. 2003. Common features of fluency-evoking conditions studied in stuttering subjects and controls: An H(2)15O PET study. *Journal of Fluency Disorders.* 28:319–335.
- 61) Stuss, D. T. 2002. Principles of frontal lobe function. New York. Oxford University Press.
- 62) Sundberg J. 1988. The science of the singing voice. Dekalb (IL): Northern Illinois University Press.
- 63) Tager-Flusberg, H. 1997. Language acquisition and theory of mind: Contributions from the study of autism..In: Adamson, LB.; Ronski, MA., editors. *Research on communication and language disorders:Contributions to theories of language development.* Pauk Brookes Publishing; Baltimore. 133-158.
- 64) Terao, Y., Mizuno, M., Sakurai, Y., Ugawa, Y., Kobayasi, S. 2006. Vocal amusia in a profesional tango Singer due to a right superior temporal cortex infarction. *Neuropsychology.* 44:479-488.
- 65) Titze, I. R. 1994. Principles of voice production. Englewood Cliffs, NJ: Prentice-Hall.
- 66) Titze, I. R. 1995. Motor and sensory components of a feedback-control model of fundamental frequency. In: Bell-Berti and Kaphael L, eds, *Producing Speech: Contemporary Issues* New York: American Institute of Physics, 1995:309-318.
- 67) Thurman, L. and Welch, G. F. 2000. Bodymind and voice: Foundations of voice education (2<sup>nd</sup> ed.). Iowa City: National Center of Voice and Speech.

- 68) Tonkinson S. 1994. The Lombard effect in choral singing. *Journal of Voice* 8:24–29.
- 69) Wan, C., Demaine, K., Zipse, L., Norton, A., Schlaug, G. 1989. From music making to speaking: Engaging the mirror neuron system in autism. 2010. Manuscript submitted for publication Weiner, W., Singer, C. Parkinson's disease and nonpharmacological treatment programs. *Journal of American Geriatric Society* 37:359–363.
- 70) Wan, C., Zipse, L., Norton, A., Demaine, K., Baars, R., Zuk, J., et al. 2009. Using an auditory-motor mapping therapy to improve expressive language abilities in nonverbal children with autism; Poster session presented at the 8th Annual Auditory Perception, Cognition, and Action Meeting; Boston, MA.
- 71) Wan, Catherine Y., Rüber, Theodor, Hohmann, Anja and Schlaug, Gottfried. 2010. The Therapeutic Effects of Singing in Neurological Disorders. Beth Israel Deaconess Medical Center and Harvard Medical School. *Music Percept.* 27(4): 287–295.
- 72) Wiens, M. E., Reimer, M. A., Guyn, H. L. 1999. Music therapy as treatment method for improving respiratory muscle strength in patients with advanced multiple sclerosis. *Rehabilitation Nursing* 24:74–80.
- 73) Williams, J. G., Higgins, J. P., Brayne, C. E. 2006. Systematic review of prevalence studies of autism spectrum disorders. *Archives of Disease in Childhood* 91:8–15.
- 74) Zarate, J.M. and Zatorre, R.J. 2008. Experience-dependent neural substrates involved in vocal pitch regulation during singing. *NeuroImage* 40:1871–1887.
- 75) Zarate, J. M. 2013. The neural control of singing. *Frontiers in human neuroscience.* 7:1-12.
- 76) Zatorre, R.J., Chen, J.L., Penhune, V. B. 2007. When the brain plays music: auditory-motor interactions in music perception and production. *Nature Reviews.* 8:547-558.

## 7. Anexos.

### 7.1. Neural activity related to discrimination and vocal production of consonant and dissonant musical intervals.

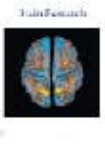
Brain Research 1643 (2016) 59–69



Contents lists available at ScienceDirect

Brain Research

journal homepage: [www.elsevier.com/locate/brainres](http://www.elsevier.com/locate/brainres)



#### Research Paper

## Neural activity related to discrimination and vocal production of consonant and dissonant musical intervals



Nadia González-García<sup>a</sup>, Martha A. González<sup>b</sup>, Pablo L. Rendón<sup>b,\*</sup>

<sup>a</sup>Hospital Infantil de México Federico Gómez, Dr. Márquez 163, México, D.F. 06702, México

<sup>b</sup>Centro de Ciencias Aplicadas y Desarrollo Tecnológico, Universidad Nacional Autónoma de México, Ciudad Universitaria, A.P. 70-486, México, D.F. 04510, México

#### ARTICLE INFO

**Article history:**  
Received 20 July 2015  
Received in revised form  
10 March 2016  
Accepted 27 April 2016  
Available online 28 April 2016

**Keywords:**  
Consonance and dissonance  
Singing  
Audio-visual integration  
fMRI  
Somatosensory  
Neural correlates of music

#### ABSTRACT

**Background:** Relationships between musical pitches are described as either consonant, when associated with a pleasant and harmonious sensation, or dissonant, when associated with an inharmonious feeling. The accurate singing of musical intervals requires communication between auditory feedback processing and vocal motor control (i.e. audio-vocal integration) to ensure that each note is produced correctly. The objective of this study is to investigate the neural mechanisms through which trained musicians produce consonant and dissonant intervals.

**Methodology:** We utilized 4 musical intervals (specifically, an octave, a major seventh, a fifth, and a tritone) as the main stimuli for auditory discrimination testing, and we used the same interval tasks to assess vocal accuracy in a group of musicians (11 subjects, all female vocal students at conservatory level). The intervals were chosen so as to test for differences in recognition and production of consonant and dissonant intervals, as well as narrow and wide intervals. The subjects were studied using fMRI during performance of the interval tasks; the control condition consisted of passive listening.

**Results:** Singing dissonant intervals as opposed to singing consonant intervals led to an increase in activation in several regions, most notably the primary auditory cortex, the primary somatosensory cortex, the amygdala, the left putamen, and the right insula. Singing wide intervals as opposed to singing narrow intervals resulted in the activation of the right anterior insula. Moreover, we also observed a correlation between singing in tune and brain activity in the premotor cortex, and a positive correlation between training and activation of primary somatosensory cortex, primary motor cortex, and premotor cortex during singing. When singing dissonant intervals, a higher degree of training correlated with the right thalamus and the left putamen.

**Conclusions/significance:** Our results indicate that singing dissonant intervals requires greater involvement of neural mechanisms associated with integrating external feedback from auditory and sensorimotor systems than singing consonant intervals, and it would then seem likely that dissonant intervals are learned by adjusting the neural mechanisms used for the production of consonant intervals. Singing wide intervals requires a greater degree of control than singing narrow intervals, as it involves neural mechanisms which again involve the integration of internal and external feedback.

© 2016 Elsevier B.V. All rights reserved.

#### 1. Introduction

It is generally acknowledged that certain combinations of musical notes are consonant, while others are considered dissonant (Hirns, 1999). When played in isolation, consonant intervals tend to be described as pleasant, while dissonant intervals are associated with an unresolved or unpleasant sensation. Although the neural response associated with listening to consonant and

dissonant intervals is the subject of a reasonably large number of studies, there has been very little work done on the neural activity which accompanies the production of consonant and dissonant intervals. In terms of sound production, it has been noted by Zarate (2013) that singing presents what is possibly the most convenient mechanism of musical production, as the instrument is already contained within the body and it is therefore not necessary to construct complicated interfaces or instruments for their use during neuroimaging. Thus, in this study we are concerned with the neural mechanisms recruited during the singing of both consonant and dissonant intervals. We expect the singing of

\* Corresponding author.

E-mail address: [pablo.rendon@ccadet.unam.mx](mailto:pablo.rendon@ccadet.unam.mx) (P.L. Rendón).



dissonant intervals to pose a greater difficulty for the singer than the singing of consonant intervals, and in principle it would then be possible to mistake the effect of increasing technical difficulty for the effect of increasing dissonance. In order to eliminate this possibility we have asked participants to sing narrow and wide intervals as well as consonant and dissonant ones. It has long been acknowledged by both researchers (Jernstrom and Sundberg, 1982, 1988) and singing teachers (Duey, 1951) that singing wide intervals is considerably more difficult than singing narrow intervals. However, the difficulty in this case has more to do with technical aspects of vocal production than in the case of dissonant intervals, where the difficulty is possibly due to an innate bias of our species favoring consonance over dissonance (Zentner and Kagan, 1996; Trainor and Heinmiller, 1998).

