



**UNIVERSIDAD NACIONAL AUTÓNOMA DE MÉXICO**  
DOCTORADO EN CIENCIAS BIOMÉDICAS  
INSTITUTO DE FISIOLÓGÍA CELULAR

**CORRELATOS NEURONALES DE LA CATEGORIZACIÓN AUDITIVA EN LA  
CORTEZA PREMOTORA MEDIAL**

**TESIS**  
QUE PARA OPTAR POR EL GRADO DE:  
DOCTOR EN CIENCIAS

PRESENTA:  
**ISAAC MORÁN MARTÍNEZ**

TUTOR PRINCIPAL:  
DR. LUIS ALONSO LEMUS SANDOVAL - INSTITUTO DE FISIOLÓGÍA  
CELULAR, UNAM

MIEMBROS DEL COMITÉ TUTOR:  
DR. HUGO MERCHANT NANCY - INSTITUTO DE NEUROBIOLOGÍA, UNAM  
DR. VÍCTOR HUGO DE LAFUENTE FLORES - INSTITUTO DE  
NEUROBIOLOGÍA, UNAM

CIUDAD UNIVERSITARIA, CD. MX., 2022  
ENERO



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.

## DEDICATORIA

A Refugio y Leticia por ser el inicio de todo esto.

## AGRADECIMIENTOS

Agradezco a la UNAM por darme la oportunidad de formarme académicamente dentro de sus aulas. A todos aquellos profesores y colegas que compartieron conmigo parte de su tiempo y conocimientos. En especial, agradezco al Dr. Luis Lemus por darme un espacio en su laboratorio para desarrollar este proyecto y brindarme la oportunidad de aprender y trabajar junto a él. Al Dr. Javier Pérez por su gran apoyo en la realización de este proyecto. A los Dres. Hugo Merchant y Víctor de Lafuente por compartir su ayuda y consejos durante mi formación como estudiante.

Le doy las gracias a mis compañeros de laboratorio por compartirme su amistad y enseñanzas: Elizabeth, Jonathan, Tonatiuh, Fabiola y Miguel. Particularmente, agradezco el apoyo de Tonatiuh en los aspectos técnicos del laboratorio. A Gabriel e Ivonne por su labor en el cuidado de los monos. A Vaporel y Xomara por su invaluable participación en el proyecto.

Sin duda, agradezco a mi familia, ya que han sido el pilar central de mi formación. A mis abuelos por su gran fortaleza y sabiduría para enfrentar la vida. A mis tías y primos por su apoyo y cariño incondicionales. A mi madre y hermanos por ser un ejemplo de fuerza y alegría. A Mili por compartirme su cariño y consejos durante este trayecto.

Finalmente, agradezco a los organismos que financiaron la realización de este proyecto: CONACYT CB-256767 y Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT) IN207919. Al Doctorado en Ciencias Biomédicas y a CONACYT por la beca 288348, recibida durante la realización del posgrado.

# ÍNDICE

RESUMEN.....	1
INTRODUCCIÓN .....	2
Fenómeno de categorización .....	2
¿Qué sabemos acerca de la representación de categorías en el cerebro? .....	3
Neurofisiología del procesamiento auditivo en primates .....	4
¿Cómo se representan las categorías en el sistema auditivo? .....	7
¿Qué tipos de procesamiento categórico ocurre en áreas subsecuentes de la jerarquía cortical? .....	8
Problemáticas para entender los procesos neuronales asociados al reconocimiento y la categorización auditiva en modelos animales .....	9
PLANTEAMIENTO DEL PROBLEMA.....	11
HIPÓTESIS .....	11
OBJETIVOS .....	11
Objetivo general.....	11
Objetivos particulares .....	11
MÉTODOS Y RESULTADOS .....	12
DISCUSIÓN .....	13
Los monos categorizan exitosamente estímulos auditivos .....	13
Las neuronas de CPM representan información de las categorías auditivas .....	14
CPM se organiza en dos estados poblacionales distintos durante los periodos auditivo y de movimiento.....	15
Los cambios lineales de los estímulos modulan el comportamiento y la actividad neuronal de forma categórica .....	16
La condición pasiva y los errores indican que CPM correlaciona con las decisiones categóricas .....	17
Consideraciones finales .....	18
CONCLUSIÓN .....	18
REFERENCIAS .....	18
APÉNDICE.....	25

## RESUMEN

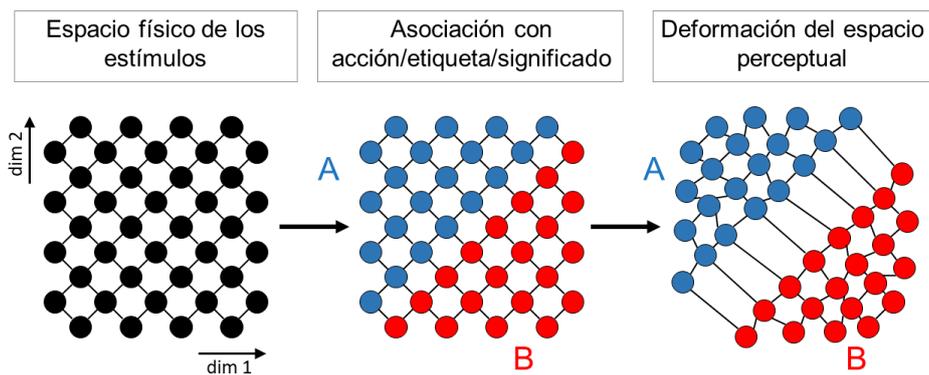
La categorización de estímulos es necesaria para formar decisiones y seleccionar acciones apropiadas para interactuar con el ambiente. Categorizar estímulos auditivos es la base de la comunicación acústica en humanos y otros animales. A pesar de su importancia, se sabe poco acerca de los mecanismos neuronales involucrados en la categorización de estímulos auditivos. En este proyecto investigamos el papel de la corteza premotora medial (CPM) durante la categorización de sonidos complejos como palabras, vocalizaciones de animales, y sonidos artificiales. Para ello, registramos la actividad eléctrica de neuronas de la CPM de dos monos Rhesus (*Macaca mulatta*) que realizaban una tarea de categorización de sonidos. En donde los monos debían agrupar estímulos auditivos dentro de una de dos categorías auditivas. Encontramos que las neuronas de la CPM representan fuertemente las categorías auditivas. También hallamos que la actividad poblacional de CPM durante el periodo en el que se presentan los estímulos auditivos es ortogonal a la ventana de tiempo en la que el animal realiza su respuesta conductual, lo que sugiere que la CPM genera cálculos distintos en cada periodo. En conjunto, nuestros resultados indican que CPM utiliza información auditiva para formar categorías conductuales.