It must be stressed that the aesthetic response to an interval or chord may be greatly affected by context, so that the appropriate musical argument can cause dissonant intervals to sound pleasant, for example. In order to avoid ambiguity and confusion, the idea of consonance as an effect devoid of context has been termed sensory consonance by Terhardt (1974). By using isolated intervals we may then assume that whatever activations are observed are related to sensory consonance only, and there is no interference from context-dependent processing (Minati et al., 2009). One well-known criterion used to define sensory dissonance is due to Helmholtz (1954), who suggested that dissonant intervals are characterized by beating between adjacent frequency components, and by a subsequent perceived roughness. More recently, however, attempts have been made to characterize consonance rather than dissonance, of which the most successful associate sensory consonance with harmonicity, which refers to the formation of an aggregate spectrum formed by the superposition of the consonant component of chords (McDermott et al., 2010; Cousineau et al., 2012).

That the mechanisms that account for distinctions between consonance and dissonance in simultaneous intervals are also engaged for sequential intervals is not immediately obvious. There is, however, evidence that listeners process intervals represented by simple frequency ratios more readily than those associated with more complex ratios, and this ease of processing is, in turn, associated with consonance (Schellenberg and Trehub, 1994, 1996). Further, Zentner and Kagan (1998) have shown that infants find melodies constructed from consonant sequential intervals more perceptually pleasing than those which contain a greater number of dissonant sequential intervals. Forde Thompson (2013) argues that this is possibly due to the combination of neural responses to sequential tones occurring within working memory, and also due to greater exposure to consonant simultaneous intervals, which translates into a preference for those same intervals in a melodic context. In fact, McLachlan et al. (2013) have demonstrated through experiments that unfamiliarity with certain intervals correlates with the perception of these intervals as dissonant. It has indeed become clear, from a number of studies (Schellenberg and Trehub, 1996; Peretz et al., 2001; Tramo et al., 1990), that physical phenomena cannot solely account for the differences between perceived consonance and dissonance.

Studies of the neural mechanisms underlying the differences between perceived consonance and dissonance indicate that these differences may derive from simple biological bases. Working with sequential pure-tone intervals, Schellenberg and Trehub (1995) observed that intervals with simpler frequency ratios, such as the octave or fifth, are processed more easily by infant listeners than intervals with more complex frequency ratios. Using evoked potentials to study listeners' reactions to, again, pairs of pure-tone intervals, Iroh et al. (2010) have found that the neural processing of noncontextual consonance is dependent on associative processing of pitch relationships in the auditory cortex. These results are in

agreement with work by Peretz et al. (2001) and Tramo et al. (1990) where the ability of patients with auditory cortex lesions to evaluate consonance is found to be severely impaired. Further, PET and fMRI studies carried out by Blood et al. (1999), Pallesen et al. (2005), Foss et al. (2007), and Minati et al. (2009) have shown that listening to dissonant intervals or chords engages a variety of cortical areas that are not activated by listening to consonant intervals or chords. There is, however, substantial variability among these studies regarding the neural correlates of consonance and dissonance perception. Increasing dissonance, for example, is variously correlated with activity in right parahippocampal gyrus and precuneus (Blood et al., 1999), in amygdala, retrosplenial cortex, brain stem and cerebellum (Pallesen et al., 2005), in inferior frontal gyrus, middle frontal gyrus, superior temporal gyrus, anterior cingulate, thalamus and cerebellum (Foss et al., 2007) and in scattered clusters in the left hemisphere (Minati et al., 2009).

Different neural networks are involved during vocalization on the one hand and during identification of sensory consonance and dissonance on the other, and we expect to observe activation of component parts of both these networks during these particular tasks. The neural network involved in vocalization and singing of both a single note (Perry et al., 1999) and a complete melody (Brown et al., 2004a; Kleber et al., 2007; Jurgens and Floog, 1970; Jurgens and Richter, 1986) includes, for vocal motor control, the face area in the primary motor cortex, the supplementary motor area (SMA), the anterior cingulate cortex (ACC), the cerebellum, the basal ganglia and the thalamus. It also includes the primary (S1) and secondary (S2) somatosensory cortices for somatosensory feedback processing, the superior temporal gyrus and sulcus for auditory feedback processing, and, according to Zarate (2013), the insula during multimodal processing of sensorial feedback.

In the present study we evaluate the ability of semiprofessional singers to recognize and sing adequately both consonant and dissonant intervals, and to sing narrow and wide intervals in tune. Singing provides a straightforward mechanism by means of which we may test differences in activation of brain areas associated with consonance and dissonance while having the participants perform a motor task. This is a novel aspect of our study, as neural correlates of consonance and dissonance production have not been widely examined. One such study (Callan et al., 2006) is concerned with perception and covert production of both song and speech. During the covert production task, the authors observed greater differential activity for singing over speech in the subcallosal cingulate, which is a region associated with consonance perception (Brown et al., 2004b).

Generally speaking, we expect the participants to find it more difficult to produce dissonant or wide intervals than consonant or narrow ones (Jernstrom and Sundberg, 1982, 1988), albeit for different reasons. Whereas being able to sing wide intervals proficiently is really a matter of technique achieved through intensive training, we believe singing dissonant intervals well and in tune is more a question of skillful self monitoring and corresponding adjustment. It is common practice for singers to produce sound, and then to listen while comparing the actual sound to the intended sound. Very often this comparison leads to small adjustments in vocal output, which become more automatic and efficient with practice, and reflect changes in the neural processes involved, leading to the conclusion that training and auditory-motor integration are causally related (Lahav et al., 2007). With regard to the more technical aspects of vocal production, there is evidence that the response of the somatosensory system related to the larynx is enhanced as a function of the training of vocal skills (Kleber et al., 2010). In a wider context, it has been shown by Bangert and Altenmüller (2003) that after a certain amount of training, there is a significant overlap in activity in auditory and

premotor areas when musicians listen to a piece they know how to play and when they actually play it, suggesting that the auditory and motor systems interact very closely during both perception and production of sound.

Even if it is by and large accepted that singing dissonant intervals is more demanding than singing consonant ones, the reason behind this greater degree of difficulty is not well established. We hypothesize that singers require a greater degree of adjustment of vocal output when singing dissonant as opposed to consonant intervals, leading to increased activation of areas related to somatosensory-motor integration during voluntary vocal production. We expect training to have a noticeable effect as well, especially regarding the neural mechanisms which allow for efficient adjustments in vocal production resulting from auditory self-monitoring. Singing wide as opposed to narrow intervals is also associated with a greater degree of difficulty, but we assume that this represents a more purely technical challenge. Thus, we anticipate that the brain regions recruited when singing intervals to which a greater degree of difficulty is associated will differ from those activated by singing simpler intervals, but that the regions of activation will not necessarily be the same in the case of dissonant or wide intervals.

2. Results

2.1. Behavioral pre-test

For each task, 100 intervals were generated randomly within the range described previously, but in such a way that each of the four test intervals occurred exactly 25 times, and were then presented to participants. During the first task, the participants were asked to identify whether an interval was consonant or not. Participants identified the intervals correctly as consonant or dissonant 74.82 ± 14.15% of the time. During the second stage, participants were asked to identify which of the four intervals mentioned above was being played to them, a task which they managed correctly 65.72 ± 21.67% of the time. Therefore, there is a significant difference between the success rate when identifying whether an interval is consonant or dissonant and the success rate when identifying particular intervals [ $p < 0.05$ ]. The success rates for identifying each of the four intervals are shown in Table 1. There was no significant difference in the success rate at identifying any of the particular intervals [ $F(3,40)=0.67, p=0.57$ ]. It was immediately apparent, however, that participants were less successful at identifying particular intervals than they were at merely identifying consonance or dissonance, as would be expected. Indeed, we did not observe from the results any trend exhibiting better identification of individual intervals.

2.2. fMRI results

2.2.1. Singing network

Table 2 shows the significant activation of certain brain regions during the singing minus listening task. Brain activated regions were labeled according to the probabilistic assignment obtained from the SPM Anatomy Toolbox (Eickhoff et al., 2005), and, when necessary, the Harvard-Oxford cortical and subcortical structural

atlas. The following abbreviations are used: S1 (primary somatosensory cortex), S2 (secondary somatosensory cortex), M1 (primary motor cortex), SMA (supplementary motor area), PM (premotor cortex), BA (Brodmann area), PAC (primary auditory cortex), and STG (superior temporal gyrus). Overall, subjects exhibited noticeable activation in S1 and S2, M1, PM and SMA, BA 44, PAC, STG, temporal pole, anterior insula, putamen, thalamus and cerebellum. During singing, then, we found activation in regions that constitute the "singing network" as defined by Özdemir et al. (2006), Zarate and Zatorre (2008), and Kleber et al. (2007, 2010, 2013) in previous studies, thus validating the basic design; see Fig. 1.

2.2.2. Consonance versus dissonance

The contrasts described in Section 4.3.5 were used to examine the differences in brain activation between singing consonant and dissonant intervals, as well as between singing narrow and wide intervals. For these analyses, these four contrasts, singing tritone minus listening to tritone, singing fifth minus listening to fifth, singing seventh minus listening to seventh, singing octave minus listening to octave, were compared using 2-way repeated measures ANOVAs with the factors consonance/dissonance and narrow/wide interval.

With regard to the production of consonant and dissonant intervals, we observed that singing dissonant intervals led to an increase in activation in several regions, most notably the PAC, S1, midbrain, amygdala, left putamen, and right posterior insula. The complete list of activations is shown in Table 3 and Fig. 2. Since the region identified as S1 seems to be located directly on the central

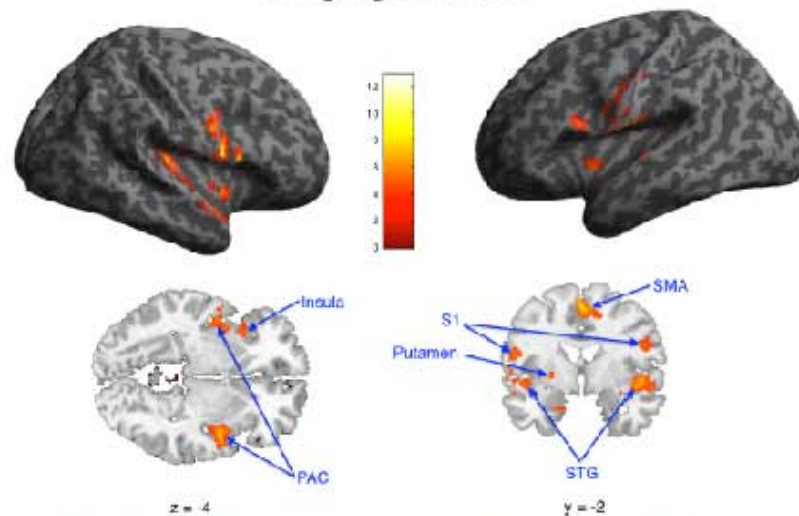
**Table 2**  
Regions of peak neural activity associated with singing minus listening. We present Z-statistic values of thresholded ( $z > 3.00$ ) cluster-corrected image ( $p=0.05$ ) at x, y, z. Cluster extent (CE) is reported in voxels. Coordinates are reported in MNI space.

Singing versus passive listening						
Region		x	y	z	F	CE
<i>Sensorimotor</i>						
S1	L	-52	-15	36	7.06	615
	R	58	10	22	8.62	458
S2	L	-62	-14	14	7.36	615
	R	50	-12	20	6.30	1096
M1	L	-46	-8	36	5.77	135
SMA		0	-2	58	7.62	216
PM	R	54	4	34	11.94	458
BA 44	L	-56	6	20	7.00	615
	R	56	8	14	12.88	458
<i>Auditory</i>						
PAC	L	-46	-12	4	5.87	615
	R	50	-8	4	6.73	1096
STG	L	-62	-16	6	6.08	615
	R	52	6	12	5.69	1096
Temporal pole	L	44	4	18	5.49	615
	R	48	14	-16	6.78	1096
<i>Multimodal</i>						
Insula	R	42	4	-10	8.35	1096
Cerebellum Lobules I–IV	L	-2	-54	0	6.22	148
Basal ganglia						
Putamen	L	-26	0	0	7.22	604
Thalamus	L	-4	-28	0	6.22	148

**Table 1**  
Results of interval identification tests, with standard deviation.

Interval	Tritone	Fifth	Major seventh	Octave	Total
Success rate (%)	65.08 ± 20.68	72.72 ± 20.60	65.44 ± 24.00	59.60 ± 21.36	65.72 ± 21.67

## Singing network



**Fig. 1.** Brain activation for singing, including all intervals. Activation in regions that constitute the “singing network”. Activation maps were superimposed using the cortex\_20484\_surf template in SPM8, cluster significance threshold of  $p=0.05$  (FWE corrected), and MRIcron software.

**Table 3**

Regions of peak neural activity associated with singing dissonant intervals (compared with singing consonant intervals). We present  $Z$ -statistic values of thresholded ( $Z > 3.00$ ,  $p=0.001$  uncorrected) image at  $x$ ,  $y$ ,  $z$ . Cluster extent (CE) is reported in voxels. Coordinates are reported in MNI space.

Singing dissonant intervals > Singing consonant intervals						
Region		$x$	$y$	$z$	$F$	CE
<i>Sensory/motor</i>						
S1	R	58	-4	24	24.00	44
<i>Auditory</i>						
PAC	R	42	-30	10	25.50	262
<i>Emotional</i>						
Amygdala	L	-24	0	-22	17.29	54
<i>Midbrain</i>						
Midbrain	L	-4	-12	-32	21.34	96
<i>Multimodal</i>						
Posterior insula	R	35	-18	-5	25.22	262
<i>Subcortical</i>						
Putamen	L	24	2	4	22.52	62

status, we have used the SPM Anatomy Toolbox (Falkthoff et al., 2005) to confirm that the cluster centered on the coordinate (58, -4, 24) falls within S1 with greater than 80% probability. Meanwhile, we did not detect significant changes in activation while singing consonant intervals.

### 2.2.3. Narrow versus wide intervals

Increases in activation of a variety of regions were observed when singing both narrow and wide intervals. These activations are shown in Table 4 and Fig. 2, where PMd is used to abbreviate dorsal premotor cortex. Singing wide as opposed to narrow intervals resulted in increased activation in the anterior insula, whereas singing narrow as opposed to wide intervals led to increased activation in dorsal premotor cortex and SMA.

### 2.2.4. The effect of training

A regression analysis was performed in the group to test for correlations between BOLD signal and training, where we have used a Bonferroni–Holm (Holm, 1979) correction for multiple comparisons to produce the adjusted  $p$ -values reported here. During the singing minus listening condition, a higher degree of training was positively

correlated with activation in S1 ( $r=0.952$ ,  $p=0.0005$ ), M1 ( $r=0.862$ ,  $p=0.0040$ ), PM ( $r=0.786$ ,  $p=0.0180$ ), SMA ( $r=0.610$ ,  $p=0.0800$ ), and PAC ( $r=0.597$ ,  $p=0.5300$ ) although only the  $p$ -values for S1, M1 and PM are statistically significant. Further, examining the singing dissonant minus consonant condition reveals that a higher level of training positively correlates with activation in putamen ( $r=0.68$ ,  $p=0.002$ ), thalamus ( $r=0.780$ ,  $p=0.002$ ), and PAC ( $r=0.583$ ,  $p=0.070$ ). In this case, only the  $p$ -values for putamen and thalamus are statistically significant. A list of these activations, for both cases, is given in Table 5, and the linear regressions for the particular cases of S1 and the thalamus, as an illustration of one example for each condition, are plotted in Fig. 3.