# INTRODUCCIÓN

El reconocimiento y categorización de sonidos es la base de la comunicación en humanos y otros animales. Esta capacidad proviene de un conjunto de operaciones cerebrales que comienzan con la entrada de señales acústicas y culminan con una conducta (Bizley & Cohen, 2013). No obstante, los procesos involucrados son poco claros. Esta tesis se centró en estudiar la forma en que la corteza premotora medial (CPM), que es un área de integración sensorio-motora, participa en la categorización de señales auditivas. Para averiguarlo registramos neuronas de la CPM de un par de monos que realizaban una tarea de categorización de estímulos auditivos. A continuación, se describen algunos antecedentes importantes del presente proyecto.

## Fenómeno de categorización

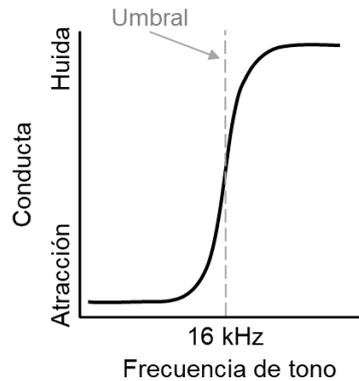
La categorización se define como la capacidad de ubicar objetos dentro de una clase o grupo. Perceptualmente, permite que las diferencias físicas de los objetos dentro una misma categoría se reduzcan, mientras que las incrementa entre objetos de categorías distintas (Figura 1; Min et al., 2020).



**Figura 1.** Deformación perceptual durante un proceso categorico. Izquierda: un grupo de estímulos en un espacio de dos dimensiones físicas. Centro: los estímulos físicos conforman dos categorías; cada categoría es una combinación específica de dos dimensiones. Derecha: durante la categorización, las representaciones perceptuales de los estímulos son deformadas; permitiendo que las diferencias de los estímulos dentro de las categorías se reduzcan y las diferencias entre categorías se agranden. Basado en Ley et al., 2014.

La categorización permite establecer umbrales a señales continuas que llegan a nuestros sentidos, lo cual hace que pequeñas variaciones en las propiedades físicas de los estímulos representen cambios repentinos a nivel perceptual (Seeger & Miller, 2010). Un claro ejemplo de este fenómeno se ha descrito en un tipo de grillo que realiza dos comportamientos opuestos en función de la frecuencia del sonido que escuchan (Figura 2). Con un tono mayor a 16 kHz los grillos huyen, con uno menor, son atraídos. La razón de esto es que sus señales de apareamiento están por debajo de los 16 kHz, mientras que

las vocalizaciones de sus depredadores tienen una frecuencia superior a ese umbral (Wytttenbach et al., 1996).



**Figura 2.** Percepción categórica en grillos. El umbral categórico define la conducta efectuada por los sujetos; si el estímulo está por encima del umbral, el sujeto huye, mientras que si está por debajo es atraído. Notar el cambio abrupto de la conducta entre categorías. Basado en Wytttenbach et al., 1996.

La agrupación categórica depende principalmente de la relevancia asignada a los estímulos. De ahí su alta flexibilidad y el hecho de que pueda ocurrir en una amplia gama de fenómenos: en cualquier modalidad sensorial (Freedman et al., 2001; Romo & Salinas, 2001; Tsunada et al., 2019), con cualquier característica física perceptible (intensidad, tamaño, forma o duración; Alvarez et al., 2015; Mendez et al., 2011; Mendoza et al., 2018; Sigala & Logothetis, 2002) o incluso empleando conceptos o reglas abstractas (p.ej. paz y amor; Engelmann & Pogosyan, 2013). La categorización es un paso crítico en procesos perceptuales y cognitivos. Es un puente que vincula procesos sensoriales como el reconocimiento y la percepción, con procesos como la toma de decisiones o la planificación de acciones (Freedman & Assad, 2011).

### ¿Qué sabemos acerca de la representación de categorías en el cerebro?

Una de las primeras evidencias que demuestran la existencia de neuronas categóricas en el cerebro fue realizada por Romo y colaboradores (Romo et al., 1997), quienes entrenaron a monos para categorizar la velocidad de estímulos somatosensoriales. Encontraron que la actividad de neuronas de CPM diferenciaba categóricamente la velocidad de los estímulos. Además, demostraron que la categorización está asociada a la toma de decisiones, ya que su actividad correlacionaba con la conducta de los sujetos.

Luego de esos descubrimientos se ha reportado que existen diversas áreas cerebrales que contienen neuronas con sintonizaciones categóricas. Estas áreas incluyen regiones corticales y subcorticales, así como áreas especializadas para el procesamiento de cada una de las modalidades sensoriales (Li et al., 2007). Por ejemplo, hay neuronas

categorías en áreas de procesamiento sensorial o perceptual, como las cortezas temporales de las modalidades visual y auditiva (Freedman et al., 2003; Tsunada et al., 2011), así como áreas parietales de la modalidad somatosensorial (Rossi-Pool et al., 2021). De igual forma, se han reportado neuronas categorías en regiones de procesamiento central, como las cortezas parietal, prefrontal y premotoras (Freedman et al., 2001; Romo et al., 1997; Zhou & Freedman, 2019).

Se piensa que la codificación categoría se construye gradual y jerárquicamente a lo largo de una red formada por distintas áreas cerebrales (Brincat et al., 2018; de Lafuente & Romo, 2006; Siegel et al., 2015). Tradicionalmente, se ha considerado que, en esta red, las estructuras más tempranas y cercanas a la entrada sensorial tienden a codificar características simples de los estímulos (ver Nienborg et al., 2012 para una postura más amplia acerca de la función de áreas sensoriales). Mientras que las posteriores representan menos las características de los estímulos y más las señales relacionadas con la percepción y la conducta.

Las señales categorías observadas en distintas regiones además conviven con procesos intrínsecos dentro de cada una de ellas (ya sean procesos sensoriales, perceptuales, cognitivos, etc.). Por ejemplo, se sabe que neuronas de la corteza inferior temporal, además de especializarse en representar información de las características físicas de estímulos visuales (como orientación, color y/o forma), modulan su tasa de disparo en función de las elecciones tomadas por los sujetos en tareas de categorización (Freedman et al., 2003; Koida & Komatsu, 2007). Algo similar se ha observado en cortezas prefrontales y premotoras, donde las señales categorías se mezclan con respuestas asociadas a la toma de decisiones y selección de acciones (Kobak et al., 2016; Mante et al., 2013; Lemus et al., 2007, 2009b, 2010).

## Neurofisiología del procesamiento auditivo en primates

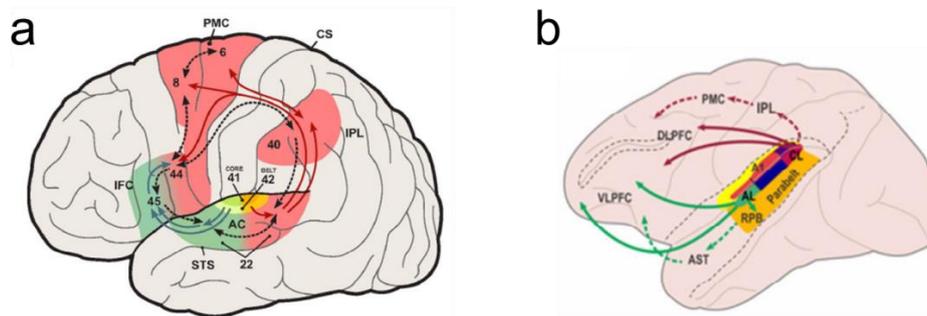
Tradicionalmente, se considera que la primera entrada de información auditiva a la corteza cerebral ocurre en el “núcleo” de la corteza auditiva primaria en el lóbulo temporal, la cual, a su vez, recibe señales directas del núcleo geniculado medial del tálamo (Kaas & Hackett, 2000; ver Camalier et al., 2012 y Hamilton et al., 2021 para una postura alternativa acerca del procesamiento de información sensorial a la corteza auditiva). La característica principal de la corteza auditiva primaria es que representa las características físicas de los sonidos a través de un mapa tonotópico, es decir, diferentes grupos de neuronas, separados espacialmente, se especializan en procesar determinadas frecuencias de los estímulos (Figura 3; Bizley, 2017).

Los dos siguientes estadios de procesamiento de la corteza auditiva corresponden al cinturón y al paracinturón, los cuales reciben ese nombre debido a que rodean al núcleo. Estas dos regiones, interconectadas entre sí, reciben su mayor entrada de información, directa o indirectamente, del núcleo de la corteza auditiva (Kaas & Hackett, 2000). Finalmente, se ha visto que tanto el cinturón como el paracinturón envían proyecciones



hicieron en animales que realizaban tareas pasivas, se sabe que estas áreas cumplen un rol crítico en funciones cognitivas complejas como la atención, memoria de trabajo, toma de decisiones y planeación de acciones (Miller et al., 2002); los cuales son cómputos vinculados con el proceso de categorización (Seger & Miller, 2010).

Se considera que durante la comunicación entre la corteza auditiva y diferentes regiones del lóbulo frontal ocurren fenómenos que incluyen el reconocimiento de objetos auditivos, la localización espacial y, en el caso de los humanos, el lenguaje (Rauschecker, 2021). De hecho, una de las preguntas fundamentales en neurociencia es entender la forma en la que la información auditiva llega a la corteza frontal, dónde se origina y cómo es utilizada para realizar conductas complejas, como la comunicación.



**Figura 4.** Flujo de procesamiento auditivo a través de las vías ventral y dorsal. a) Esquema de un cerebro humano en donde se representa el flujo de información desde la corteza auditiva hasta cortezas frontales y parietales. La vía ventral se muestra en verde, la dorsal en rojo. b) Esquema de un cerebro de macaco donde se representan las vías ventral y dorsal auditivas. En ambos cerebros, se ha propuesto que las vías comienzan en áreas sensoriales auditivas y finaliza en cortezas prefrontales y premotoras. Modificado de Rauschecker, 2018.

Se ha postulado que el flujo de procesamiento auditivo en monos se segrega paralelamente a lo largo de dos vías cerebrales (Figura 4; Rauschecker, 2021; Rauschecker & Tian, 2000; Romanski et al., 1999). La vía dorsal se especializa en la extracción de componentes espaciales de los estímulos (“¿Dónde está el sonido?”); mientras que la vía ventral procesa componentes no espaciales y más relacionados con el significado de los estímulos (“¿Qué es ese sonido?”).

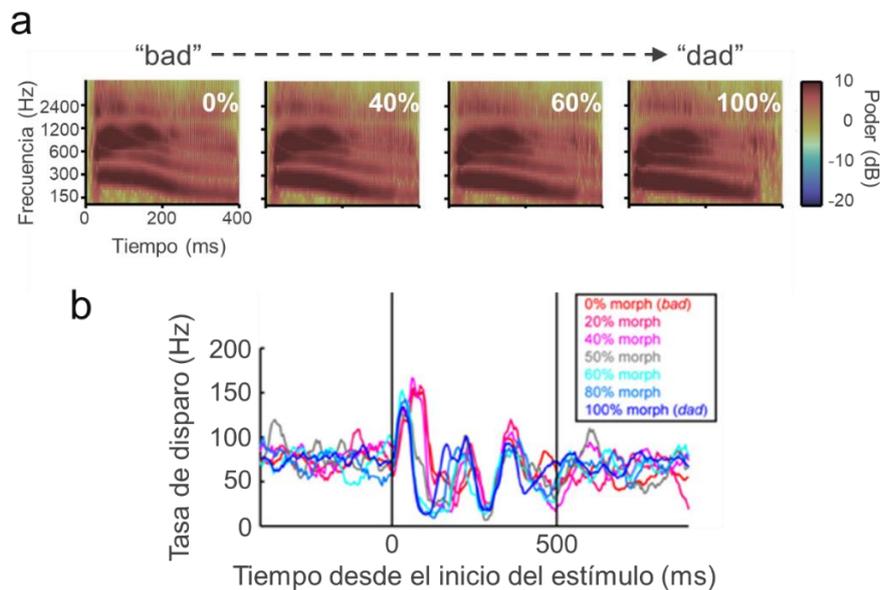
Las áreas involucradas de la vía dorsal incluyen las secciones caudales del cinturón y paracinturón, el plano temporal, el surco intraparietal y, en el lóbulo frontal, la corteza prefrontal y premotora (Rauschecker, 2018). En cambio, las áreas asociadas a la vía ventral incluyen la parte rostral del núcleo auditivo, el cinturón anterolateral y el paracinturón rostrottemporal; el cual, proyecta a la corteza prefrontal ventrolateral (Romanski et al., 1999). Aunque se considera que la vía ventral está involucrada en la categorización de vocalizaciones y palabras humanas, se sabe que existen procesos relacionados al significado de los sonidos en la vía dorsal, así como cómputos de información espacial en la vía ventral (Rauschecker, 2018). En humanos, se piensa que

la vía dorsal está muy involucrada en el procesamiento audiomotor del habla (Hickok & Poppel, 2007; Lima et al., 2016; Rauschecker, 2018).

### ¿Cómo se representan las categorías en el sistema auditivo?

Se piensa que el proceso de categorización de estímulos auditivos comienza desde etapas iniciales de la vía auditiva. Se ha sugerido que el núcleo de la corteza auditiva primaria procesa información categórica para distinguir estímulos auditivos simples como tonos puros (Christison-Lagay & Cohen, 2018; Tsunada et al., 2016; Xin et al., 2019). No obstante, otros reportes sugieren que las neuronas del núcleo auditivo codifican información de las características acústicas de los estímulos y no la pertenencia a cierta categoría (Lemus et al., 2009a).