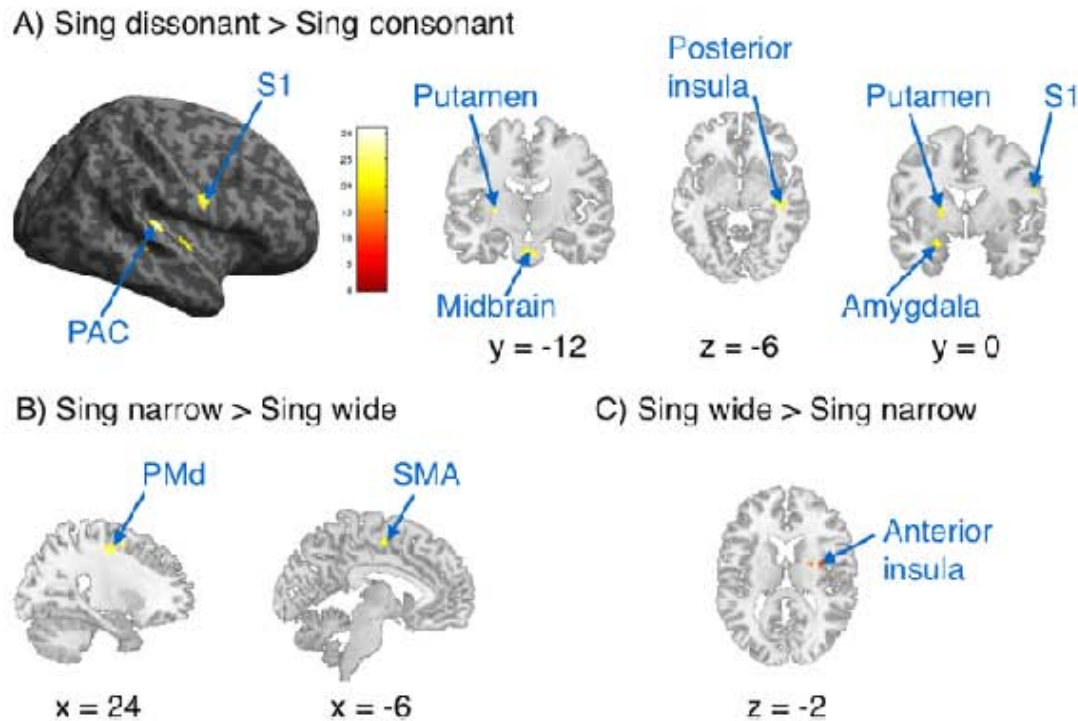
### 2.2.5. Singing accuracy

By performing a linear regression we found a significant correlation between the value  $\delta$ , as defined below, calculated for all four intervals, and the activity within the dorsal premotor cortex ( $xyz=[-32\ 20\ 54]$ ,  $Z=4.30$ ,  $p(\text{cluster})=0.005$ ,  $r=0.7$ ,  $p=0.015$ ). These results are plotted in Fig. 4.

As regards the behavioral results from the fMRI session, the mean value of  $\delta$ , calculated for all four intervals, was 22.74 cents, which is slightly less than a quarter of a semitone. The values of  $\delta$  for each individual interval are given in Table 6. There was no significant difference between the values of  $\delta$  for each of these intervals [ $F(3,40)=0.35$ ,  $p=0.79$ ], so we may conclude that in terms of performance there was no difference between either the tuning of dissonant and consonant intervals, or the tuning of large and narrow intervals.

## 3. Discussion

In this study, we have investigated both the recognition and the production of consonant and dissonant, narrow and wide intervals in semiprofessional singers. A limitation of the present study is the relatively small number of trained musicians involved. Our sample size, however, is not very different from those used in similar studies (Foss et al., 2007; Kleber et al., 2007; Zarate and Zatorre, 2008; Parkinson et al., 2012) and, furthermore, our sample group is quite homogeneous in terms of age, years of training, and general proficiency. As we have mentioned before, singing requires a



**Fig. 2.** Contrast between singing of dissonant and consonant, wide and narrow intervals. (A) Contrast between singing of dissonant as opposed to consonant intervals, showing increased BOLD activation within right S1, right PAC, left putamen and left amygdala. (B) Contrast between singing a narrow as opposed to singing a wide interval, showing increased activation within left SMA, left PAC and right PMd. (C) Contrast between singing a wide as opposed to singing a narrow interval, showing increased activation within right anterior insula and left BA. All maps were thresholded for visual display at threshold of  $p=0.0001$  and extent threshold of 30 voxels.

**Table 4**  
Regions of peak neural activity associated with singing narrow intervals (compared with singing wide intervals) and with singing wide intervals (compared with singing narrow intervals). We present Z-statistic values of thresholded ( $z > 3.00$ ,  $p=0.001$  uncorrected) image at x, y, z. Cluster extent (CE) is reported in voxels. Coordinates are reported in MNI space.

Region		x	y	z	T	CE
Singing narrow intervals > Singing wide intervals						
<i>Motor</i>						
PMd	R	24	-10	24	21.61	149
SMA	L	-6	-6	56	18.56	59
Singing wide intervals > Singing narrow intervals						
<i>Multimodal</i>						
Anterior insula	R	36	-2	14	16.00	25

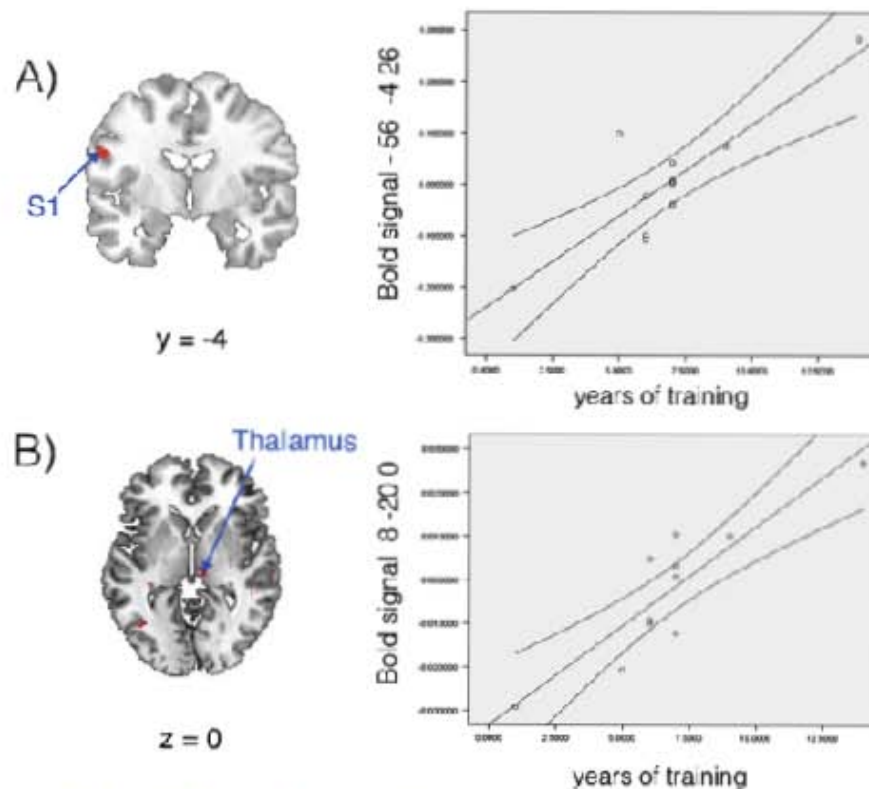
**Table 5**  
Regression analysis testing for correlation between BOLD activity and years of training, presenting Z-statistic values of thresholded ( $z > 3.00$ ,  $p=0.001$  uncorrected) image at x, y, z. Cluster extent (CE) is reported in voxels. Coordinates are reported in MNI space.

Region		x	y	z	Z	CE
Singing minus listening						
<i>Sensory/auditory</i>						
S1	L	-55	-4	26	9.98	959
M1	R	56	2	28	12.57	989
PM	L	-44	-6	50	7.16	989
Singing dissonant intervals minus singing consonant intervals						
<i>Subcortical</i>						
Thalamus	R	8	20	0	5.70	24
Putamen	L	24	2	4	3.43	29

complex interaction between the vocal motor and auditory feedback systems by means of which vocal output is monitored. When singing the required intervals, the participants listen to the sounds produced, then compare to the intended sounds, and finally try to adjust the produced sounds to approach the target sounds as closely as possible.

Behavioral results show that participants identified consonance and dissonance more successfully than they were able to identify exact interval distance. Moreover, no difference in detection and production accuracy was found between consonant and dissonant intervals. That is, neither listening to nor singing the required intervals posed an inordinate challenge to participants. Here we point out that a more reliable source of ad hoc data regarding the difficulty of singing dissonant or wide intervals than behavioral analysis would be the recording of a laryngeal electromyography during vocal production.