Es posible que la categorización de sonidos realmente comience en las neuronas del cinturón y el paracinturón. Se ha visto que la actividad neuronal del cinturón lateral en monos y del giro temporal superior en humanos codifica categorías de sonidos como sílabas de palabras humanas y vocalizaciones de monos (Jiang et al., 2018; Tsunada et al., 2011; Yi et al., 2019). Por ejemplo, en monos entrenados en una tarea donde debían categorizar sonidos que contenían porcentajes variables de información entre dos fonemas (“bad” vs “dad”; estímulos conocidos como sonidos “morfeados”), se encontró que la actividad neuronal del cinturón lateral cambiaba discretamente en función del límite entre las dos categorías (Figura 5; Tsunada et al., 2011).



**Figura 5.** La actividad neuronal del cinturón lateral auditivo de presenta respuestas categóricas. a) Ejemplos de espectrogramas de estímulos “morfeados” que varían lineal y gradualmente desde el fonema “bad” al fonema “dad”. El porcentaje de contribución de “dad” se muestra en la parte superior de cada espectrograma. El porcentaje de contribución de “bad” es complementario al valor mostrado. b) Ejemplo de la tasa de disparo de una neurona categórica del cinturón lateral de macaco. El color de cada trazo corresponde al porcentaje de sonido “morfeado” mostrado en el recuadro. Tomado y modificado de Tsunada et al., 2011.

Fuera de las cortezas auditivas, se ha señalado que la corteza prefrontal ventral es un núcleo importante en el procesamiento categórico auditivo. Por ejemplo, en experimentos en monos durante tareas pasivas se han reportado neuronas en la corteza prefrontal que tienen selectividad a distintos tipos de vocalizaciones con específicas como “coos”, “grunts”, “barks”, etc. (Gifford et al., 2005; Romanski & Goldman-Rakic, 2002). Además, en tareas psicofísicas en donde monos son entrenados para categorizar fonemas, se ha visto que la actividad neuronal de esta área representa la transición entre una categoría y otra (Russ et al., 2008). Además, se ha reportado que neuronas de la corteza prefrontal correlacionan con la conducta de los sujetos durante tareas de discriminación auditiva (Tsunada et al., 2019).

Estos resultados concuerdan con lo encontrado en experimentos realizados en otras modalidades sensoriales en donde se ha visto que la corteza prefrontal está involucrada fuertemente en procesos de categorización, así como en fenómenos relacionados con el control cognitivo, como memoria de trabajo (Romo et al., 1999), toma de decisiones (Lemus et al., 2007) y selección de acciones (Miller et al., 2002). Por esta razón, se piensa que la corteza prefrontal participa en la generación de categorías abstractas. Por ejemplo, en humanos, esta área lleva información semántica y de valencia emocional (Muhle-Karbe & Flesch, 2020). En monos, se ha visto que separa en categorías discretas a vocalizaciones relacionadas con la calidad del alimento (Gifford et al., 2005).

Las evidencias apuntan a que la categorización auditiva se enclava en una red de áreas y procesos que se organizan de forma distribuida y jerárquica dentro del cerebro. Por ejemplo, en un artículo reciente realizado por el grupo de investigación de Shamma y Fritz (Yin et al., 2020), se describió que la información auditiva fluye ascendentemente a lo largo de una vía que involucra cortezas sensoriales auditivas, así como áreas prefrontales y premotoras. Los autores observaron que las señales categóricas son más fuertes y ocurren primero en las áreas frontales y que la actividad de estas áreas está asociada a las demandas conductuales durante la tarea. Además, encontraron que existe una cascada descendente en donde la información categórica fluye en una dirección que comienza desde las cortezas premotoras y prefrontales, luego la corteza auditiva secundaria y finaliza en la corteza auditiva primaria. Estos resultados concuerdan con tareas de categorización de estímulos visuales, en donde se ha visto que la actividad sensorial transita primero y rápidamente en una dirección “bottom-up”, seguida de un flujo inverso de información categórica que viaja en dirección “top-down” (Siegel et al., 2015).

### ¿Qué tipos de procesamiento categórico ocurre en áreas subsecuentes de la jerarquía cortical?

Más allá de los estudios realizados en las cortezas auditiva y prefrontal, hay pocos reportes en donde se investigue la participación de otras áreas cerebrales en procesos basados en información auditiva. Aun así, hay algunas áreas de interés que podrían estar

vinculando información proveniente de las cortezas auditivas y prefrontal, con procesos de salida como la toma de decisiones y la planeación y selección de acciones.

En particular, las cortezas premotoras son áreas que han atraído la atención debido a que están involucradas en procesos que enlazan la información sensorial con la salida motriz. Tradicionalmente, se sabe que las cortezas premotoras participan en el control cognitivo de movimientos y de acciones voluntarias (Lara et al., 2018; Nachev et al., 2008; Shima & Tanji, 2000). Sin embargo, aunque la actividad de estas áreas se ha relacionado canónicamente con el control motor, varios estudios indican que las cortezas premotoras participan activamente en procesos de memoria de trabajo y toma de decisiones en tareas de discriminación sensorial (Fine & Hayden, 2021; Hernández et al., 2002; Lemus et al., 2007; Romo et al., 1999; Romo et al., 2004). De hecho, en experimentos realizados por Romo y colaboradores, se han encontrado indicios de que la corteza premotora ventral y medial forman parte de circuitos que vinculan información auditiva y audiotáctil con cómputos relacionados a la toma de decisiones (Lemus et al., 2009b; Vergara et al., 2016).

CPM es un interesante candidato para entender los procesos de categorización auditiva. Es parte de un circuito de varias áreas corticales y subcorticales que participan en la percepción auditiva (Lima et al., 2016). Recibe proyecciones directas de los ganglios basales (Akkal et al., 2007; Catani et al., 2012), la ínsula, el surco temporal superior, el giro temporal superior y las cortezas parietal y frontal inferiores; conocidas en humanos como áreas del habla de Wernicke y Broca, respectivamente (Luppino et al., 1993).

Además, recientemente se ha encontrado que la CPM está involucrada en la percepción de sonidos que van dirigidos a generar acciones (Archakov et al., 2020) y a la formación de imágenes mentales acústicas, la selección de palabras y expresión de sonidos y palabras de relevancia social y emocional (Bestelmeyer et al., 2014; Binder et al., 2008; Lima et al., 2016), así como en procesos rítmicos, incluyendo al habla y el baile (Adank et al., 2013; Herrmann et al., 2014; Merchant et al., 2015).

## Problemáticas para entender los procesos neuronales asociados al reconocimiento y la categorización auditiva en modelos animales

En comparación con los sistemas visual y somatosensorial, las bases neurofisiológicas del procesamiento auditivo en primates no humanos han tenido algunos obstáculos importantes para ser estudiadas. Una de las razones es la dificultad para entrenar monos en tareas psicofísicas auditivas. Algunos estudios han encontrado que los monos tienen una habilidad reducida para recordar sonidos (Fritz et al., 2005; Scott et al., 2012). Esto ha ocasionado que muchas investigaciones acerca del procesamiento auditivo usen animales anestesiados o animales que realizan tareas pasivas. No obstante, estas estrategias conllevan limitaciones importantes, ya que sabemos que muchos procesos cognitivos relacionados a la percepción de señales sensoriales dependen de un reconocimiento activo de los sujetos.

Aunque los monos Rhesus han sido el modelo predilecto para entender procesos cerebrales en otras modalidades sensoriales, algunos grupos de investigación de neurociencia auditiva han optado por usar otro tipo de animales como modelo auditivo (ej. modelos murinos, hurones y aves; Gentner & Margoliash, 2003; Town et al., 2018; Yin et al., 2020). No obstante, algunos fenómenos, como el reconocimiento de señales naturalísticas (como las palabras), son difíciles de abordar con estos animales. Además de que, al ser modelos con cerebros sustancialmente distintos al de los humanos, en ocasiones es difícil extrapolar directamente los resultados neuronales.

Finalmente, otra problemática para estudiar procesos auditivos es que algunas áreas corticales relacionadas con la percepción auditiva son de difícil acceso para realizar experimentos electrofisiológicos. Estas áreas incluyen secciones del cinturón y el paracinturón auditivo, sobre todo las secciones más anteriores del lóbulo temporal. Este inconveniente ha impedido la descripción detallada de ciertos procesos de la vía ventral auditiva (la vía relacionada con el significado de los sonidos). No obstante, recientemente ha habido diseños experimentales exitosos en el registro neuronal de estas regiones (Fukushima et al., 2014).

## PLANTEAMIENTO DEL PROBLEMA

Por todo lo aquí descrito, este proyecto plantea que la actividad neuronal de CPM participa en la integración de información auditiva para su transformación en categorías discretas que ayuden a generar una conducta.

## HIPÓTESIS

Si CPM representa información categórica auditiva, entonces su actividad, más que codificar los atributos individuales de cada estímulo, se modulará por la información relacionada con la elección categórica reportada por los sujetos.

## OBJETIVOS

### Objetivo general

Estudiar los correlatos neuronales de la categorización auditiva en la corteza premotora medial en monos Rhesus.

### Objetivos particulares

1. Entrenar a dos monos Rhesus en una tarea de categorización auditiva usando estímulos naturalísticos.
2. Registrar los potenciales de acción de neuronas individuales de la corteza premotora medial mientras los animales realizan la tarea de categorización.
3. Analizar la forma en la que las neuronas de la corteza premotora medial representan la información de los parámetros de la tarea psicofísica (codificación de estímulos individuales, de las categorías y la relación entre la actividad conductual y neuronal).

## MÉTODOS Y RESULTADOS

Los métodos y resultados expuestos en este proyecto se encuentran publicados en Morán et al., 2021 (ver APÉNDICE). Además del artículo publicado, a continuación se presenta una breve síntesis de los experimentos y resultados encontrados.

En este proyecto diseñamos cuatro experimentos para determinar si CPM representa categorías auditivas:

- 1) Usamos un grupo de estímulos auditivos naturalísticos agrupados arbitrariamente en una de dos categorías. Los estímulos naturalísticos poseen características físicas multidimensionales; de modo que, si las neuronas de CPM representan información categórica, su actividad no llevará información acerca de los atributos individuales de cada estímulo, pero sí de la elección categórica reportada por los sujetos.
- 2) Usamos estímulos “morfeados”, es decir, estímulos auditivos discretos que cambian lineal y paulatinamente a lo largo de un continuo desde un estímulo de una categoría a un estímulo de otra categoría. Esto nos sirvió para evaluar si la actividad de CPM cambia linealmente en función de las dimensiones físicas de los estímulos o si refleja el resultado del reporte conductual.
- 3) Comparamos la actividad entre ensayos correctos e incorrectos para determinar si las neuronas de CPM predicen la elección conductual de los sujetos.
- 4) Implementamos un set pasivo en donde los sujetos escuchaban los estímulos sin generar ninguna conducta asociada a la tarea. Esto nos permitió determinar si la actividad neuronal depende de un reconocimiento activo de los estímulos.

Los registros electrofisiológicos de CPM se analizaron a nivel de células únicas y poblacionales. Llegamos a las siguientes conclusiones:

- 1) La actividad de las neuronas de CPM representan las categorías auditivas aprendidas por los monos. La respuesta neuronal agrupa a los estímulos dependiendo de su clase y muy poco en función de sus características físicas.
- 2) Aunque el diseño de nuestra tarea no nos permite identificar directamente si la actividad categórica observada en CPM tiene un origen sensorial o motor. Usando métodos de análisis poblacional, determinamos que las neuronas de CPM se organizan en dos estados poblacionales ortogonales, uno asociado al momento en el que el animal recibe el estímulo auditivo y otro relacionado con el momento en el que el animal reporta su respuesta. Esto sugiere que la misma población neuronal realiza cómputos distintos en cada uno de los dos periodos.
- 3) La actividad en respuesta a los estímulos “morfeados” mostró una transición desde una codificación ligeramente sensorial hasta una representación categórica correlacionada con la conducta de los sujetos.
- 4) Durante los errores, observamos que la actividad neuronal sufrió una inversión en la respuesta categórica, lo cual correlaciona con la decisión tomada por los sujetos.

- 5) Observamos que, durante los sets pasivos, las respuestas neuronales no llevan información acerca de los estímulos ni de las categorías. Esto sugiere que la dinámica de CPM depende de un estado conductual activo.

## DISCUSIÓN

En este proyecto investigamos si las neuronas de CPM codifican información útil para generar categorizaciones auditivas. Registramos neuronas de CPM mientras un par de monos realizaban una tarea de categorización de estímulos auditivos naturalísticos. El resultado principal del proyecto sugiere que la CPM usa información auditiva para producir señales categóricas útiles para la conducta.

Tras explorar el papel de las señales neuronales de CPM a través de diferentes métodos experimentales y analíticos, llegamos a las siguientes conclusiones. 1) A nivel de neuronas únicas y en las proyecciones poblacionales sobre descomposiciones de los componentes principales (dPCAs), encontramos que las señales categóricas emergen durante el periodo auditivo y se mantienen durante el periodo de retardo hasta alcanzar un máximo en el periodo de movimiento. 2) Aunque las señales categóricas están presentes durante estos tres periodos, encontramos que CPM se organiza en dos estados poblacionales distintos, uno durante el periodo auditivo y otro durante el de movimiento, lo que sugiere que ocurren dos cómputos diferentes en cada periodo. 3) Durante los sets con estímulos “morphing”, encontramos que la actividad neuronal transita desde una actividad que depende más de la información auditiva, hasta una actividad que correlaciona más con la respuesta conductual de los sujetos. 4) Durante los ensayos incorrectos, las señales categóricas se invertían y correlacionaban con el error del animal. Esto sugiere que las señales categóricas en CPM están íntimamente relacionadas con la elección tomada por los sujetos. 5) En este sentido, encontramos que las señales categóricas emergen sólo cuando los sujetos resuelven activamente la tarea y no cuando se presentan los estímulos de forma pasiva. A continuación, discutiremos estos resultados.