Our fMRI results indicated that during the singing of dissonant intervals, as opposed to singing of consonant intervals, there was significant activation of both the primary auditory cortex and the right posterior insula, which form part of a cluster identified by Schirmer et al. (2012) as participating in voice processing, with the posterior insula also playing a role in music perception and singing stability (Terao et al., 2006) and in music-evoked emotions (Knäuper, 2014). Since we also observed activation in this case of the primary somatosensory cortex, a relationship can be established between motor commands and their sensory consequences. It appears then that singing dissonant intervals involves a greater degree of the feedback integration of somatosensory information responsible for articulatory movements than singing consonant intervals. The manner in which the integration of auditory feedback mechanisms with vocal production contributes to proper singing has been the subject



**Fig. 3.** Correlation between years of training and activation of brain regions. (A) Linear regression expressing the correlation between years of training and the parameter estimates of the contrast of singing minus listening in the peak voxel of the S1 ( $y = -4$ ,  $z = 26$ ),  $r = 0.952$ ,  $p = 0.0005$ ,  $p_{\text{FWE}} < 0.05$  SVC). (B) Linear regression expressing the correlation between years of training and the parameter estimates of the contrast of singing dissonant minus consonant intervals in the peak voxel of the thalamus ( $y = -20$ ,  $z = 0$ ),  $r = 0.780$ ,  $p = 0.002$ ,  $p_{\text{FWE}} < 0.05$  SVC). BOLD responses are reported in arbitrary units.

of neuroimaging studies by Zarate and Zatorre (2008), Parkinson et al. (2012), and Kleber et al. (2013). Singing dissonant intervals also involves greater activation of the left putamen than singing consonant intervals. Putamen activity is thought to accompany the recall of a well-learned motor program for correcting inaccurate vocal output (Doyon et al., 2003). It has been suggested that this putamen activity contributes to people learning to monitor their auditory feedback and adjust their vocal output accordingly. The left putamen, in particular, has been shown to play a part in processing auditory feedback towards adjusting the pitch of sung notes, with greater activation in musicians than in non-musicians (Zarate and Zatorre, 2008). Thus, it would then seem likely that dissonant intervals are produced by adjusting the mechanisms already used for the vocal production of consonant intervals. The increased functional activation of the amygdala is consistent with observations made by Pallesen et al. (2005) when comparing listener response to consonant and dissonant chords, and also with the study by Koelsch et al. (2006) where activations in the amygdala accompany unpleasant musical information, which, in turn, is associated with dissonance (Gosselin et al., 2006). The amygdala forms part of the brain structures which are known to be associated with emotion, and which are susceptible to have their activity modulated by musical stimuli (Koelsch, 2014). Blood et al. (1998) find that increasing consonance and dissonance both correlate well with changes in areas known to be involved in affective processing. They observe that increasing dissonance correlates with greater activity in the right parahippocampal gyrus, which, although normally associated with memory and learning, has strong reciprocal connections with the amygdala.

We have also established that longer training correlates positively with increased activity in the right thalamus and the left

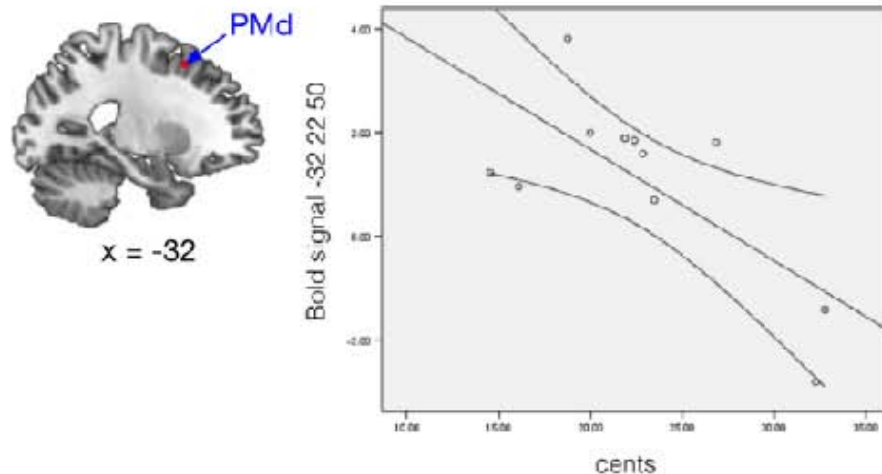
**Table 6**

Values of  $\delta$  for each sung interval, accompanied by standard deviation.

Interval	Tri-tone	Fifth	Major seventh	Octave	Total
Mean ( $\delta$ ) (cents)	24.24	21.61	22.42	22.65	22.74
Std. deviation	3.57	5.97	6.03	6.85	5.98

putamen when singing dissonant intervals. The role of the left putamen in pitch adjustment has been described above, but since the ventrolateral thalamus is thought to form, together with the motor cortex, a system for adaptive motor control (Kleber et al., 2010), we can reasonably argue that the ability to adjust intervals while singing increases with appropriate training. The activation of the putamen is also significant in that the basal ganglia has been identified as one of the regions central to performing pitch and temporal events in the correct order during expert performance (Brown et al., 2015). Increased activity in the thalamus has also been linked with passive listening to dissonant chords by Foss et al. (2007). More generally, the effect of training on vocal production is expected to give rise to increased activation of the bilateral somatosensory cortex (Kleber et al., 2010), which we did observe very clearly, and also to a series of functional adaptations accompanied by anatomical changes in motor and premotor regions (Bengtsson et al., 2005), where we also detected increased activation.

With regard to the ability to sing correctly both consonant and dissonant intervals, we have found that increased activity in the left dorsal premotor cortex is correlated with singing in tune. This finding is in agreement with previous studies showing that



**Fig. 4.** Correlation between singing in tune and activation of dorsal premotor cortex. Linear regression expressing the correlation between  $\delta$  and the parameter estimates of the contrast of singing minus listening in the peak voxel of dorsal premotor cortex activation ( $xyy = [-32, 22, 50]$ ,  $r = 0.700$ ,  $p = 0.015$ ,  $p < 0.05$  SVC). BOLD responses are reported in arbitrary units.

activation of the left dorsal premotor cortex represents the selection of an upward or downward vocal adjustment to counteract the pitch shift (Zarate and Zatorre, 2008). In fact, a number of neuroimaging studies (Özdemir et al., 2006; Zarate and Zatorre, 2008; Kleber et al., 2010) have shown that the premotor cortex is key to achieving integration between sensory information and motor action during speech and singing. Since there is no significant difference observed between the ability to sing consonant and dissonant intervals in tune, we may only conclude that in both cases some degree of motor adjustment is needed in order to produce intervals sung at correct pitch.

If we now examine the effect of singing wide and narrow intervals without taking into account whether they are consonant or dissonant, we initially report no correlation between singing wide or narrow intervals and significant activation of any brain region. Again complementing these results by conventional subtraction analysis, we observed that singing narrow intervals as opposed to singing wide intervals resulted in the activation of SMA and PMd, which are mainly associated with motor control of vocal production. Conversely, we observed significant activation of the right anterior insula when participants were asked to sing wide as opposed to narrow intervals. In the case of trained singers, the insula serves as a gateway for somatosensory (Arckermann and Riecker, 2010; Jürgens, 2002) and auditory (Augustine, 1996) information. Based on the observation that the auditory cortex, the insula, the anterior cingulate cortex (ACC), and the intraparietal sulcus (IPS) are all functionally connected, Zarate and Zatorre (2008) and Zarate et al. (2010) suggest that the anterior insula may act as a hub for the voluntary vocal motor adjustments necessary to maintain an originally produced note, or correct for a perceived shift in pitch. In a recent study by Kleber et al. (2013), somatosensory feedback was reduced in both singers and non-musicians by means of the application of a topical anesthetic to the vocal chords prior to singing. The authors found that singers were able, through vocal training and practice, to regulate activity in the right anterior insula in order to ignore temporarily somatosensory feedback, and fall back instead on internal models of motor control which were then deemed more reliable. These results are perfectly consistent with Critchley et al.'s (2004) finding that the right anterior insula mediates explicit awareness of internal bodily processes. Also noteworthy in this context is the suggestion made by Uddin (2014) that the anterior insula is part of the network in charge of salience processing in the brain. Anterior insula activity only accompanies

the singing of wide intervals, which we confirm to be more complicated from a technical standpoint than the singing of narrow intervals because a greater degree of vocal control is required to adjust the sung notes.

In conclusion, we suggest that singing dissonant or wide intervals requires greater involvement of neural mechanisms associated with integrating internal and external feedback from auditory and sensorimotor systems than singing consonant or narrow intervals. The neural networks involved in singing wide and dissonant intervals, however, are quite different, and so we postulate that singing these types of intervals is difficult for different reasons. We obtained evidence that dissonant intervals are intoned by adjusting the neural mechanisms used for the production of consonant intervals. On the other hand, singing wide intervals does seem to require integration of auditory and sensorimotor systems, while singing narrow intervals does not. The training of vocal skills seems to be related to an increased ability to adjust the singing voice efficiently in response to auditory monitoring.

## 4. Methods and materials

### 4.1. Subjects

Eleven healthy subjects were recruited from conservatory. Both Itoh et al. (2010) and Foss et al. (2007) report significant differences in evaluation of noncontextual dissonance between musicians and non-musicians, suggesting some degree of plasticity in the underlying neural mechanisms, and motivating our choice of subjects as only musicians who were further required to be able to sing precise intervals. Thus, all subjects were right-handed women (15–21 years old) with 7–15 years of formal training in Western tonal music and singing, had normal hearing, and were devoid of neurological or psychological disorders and contraindications for fMRI techniques. Audiometric tests were performed to establish that the subjects' hearing was normal, and it was indeed confirmed that none of the subjects suffered from pronounced hearing loss at any frequency range. Only female subjects of approximately the same tessitura (soprano) participated in the study so that the difficulty of singing wide intervals within an appropriate range of notes would be reasonably the same for all subjects. All subjects gave written informed consent to participate in this study, in accordance with procedures approved by the

Research, Ethics and Safety Committee of the Hospital Infantil de México Federico Gómez, which, further, specifically approved this study.