### Los monos categorizan exitosamente estímulos auditivos

Tradicionalmente se considera que los monos Rhesus tienen deficiencias para aprender tareas de reconocimiento auditivo debido a limitaciones para mantener estímulos auditivos en memoria de corto y largo plazo (Fritz et al., 2005; Scott et al., 2012). Incluso algunos autores han cuestionado la practicidad de usar monos Rhesus como modelo para estudiar los mecanismos neuronales relacionados a la percepción auditiva. No obstante, otros estudios han reportado haber logrado entrenar a monos en tareas psicofísicas de reconocimiento auditivo (Napoli et al., 2021; Ng et al., 2014; Tsunada et al., 2011; Tsunada et al., 2016).

En nuestro trabajo, a pesar de que el tiempo de entrenamiento de los monos fue relativamente largo (en comparación con paradigmas de las modalidades táctil y visual),

los monos demostraron ser capaces de reconocer y categorizar de forma eficiente diferentes estímulos auditivos naturalísticos durante la tarea psicofísica de categorización. En nuestro protocolo, el reconocimiento auditivo está basado en información almacenada a largo plazo, a diferencia de otros paradigmas donde se comparan estímulos usando memoria de corto plazo. Esto permitió que los monos aprendieran a asociar estímulos que arbitrariamente pertenecían a una de dos categorías (target y no-target) y generar una respuesta motriz toda vez que ocurriera un estímulo asociado a la recompensa (target). Con ello, además de demostrar que los monos son capaces de aprender asociaciones arbitrarias de estímulos auditivos, logramos hacer que reconocieran estímulos complejos como palabras, vocalizaciones y sonidos naturales.

En ese sentido, el planteamiento del proyecto fue que, si CPM representa información categórica, entonces su actividad no codificaría los atributos individuales de cada estímulo, pero sí la información acerca de la elección categórica reportada por los sujetos. Es decir, representaría información del significado en vez de información acústica, lo cual es de sumo interés para entender fenómenos relacionados con el habla y la comunicación.

## Las neuronas de CPM representan información de las categorías auditivas

Nuestros resultados demuestran que hay una proporción significativa de neuronas de CPM que codifican robustamente la categoría conductual de los estímulos auditivos (target vs no-target; 68.5% y 85.9% para cada mono). Particularmente, observamos que las señales categóricas, tanto a nivel de neuronas únicas como poblacional, comienzan a aparecer después de la primera mitad de la duración de los estímulos auditivos. Esto sugiere que, durante este tiempo, los sujetos integran información suficiente para reconocer y categorizar a los sonidos. Además, concuerda con lo encontrado en la corteza somatosensorial primaria durante una tarea de discriminación de frecuencias vibro-táctiles, en donde se vio que las neuronas integran información sensorial predominantemente durante el inicio de los estímulos (Luna et al., 2005). También coincide con experimentos de registros electrocorticográficos realizados en humanos, donde se ha visto que existen señales en el lóbulo temporal que marcan selectivamente el inicio de sonidos de palabras habladas (Hamilton et al., 2018).

De manera interesante, encontramos que una pequeña proporción de neuronas codifican significativamente la información individual de los estímulos auditivos; es decir, que distinguen entre diferentes estímulos que pertenecen a la misma categoría. Aunque este tipo de respuestas no son comúnmente reportadas en CPM durante otras tareas de categorización (Romo et al., 1997), existen experimentos de discriminación de estímulos somatosensoriales en donde se ha visto que neuronas de esta área representan la identidad de estímulos unidimensionales de forma monotónica (Romo et al., 1999; Hernández et al., 2002). Es posible que este tipo de respuestas puedan tener un origen en señales provenientes de la corteza auditiva (Lima et al., 2016).

Conforme transcurre el ensayo, observamos que, a nivel de neuronas únicas, la información categórica se mantiene estable durante el periodo de retardo y alcanza un pico máximo durante el periodo de respuesta motora. Proyectando la tasa de disparo poblacional sobre los dPCAs, encontramos al menos dos tipos de dinámicas en los componentes asociados a las categorías de los estímulos: una dinámica persistente, que incrementa después del estímulo auditivo, se mantiene durante el retardo y alcanza una máxima amplitud durante el periodo de movimiento; y una dinámica transitoria, que incrementa durante el estímulo y luego cambia de signo, ya sea durante el retardo o en el periodo de movimiento.

Observamos que los componentes categóricos de los dPCAs en CPM capturan una proporción significativamente alta del total de la varianza poblacional. De cierto modo es inesperado, ya que, en otros estudios, usualmente se ha visto que la varianza de los componentes temporales tiene un peso mayor en la dinámica de la población, en comparación con los componentes (dPCAs) asociados con los estímulos o la decisión (Kobak et al., 2016; Rossi-Pool et al., 2017).

### CPM se organiza en dos estados poblacionales distintos durante los periodos auditivo y de movimiento

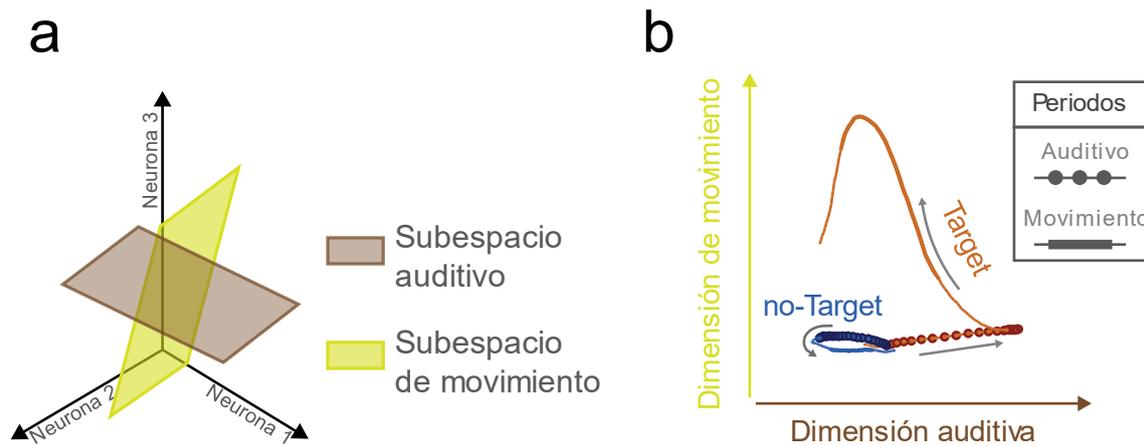
Encontramos que al proyectar la actividad neuronal sobre los dos componentes principales se observa una dinámica rotacional (Fig. 3G-H y 4H en Moran et al., 2021). Esto es similar a lo hallado en experimentos que requieren de evaluación sensorial, memoria de trabajo, planeación motora y/o generación de movimientos (Churchland et al., 2012; Mante et al., 2013). En nuestro experimento, suponemos que esta dinámica podría ser útil para llevar información acerca de la categoría de los sonidos durante el retardo y su vinculación con la generación de un plan motor para reportar la respuesta conductual. Concuera con la propuesta de que los circuitos neuronales usan dinámicas rotacionales para transformar información sensorial en información de memoria, mientras que se evitan que la traza de memoria sea deteriorada por la información sensorial (Libby & Buschman, 2021).

De manera interesante, se ha reportado que, en cortezas motoras, esta dinámica rotacional permite la formación de un espacio nulo en donde la actividad relacionada con la preparación de un movimiento ocurre sin realmente generar un movimiento. Aunque este espacio nulo es ortogonal al espacio del movimiento potencial, se vincula con él para llevar la información suficiente acerca de la salida motriz (Kaufman et al., 2014). En nuestro experimento, es posible que las dinámicas rotacionales en CPM sean parte de un proceso que vincule el reconocimiento y la categorización auditiva con la planeación y salida motora, evitando que haya interferencia de información en ambos tipos de fenómenos.

Una limitación importante en nuestra tarea es que su diseño impide separar directamente señales asociadas con el reconocimiento y categorización de los estímulos

auditivos de las señales relacionadas con la toma de decisiones y la planeación y ejecución motora. La razón de esto es que no existen límites temporales que nos permitan distinguir entre un fenómeno y otro. No obstante, inspirados en un trabajo reciente de Elsayed y colaboradores (2016), quisimos saber si ocurrían procesos distintos en dos periodos clave de la tarea: el periodo auditivo y el periodo de movimiento. El primero presuntamente relacionado con el reconocimiento de los sonidos y el segundo, vinculado con procesos de planeación y ejecución motriz.

Al aplicar los métodos desarrollados por Elsayed y colaboradores a nuestra tarea, encontramos que, similar a lo descubierto por ellos, existe una organización poblacional distinta entre los periodos auditivo y de movimiento. De hecho, encontramos que la dinámica durante el periodo auditivo es ortogonal a la dinámica del mismo pool de neuronas durante el periodo de movimiento. Estos resultados, junto a las dinámicas rotacionales, sugieren que, a pesar de no poder separar temporalmente procesos distintos, la misma población de neuronas en CPM se organiza como dos circuitos que realizan cálculos diferentes durante cada uno de estos periodos, uno más asociado al reconocimiento y categorización del estímulo auditivo y otro a la preparación y ejecución motora (Figura 6).



**Figura 6.** CPM se organiza en dos estados poblacionales ortogonales durante los periodos auditivos y movimiento de la tarea. a) Esquema hipotético de los subespacios formados por tres neuronas de CPM. Los subespacios auditivo y de movimiento se organizan como dos modos ortogonales entre sí. b) Dinámica de la actividad neuronal poblacional de CPM en la primer dimensión de los subespacios auditivo y de movimiento.

## Los cambios lineales de los estímulos modulan el comportamiento y la actividad neuronal de forma categórica

Hasta aquí, los resultados expuestos son un análisis de la actividad neuronal durante la categorización de estímulos naturalísticos. A continuación, evaluamos cómo se modula la conducta y la actividad neuronal de CPM durante la categorización de sonidos que cambian gradualmente de una categoría a otra. Este tipo de estímulos, llamados “morphing”, han sido utilizado ampliamente para entender si la actividad neuronal se

modula en función de las dimensiones físicas de los sonidos o, si en vez de eso, refleja el resultado del reporte conductual categórico hecho por los sujetos (Jiang et al., 2018; Tsunada et al., 2011; Yi et al., 2019).

Nuestros resultados muestran que durante el periodo auditivo existe una correlación más fuerte con los parámetros físicos de los estímulos y, conforme avanza el ensayo, la actividad neuronal se vuelve más categórica, alcanzando una correlación máxima con la conducta de los sujetos durante el periodo de movimiento.

Previamente, en estudios que han utilizado sonidos “morphing”, se ha visto que existe un proceso categórico que ocurre gradualmente a lo largo de diferentes etapas, en donde en los puntos extremos están la corteza auditiva secundaria y las cortezas frontales. Se piensa que las modulaciones categóricas de las áreas auditivas están asociadas con procesos sensoriales durante la formación de los objetos auditivos, mientras que las modulaciones en cortezas frontales se vinculan con procesos de toma de decisiones y selección de acciones (Yin et al., 2020). Esto concuerda con la idea de que la información categórica no ocurre en un solo sitio del cerebro, sino que se distribuye y transita a lo largo de áreas relacionadas con el procesamiento sensorial y perceptual hasta las áreas asociadas con la toma de decisiones categóricas y la planeación de acciones (Siegel et al., 2015). En nuestro experimento, sólo pudimos observar el procesamiento de información categórica auditiva en un área cerebral de salida o preparación motriz. Sin embargo, futuros experimentos podrán esclarecer la forma en la que una señal sensorial auditiva es transformada en una categoría en otras áreas cerebrales más cercanas a la entrada sensorial.

## La condición pasiva y los errores indican que CPM correlaciona con las decisiones categóricas

El análisis de los ensayos en donde el animal cometió errores nos fue útil para entender si la actividad de CPM está asociada con la elección categórica de los sujetos. De manera consistente con otros estudios, observamos que las señales categóricas se invierten durante los errores, lo cual es indicativo de su dependencia con las elecciones tomadas por los sujetos. Por ejemplo, la actividad incrementa cuando el mono considera que el estímulo presentado fue un sonido target, independientemente de si la señal sensorial realmente correspondía a un estímulo target. Particularmente, observamos que este fenómeno es más fuerte durante el retardo y el periodo de movimiento. La inversión de las señales categóricas durante la realización de ensayos erróneos es un fenómeno encontrado en cortezas frontales (Lemus et al., 2007, 2009b; Rossi-Pool et al., 2017).

Por otra parte, al analizar la actividad de CPM durante la presentación de los estímulos auditivos en una condición pasiva, encontramos que la actividad categórica desaparece por completo, lo cual sugiere que la dinámica de CPM depende mayoritariamente de una condición en donde el animal esté reconociendo y categorizando “activamente” a los estímulos. Esto mismo se ha encontrado en otros estudios donde se ha visto que las

señales neuronales de áreas frontales son dependientes de la realización “activa” de tareas cognitivas (Salinas & Romo, 1998; Yin et al., 2020).

## Consideraciones finales

Las evidencias mostradas aquí ilustran la complejidad para entender los cálculos que realiza el cerebro para realizar tareas de categorización auditiva. Queda claro que para estudiar este proceso se deben enfocar los esfuerzos para tratar de separar y distinguir cuidadosamente los diferentes tipos de señales que ocurren simultáneamente durante la categorización. Por ejemplo, distinguir entre las representaciones perceptuales de “bajo nivel” y “alto nivel”, así como los procesos de retroalimentación basados en la toma de decisiones y en las señales relacionadas con la resolución de la tarea psicofísica.