#### 4.2. Behavioral pre-test

In a first session, the subjects' capacity to identify the intervals themselves, and to determine whether they were consonant or not was evaluated. We presented the intervals entirely devoid of musical context. During this session, each subject sat in front of a laptop computer screen and wore headphones, through which all auditory stimuli were delivered. Pure tones with a duration of one second were generated with Matlab producing a sine wave of the desired frequency  $f$  sampled at a rate of 44,100 Hz. The frequencies were calculated according to standard equal temperament, with  $A_4$  tuned at 440 Hz. Intervals were then produced by the straightforward juxtaposition of two of these tones, so that the total duration of the interval was two seconds. These intervals were then played back to the test subjects via headphones directly from a laptop computer. They were then asked to respond by pressing appropriate keys on a laptop keyboard. Two consonant and two dissonant intervals were selected for these identification tests according to criteria explained in the next section. In a two-alternative forced-choice task, subjects were first asked to identify whether the intervals they heard were consonant or dissonant, and then, in a four-alternative forced-choice task, asked to identify the intervals themselves. The number of successful attempts and errors was measured and analyzed. The analysis of accuracy was conducted through Student's  $t$  test for two-alternative tasks, and using ANOVAs for four-alternative tasks. This behavioral interval perception task was used as confirmation that the participants had no trouble identifying the intervals, and could therefore perform adequately once inside the MR scanner.

#### 4.3. fMRI session

In a second session, the same subjects entered a Philips Achieva 1.5-T magnetic resonance (MR) scanner. During data acquisition the subjects performed singing tasks for which they were not previously trained. The control condition involved passive listening, during which the subjects were presented with a musical interval, which they did not have to sing back, where two notes with a duration of one second each were played one immediately after the other. For the singing task, after hearing the two second target interval subjects were cued to reproduce the same notes for the same duration of time. Subjects were asked to hum, rather than sing the notes, producing an "m" sound while keeping the mouth closed and the lips still. In this manner, sound emerges from the nostrils, allowing us to reduce movement artifacts during the fMRI session.

##### 4.3.1. Stimuli

In order to test both the ability to recognize and to sing consonant and dissonant intervals, we chose sets of two representative intervals for each case: an octave and a fifth as consonant intervals, and a major seventh and a tritone, or augmented fourth, as dissonant intervals. The intervals were produced as juxtaposed pure tones with a duration of one second each, and were tuned according to equal temperament, so that the ratios associated with each interval, with the exception of the octave, were only approximate. The ratios corresponding to the selected intervals are shown in Table 7. Notice that the ratios associated with the fifth (3:2) and octave (2:1) do involve smaller integers than those associated with the tritone (45:32) and major seventh (15:8). The octave and the fifth are, for a variety of reasons, considered to be the most consonant intervals, and, conversely, the

**Table 7**  
Western musical intervals selected. The intervals are given by their most common name. The number of equal-temperament semitones (ST) each of them comprises is indicated, as is the approximate frequency–also each of them is associated to.

ST	Name	Ratio
5	Tritone	45:32
7	Perfect fifth	3:2
11	Major seventh	15:8
12	Octave	2:1

tritone has historically served as the archetype of a dissonant interval. Along with the tritone, the major seventh and the minor second are generally cited as examples of dissonant intervals (Beldeman and Krishnan, 2009; Lutz and Stone, 2008). As we have discussed in the introduction, since all participants have been trained to listen to and sing consonant intervals more often than dissonant intervals, the variables dissonance/consonance and unfamiliarity/familiarity are to some degree inherently confounded. That we used two consonant and two dissonant intervals during our tests is due to the fact that correctly singing wide intervals, such as the octave and the major seventh, is technically more demanding than singing narrower ones, such as the tritone and the fifth. The notes produced were all in the range comprised between C4 and C6. This range fell within the natural tessitura of the participants; all of them female, and was chosen to allow them to sing the notes evenly and without undue effort. Thus, by using these four intervals we expected to be able to distinguish between differences, if any, which correlated with consonance and dissonance, and those which correlated with the degree of difficulty of singing an interval.

##### 4.3.2. Equipment

Stimuli were generated by a PC laptop computer running E-PRIME (Psychology Software Tools, Pittsburgh, USA) and were delivered to the subject binaurally via one of two types of headphones. When played to participants outside the imaging room, a pair of Sony MDR-7555 stereo headphones were used, and inside the scanner participants were equipped with a custom-made pneumatic, non-magnetic MR-compatible headset. Sound levels were again adjusted to comfortable levels for each subject before the experiment began. With regard to the actual recording of the sung voice as produced inside the MR-scanner, it has been established by Hurwitz et al. (1989) that condenser microphones are suitable for use in the presence of the intense magnetic fields in the imaging environment because their sensitivity to acoustic signals is not affected by such fields. We employed the same microphone used by Ravetz et al. (2000), a Shure SM93 pro audio condenser microphone, connected to a custom bias voltage power supply located outside the room housing the scanner through a 10 m shielded twisted-triplet cable. This cable acted, effectively, as a Faraday cage and prevented the appearance of imaging artifacts due to the presence of the cable. The response of the microphone to a test tone inside and outside the scanner was compared in order to establish that the microphone was behaving as expected inside the scanner. Since the participants were asked to hum rather than sing, and were thus asked not to open their mouths, the microphone was taped directly to the participants' neck just below the larynx. Although there was a significant amount of background noise during scanning, as established by, among others, Cho et al. (1997) and Ravetz et al. (2000) (peak noise levels of 123 dB are registered inside a 1.5T MR-scanner), this setup permitted a sufficiently faithful recording of the participants' voices through a Tascam 135-144 MKII interface connected to laptop computer. The acquired signals were then processed using the Cool Edit Pro 2.1 suite, essentially filtering the signals through a Hamming

window. It then became straightforward to identify the sung fundamentals using an autocorrelation algorithm.

#### 4.3.3 Procedure

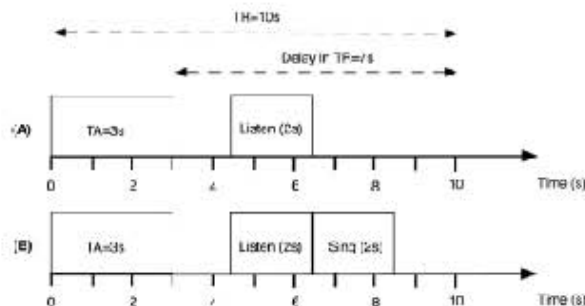
During scanning, subjects were exposed first to the listening and then to the singing conditions described previously in Section 4.3. The listening condition was presented 10 times at the beginning of each run, with the four different intervals appearing in random order, followed by 40 iterations of the singing condition, with the four intervals appearing in pseudorandomized order, but in such a manner that eventually each interval was presented exactly 10 times. Each subject went through three experimental runs, each one lasting 500 s, resulting in a total of 30 acquisitions for the listening condition and 120 acquisitions for the singing condition.

#### 4.3.4 MR data acquisition

We utilized a sparse sampling, event-related experimental design, in which tasks were performed during the silences between image acquisitions to prevent scanner noise from interfering with the auditory stimuli and to reduce any effect of movement (Belin et al., 1999). Timings of task presentations were systematically varied or "jittered" by  $\pm 500$  ms to maximize the likelihood of obtaining the peak of the hemodynamic response for each task. Further, we initially acquired and discarded three EPI "dummy" scans to allow for T1-saturation effects. Functional T2\* weighted whole-head scans (150 volumes) were collected with blood oxygen level dependent (BOLD) contrast (TE=40 ms, TR=10 s, TA=3 s, delay in TR=7 s, flip angle=90°, FOV=256 mm<sup>2</sup>, and matrix dimensions of 64 × 64), covering the entire brain (34 interleaved slices, 4 mm thickness, parallel to the anterior-posterior commissural plane, and with a voxel size of 4 × 4 × 4 mm<sup>3</sup>). Thus, listening and singing took place within an interval of 7 s of silence, leaving 3 s for acquisition. The sparse sampling design is shown in Fig. 5. Prior to the functional scan, a T1-weighted anatomical volume was acquired using a gradient echo pulse sequence (TR=10.2 ms, TE=4.2 ms, flip angle=90°, and a voxel size of 1 × 1 × 1 mm<sup>3</sup>).

#### 4.3.5 MR data analysis

Functional data was pre-processed using SPM8 (Wellcome Department of Cognitive Neurology, London, UK; see <http://www.fil.ion.ucl.ac.uk/spm>), following standard procedure. Images were realigned to the first volume, resampled and spatially normalized (final voxel size 2 × 2 × 2 mm<sup>3</sup>) to standard Montreal Neurological Institute (MNI) stereotactic space (Evans et al., 1994) and smoothed using an isotropic 8 mm FWHM Gaussian kernel. The



**Fig. 5.** 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. Participants listen to a two-tone interval (2 s) and then match the target tones with their singing voice (2 s). Each block consists of 10 randomized listen-only trials, followed by 40 listen-then-sing trials, pseudorandomized in order to present exactly 10 trials corresponding to each of the 4 intervals.

movement parameters obtained during preprocessing were also included in the model in what is effectively a realignment procedure.

The passive listening and singing tasks were modeled with silence scans serving as an implicit baseline. The 10 s long time interval between scans was divided into 80 time-bins. In order to diminish the variance associated with both noise and predictors, we have high-pass filtered (1000 s for the "singing network," as modelling all singing tasks together will amount to a block of 400 s, and 360 s elsewhere) the blood-oxygen-level dependence (BOLD) response for each event. Since we wanted to make as few assumptions as possible with respect to the hemodynamic response, the BOLD response was modelled using a single-bin finite impulse response (FIR) basis function (order 1), or boxcar function, spanning the time of volume acquisition (3 s).

First, we sought to identify activity during the singing condition in regions constituting the "singing network" as defined in a number of previous studies (Kleber et al., 2007, 2010, 2013; Zarate and Zatorre, 2008), in this way validating the basic design. To this end, for each participant, on first level analysis, we calculated a linear contrast of interest relative to the effect of singing minus listening [(sing tritone + sing fifth + sing seventh + sing octave) - listening], where we used one-sample t-tests with statistical maps thresholded by means of clusters determined by  $Z > 3$  and a (corrected) cluster significance threshold of  $p=0.05$ , as in Worsley (2001). Note that we did not use the "singing network" as a mask at any moment later on.

Afterwards, four further linear contrasts of interest were calculated for each participant, on first level analysis, corresponding to the effects of singing tritone minus listening to tritone [(sing tritone) - (listen tritone)]; singing fifth minus listening to fifth [(sing fifth) - (listen fifth)]; singing seventh minus listening to seventh [(sing seventh) - (listen seventh)]; singing octave minus listening to octave [(sing octave) - (listen octave)]. Using these contrasts, we compared the singing of narrow and wide intervals as well as the singing of consonant and dissonant intervals in order to be able to take into account the possible effects of an increased difficulty associated with the singing of the wider intervals. Each of these linear contrasts was taken to a second level random-effects model using whole brain analysis, namely, a set of 2-way repeated measures ANOVAs with the factors consonance/dissonance and narrow/wide interval in order to determine the areas activated or deactivated for the various possible interactions. Activations of interest were identified according to the significance threshold  $p < 0.001$ , uncorrected for multiple comparisons, and an extent threshold of 20 voxels for subsequent post hoc analyses and plotting.

We also examined the correlation between years of training and brain activity. To do so, we again used the singing minus listening contrast described above, and we calculated two further contrasts for each participant, on first level analysis: singing dissonant minus consonant intervals [(sing tritone + sing seventh) - (sing fifth + sing octave)], and singing wide minus narrow intervals [(sing seventh + sing octave) - (sing tritone + sing fifth)]. These contrasts were first taken to a second-level group whole-brain random-effects multiple regression analysis to test for correlations between BOLD response and years of training. First-level parameter estimates for the condition of interest were then extracted from the peak voxel of the significant activation for each subject, and finally analyzed, post hoc, using Spearman's rank-order correlation with SPSS.

Finally, we investigated the correlation between singing in tune and brain activity. In order to establish this correlation, we first determined, for each trial in the singing condition, a measure of how well each student manages to sing in tune by calculating the root-mean-square (RMS) value of the differences between the



presented and the sung notes, where the mean fundamental frequency was extracted through an autocorrelation algorithm, and the expected notes tuned according to standard equal temperament. For the sake of simplicity we shall call this parameter  $\delta$ , with cents as units. Thus, a high value of  $\delta$  is associated with singing out of tune while a small value of  $\delta$  is associated with singing in tune. Before carrying out all statistical analyses, we eliminated outliers by excluding any sung intervals for which the value of  $\delta$  exceeds 50 cents. That is, we excluded sung intervals which approached an interval different from the target interval more closely than the target interval itself. The regression analysis by means of which we determined those areas in which the BOLD signal was significantly correlated with an interval being sung in tune was identical to that described above, using, once more, the singing minus listening contrast.

### Conflict of interest

The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

### Acknowledgements

We thank Pilar Díez Suárez and Manuel Obregón Espejos for their invaluable assistance during fMRI acquisition. We would also like to acknowledge Ricardo Durantes Escamilla and Antonio Pérez López, who worked on the microphone and cables employed during fMRI acquisition, and Hugo Ezequiel Gutiérrez Hernández, who advised on the processing and analysis of recorded sound.

The authors gratefully acknowledge financial support for this research from Secretaría de Salud de México (HIM/2011/058 SSA-1009), CONACYT (SALUD-2012-01-182160), and from DGAPA UNAM (PAPIIT IN109214). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

### References

Ackermann, H., Riecker, A., 2010. The contribution(s) of the insula to speech production: a review of the clinical and functional imaging literature. *Brain Struct. Funct.* 214, 419–431.

Augustine, J.R., 1996. Circuitry and functional aspects of the insular lobe in primates including humans. *Brain Res. Brain Res. Rev.* 22, 229–244.

Bangert, M., Altenmüller, E.C., 2003. Mapping perception to action in piano practices: a longitudinal de eeg study. *BMC Neurosci.* 4, 1–14.

Belin, P., Zatorre, R.J., Hoge, R., Evans, A.C., Pike, B., 1999. Event-related fMRI of the auditory cortex. *Neuroimage* 10 (4), 417–429.

Bengtsson, S.L., Nagy, Z., Skarn, S., Forsman, I., Forsberg, H., Hillen, F., 2005. Extensive piano practicing has regionally specific effects on white matter development. *Nat. Neurosci.* 8, 1148–1150.

Bidelman, G.M., Krishnan, A., 2009. Neural correlates of consonance, dissonance, and the hierarchy of musical pitch in the human brainstem. *J. Neurosci.* 29 (October (42)), 13165–13171.

Blond, A.J., Zatorre, R.J., Bermudez, P., Evans, A.C., 1999. Emotional responses to pleasant and unpleasant music correlate with activity in paralimbic brain regions. *Nat. Neurosci.* 2, 382–387.

Brown, R.M., Zatorre, R.J., Penhune, V., 2015. Expert music performance: cognitive, neural and developmental bases. *Prog. Brain Res.* 217, 57–86.

Brown, S., Martinez, M.J., Hodges, D.A., Fox, P.T., Parsons, L.M., 2004a. The song system of the human brain. *Cogn. Brain Res.* 20 (3), 363–375.

Brown, S., Martinez, M.J., Parsons, L.M., 2004b. Passive music listening spontaneously engages limbic and paralimbic systems. *Neuroreport* 15, 2033–2037.

Burns, E.M., 1999. Intervals, scales, and tuning. In: Deutsch, D. (Ed.), *The Psychology of Music*, 2nd edition Academic Press, London, pp. 215–254.

Callan, D.E., Tsytarev, V., Hasekawa, T., Galen, A.M., Katsuhara, M., Fukuyama, H., Turner, R., 2006. Song and speech: brain regions involved with perception and covert production. *Neuroimage* 31, 1327–1342.

Choi, Z.H., Park, S.H., Kim, J.H., Chung, S.C., Chung, S.T., Chung, J.Y., Moon, C.W., Yi, J. H., Sin, C.H., Wong, F.K., 1997. Analysis of acoustic noise in MRI. *Magn. Reson.*

*Imaging* 15, 815–822.

Courineau, M., McDermott, J.H., Peretz, I., 2012. The basis of musical consonance as revealed by congenital amusia. *Proc. Natl. Acad. Sci. USA* 109 (48), 19858–19863.

Critchley, H.D., Wiens, S., Richlstein, P., Östman, A., Dolan, R.J., 2004. Neural systems supporting interoceptive awareness. *Nat. Neurosci.* 7, 189–195.

Deyon, J., Penhune, V., Ungerleider, L.G., 2003. Distinct contribution of the corticostriatal and cortico-cerebellar systems to motor skill learning. *Neuropsychologia* 41 (2), 252–262.

Druy, P.A., 1951. *Bel Canto in its Golden Age: A Study of its Teaching Concepts*. King's Crown Press, New York.

Eickhoff, S.B., Stephan, K.E., Mohler, H., Grefkes, C., Fink, G.R., Amunts, K., Eles, K., 2005. A new spin toolbox for combining probabilistic cytoarchitectonic maps and functional imaging data. *Neuroimage* 25, 1325–1335.

Evans, A.C., Kamber, M., Collins, D., McDonald, D., 1994. An MRI-based probabilistic atlas of neuroanatomy. In: Shorvon, S.D., Fish, D.R., Andermann, E., Bydder, G. M., Stefan, H. (Eds.), *Magnetic Resonance Scanning and Epilepsy* vol. 264 Plenum Press, New York, pp. 263–274.

Forde Thompson, W., 2013. Intervals and scales. In: Deutsch, D. (Ed.), *The Psychology of Music*. Academic Press, Amsterdam, pp. 107–140.

Foss, A.H., Aitchison, E.L., James, K.H., 2007. Neural correlates of the pythagorean ratio rules. *Neuroreport* 18, 1521–1525.

Cosselin, N., Samson, S., Adolphs, R., Noulhiane, M., Roy, M., Hasboun, D., Bault, M., Peretz, I., 2006. Emotional response to unpleasant music correlates with damage to the parahippocampal cortex. *Brain* 129, 2585–2592.

Helmholtz, H., 1954. *On the Sensations of Tone*. Springer, New York.

Holm, S., 1979. A simple sequentially rejective multiple test procedure. *Scand. J. Stat.* 6, 65–70.

Hurwitz, R., Lane, S.R., Bell, R.A., Brant-Zawadzki, M.N., 1989. Acoustic analysis of gradient-coil noise in MR imaging. *Radiology* 173 (2), 545–549.

Itoh, K., Sawazono, S., Nakada, T., 2010. Central auditory processing of non-contextual consonance in music: an evoked potential study. *J. Acoust. Soc. Am.* 128, 3781–3787.

Jürgens, U., 2002. Neural pathways underlying vocal control. *Neurosci. Biobehav. Rev.* 26, 235–258.

Jürgens, U., Ploog, D., 1970. Cerebral representation of vocalization in squirrel monkey. *Exp. Brain Res.* 10 (5), 532–554.

Jürgens, U., Richter, K., 1986. Glutamate-induced vocalization in the squirrel monkey. *Brain Res.* 373 (1–2), 349–358.

Kleber, B., Birbaumer, N., Lotze, M., Zatorre, R.J., Lotze, M., 2007. Quiet and imagined singing of an Italian aria. *Neuroimage* 26, 889–900.

Kleber, B., Yek, R., Birbaumer, N., Gruzelier, J., Lotze, M., 2010. The brain of opera singers: experience-dependent changes in functional activation. *Cereb. Cortex* 20 (5), 1144–1152.

Kleber, B., Zeigler, 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 (14), 6070–6080.

Koelsch, S., 2014. Brain correlates of music-evoked emotions. *Nat. Rev. Neurosci.* 15, 170–180.

Koelsch, S., Fritz, T., Cramon, D.Y.V., Müller, K., Friederici, A.D., 2006. Investigating emotion with music: an fMRI study. *Hum. Brain Mapp.* 27, 239–250.

Lahav, A., Saltzman, E., Schlaug, G., 2007. Action representation of sound: audio-motor recognition network while listening to newly acquired actions. *J. Neurosci.* 27 (2), 308–314.

Lee, I.S., Shin, I., 2008. Perception of musical consonance and dissonance: an outcome of neural synchronization. *J. R. Soc. Interface* 5, 1479–1484.

McDermott, J.H., Lehr, A.J., Oxenham, A.J., 2010. Individual differences reveal the basis of consonance. *Curr. Biol.* 20 (June (11)), 1035–1041.

McLachlan, N., Marco, D., Light, M., Wilson, S., 2013. Consonance and pitch. *J. Exp. Psychol. - Gen.* 142, 1142–1158.

Minati, L., Rosazza, C., D'Inceri, L., Pietroschi, E., Valentini, L., Scaini, V., Loveday, C., Bruzzone, M.C., 2009. Functional MRI event-related potential study of sensory consonance and dissonance in musicians and nonmusicians. *Neuroreport* 20, 87–92.

Özdemir, E., Norton, A., Schlaug, G., 2006. Shared and distinct neural correlates of singing and speaking. *Neuroimage* 33 (2), 628–635.

Pallebage, K.J., Brattico, E., Bailey, C., Norvén, A., Nivola, J., Gjedde, A., Carlson, S., 2005. Emotion processing of major, minor, and dissonant chords: a functional magnetic resonance imaging study. *Ann. N.Y. Acad. Sci.* 1060, 450–453.

Parkinson, A.L., Hagenauer, S.G., Manes, J.L., Larson, C.R., Rogers, B., Robin, D.A., 2012. Understanding the neural mechanisms involved in sensory control of voice production. *Neuroimage* 61, 314–322.

Peretz, I., Bédard, A.J., Penhune, V., Zatorre, R., 2001. Cortical deafness to dissonance. *Brain* 124, 928–940.

Perry, D.W., Zatorre, R.J., Penhune, M., Alvarado, B., Meyer, E., Evans, A.C., 1999. Localization of cerebral activity during simple singing. *Neuroreport* 10, 3979–3984.

Rauve, M.F., Moller, J.R., Klug, N.Y.S., 2000. Acoustic noise during functional magnetic resonance imaging. *J. Acoust. Soc. Am.* 108 (4), 1687–1696.

Schellenberg, E.G., Trehub, S.E., 1994. Frequency ratios and the discrimination of pure tone sequences. *Percept. Psychophys.* 56, 472–478.

Schellenberg, E.G., Trehub, S.E., 1995. Children's discrimination of melodic intervals. *Dev. Psychol.* 32, 1039–1050.

Schirmer, A., Mickle-Fox, P., Grandjean, D., 2012. On the spatial organization of sound processing in the human temporal lobe: a meta-analysis. *Neuroimage* 63, 137–147.

- Terao, Y., Mizuno, T., Shindo, M., Saizawa, Y., Ugawa, Y., Kobayashi, S., 2006. Vocal amusia in a professional tango singer: due to a right superior temporal cortex infarction. *Neuropsychologia* 44, 479–488.
- Terhardt, E., 1974. Pitch, consonance, and harmony. *J. Acoust. Soc. Am.* 55 (5), 1061–1069.
- Ternström, S., Sundberg, J., 1982. Acoustical factors related to pitch precision in choir singing. *Speech Music Hear. Q. Prog. Status Rep.* 23, 75–90.
- Ternström, S., Sundberg, J., 1988. Intonation precision of choir singers. *J. Acoust. Soc. Am.* 84, 59–69.
- Trainor, L.J., Heimiller, R.M., 1998. The development of evaluative responses to music: infants prefer to listen to consonance over dissonance. *Infant Behav. Dev.* 21, 77–88.
- Tsuno, M.I., Bharucha, H., Krumhansl, C.L., 1990. Music perception and cognition following bilateral lesions of auditory cortex. *J. Cogn. Neurosci.* 2, 195–212.
- Uddin, L.Q., 2014. Salience processing and insular cortical function and dysfunction. *Nat. Rev. Neurosci.* 16, 55–61.
- Worsley, K.J., 2001. Statistical analysis of activation images. In: Jefferies, P., Matthews, P.M., Smith, S.M. (Eds.), *Functional MRI: An Introduction to Methods*. Oxford University Press, Oxford, pp. 251–270 (Chapter 14).
- Zarate, J.M., 2013. The neural control of singing. *Front. Hum. Neurosci.* 7, 1–12.
- Zarate, J.M., Wood, S., Zatorre, R.J., 2010. Neural networks involved in voluntary and involuntary vocal pitch regulation in experienced singers. *Neuropsychologia* 48 (2), 607–618.
- Zarate, J.M., Zatorre, R.J., 2008. Experience-dependent neural substrates involved in vocal pitch regulation during singing. *Neuroimage* 40 (4), 1871–1887.
- Zentgraf, M.R., Kagan, J., 1996. Perception of music by infants. *Nature* 383, 29.
- Zentgraf, M.R., Kagan, J., 1998. Infants' perception of consonance and dissonance in music. *Infant Behav. Dev.* 21, 433–492.