Este trabajo se enfocó en entender la dinámica del reconocimiento auditivo en un área premotora, sin embargo, aún falta comprender cuáles son los mecanismos implicados en gran parte de la cadena de información auditiva, sobre todo en etapas iniciales e intermedias. Nuestro trabajo es un antecedente para entender los procesos de categorización auditiva en el cerebro.

## CONCLUSIÓN

Esta tesis muestra cómo se procesa la información auditiva en CPM durante una tarea psicofísica de categorización de sonidos. Nuestros resultados indican que CPM representa robustamente la información categórica producida por el reconocimiento de sonidos, y que se organiza ortogonalmente en los periodos auditivos y de movimiento de la tarea. Esto sugiere que la misma población neuronal realiza dos cálculos distintos en cada periodo. Además, demostramos que la actividad categórica de la CPM correlaciona con la conducta de los sujetos. En conjunto, estos resultados sugieren que CPM usa información sensorial para generar señales categóricas útiles para un reporte conductual.

## REFERENCIAS

- Adank, P., Stewart, A. J., Connell, L., & Wood, J. (2013). Accent imitation positively affects language attitudes. *Frontiers in Psychology, 4*, 280.
- Akkal, D., Dum, R. P., & Strick, P. L. (2007). Supplementary motor area and presupplementary motor area: targets of basal ganglia and cerebellar output. *Journal of Neuroscience, 27*(40), 10659-10673.
- Alvarez, M., Zainos, A., & Romo, R. (2015). Decoding stimulus features in primate somatosensory cortex during perceptual categorization. *Proceedings of the National Academy of Sciences, 112*(15), 4773-4778.

- Archakov, D., DeWitt, I., Kuśmierk, P., Ortiz-Rios, M., Cameron, D., Cui, D., ... & Rauschecker, J. P. (2020). Auditory representation of learned sound sequences in motor regions of the macaque brain. *Proceedings of the National Academy of Sciences*, *117*(26), 15242-15252.
- Bestelmeyer, P. E., Maurage, P., Rouger, J., Latinus, M., & Belin, P. (2014). Adaptation to vocal expressions reveals multistep perception of auditory emotion. *Journal of Neuroscience*, *34*(24), 8098-8105.
- Binder, J. R., Swanson, S. J., Hammeke, T. A., & Sabsevitz, D. S. (2008). A comparison of five fMRI protocols for mapping speech comprehension systems. *Epilepsia*, *49*(12), 1980-1997.
- Bizley, J. K. (2017). Audition. In *Conn's Translational Neuroscience* (pp. 579-598). Academic Press.
- Bizley, J. K., & Cohen, Y. E. (2013). The what, where and how of auditory-object perception. *Nature Reviews Neuroscience*, *14*(10), 693-707.
- Brincat, S. L., Siegel, M., von Nicolai, C., & Miller, E. K. (2018). Gradual progression from sensory to task-related processing in cerebral cortex. *Proceedings of the National Academy of Sciences*, *115*(30), E7202-E7211.
- Camalier, C. R., D'Angelo, W. R., Sterbing-D'Angelo, S. J., Lisa, A., & Hackett, T. A. (2012). Neural latencies across auditory cortex of macaque support a dorsal stream supramodal timing advantage in primates. *Proceedings of the National Academy of Sciences*, *109*(44), 18168-18173.
- Catani, M., Dell'Acqua, F., Vergani, F., Malik, F., Hodge, H., Roy, P., ... & De Schotten, M. T. (2012). Short frontal lobe connections of the human brain. *cortex*, *48*(2), 273-291.
- Christison-Lagay, K. L., & Cohen, Y. E. (2018). The contribution of primary auditory cortex to auditory categorization in behaving monkeys. *Frontiers in neuroscience*, *12*, 601.
- Churchland, M. M., Cunningham, J. P., Kaufman, M. T., Foster, J. D., Nuyujukian, P., Ryu, S. I., & Shenoy, K. V. (2012). Neural population dynamics during reaching. *Nature*, *487*(7405), 51-56.
- de Lafuente, V., & Romo, R. (2006). Neural correlate of subjective sensory experience gradually builds up across cortical areas. *Proceedings of the National Academy of Sciences*, *103*(39), 14266-14271.
- Elsayed, G. F., Lara, A. H., Kaufman, M. T., Churchland, M. M., & Cunningham, J. P. (2016). Reorganization between preparatory and movement population responses in motor cortex. *Nature communications*, *7*(1), 1-15.
- Engelmann, J. B., & Pogosyan, M. (2013). Emotion perception across cultures: the role of cognitive mechanisms. *Frontiers in psychology*, *4*, 118.
- Fine, J. M., & Hayden, B. Y. (2021). The whole prefrontal cortex is premotor cortex. *arXiv preprint arXiv:2106.04651*.

- Freedman, D. J., & Assad, J. A. (2011). A proposed common neural mechanism for categorization and perceptual decisions. *Nature neuroscience*, *14*(2), 143-146.
- Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2001). Categorical representation of visual stimuli in the primate prefrontal cortex. *Science*, *291*(5502), 312-316.
- Freedman, D. J., Riesenhuber, M., Poggio, T., & Miller, E. K. (2003). A comparison of primate prefrontal and inferior temporal cortices during visual categorization. *Journal of Neuroscience*, *23*(12), 5235-5246.
- Fritz, J., Mishkin, M., & Saunders, R. C. (2005). In search of an auditory engram. *Proceedings of the National Academy of Sciences*, *102*(26), 9359-9364.
- Fukushima, M., Saunders, R. C., Leopold, D. A., Mishkin, M., & Averbeck, B. B. (2014). Differential coding of conspecific vocalizations in the ventral auditory cortical stream. *Journal of Neuroscience*, *34*(13), 4665-4676.
- Gentner, T. Q., & Margoliash, D. (2003). Neuronal populations and single cells representing learned auditory objects. *Nature*, *424*(6949), 669-674.
- Gifford III, G. W., MacLean, K. A., Hauser, M. D., & Cohen, Y. E. (2005). The neurophysiology of functionally meaningful categories: macaque ventrolateral prefrontal cortex plays a critical role in spontaneous categorization of species-specific vocalizations. *Journal of cognitive neuroscience*, *17*(9), 1471-1482.
- Hamilton, L. S., Edwards, E., & Chang, E. F. (2018). A spatial map of onset and sustained responses to speech in the human superior temporal gyrus. *Current Biology*, *28*(12), 1860-1871.
- Hamilton, L. S., Oganian, Y., Hall, J., & Chang, E. F. (2021). Parallel and distributed encoding of speech across human auditory cortex. *Cell*, *184*(18), 4626-4639.
- Hernández, A., Zainos, A., & Romo, R. (2002). Temporal evolution of a decision-making process in medial premotor cortex. *Neuron*, *33*(6), 959-972.
- Herrmann, B., Henry, M. J., Scharinger, M., & Obleser, J. (2014). Supplementary motor area activations predict individual differences in temporal-change sensitivity and its illusory distortions. *Neuroimage*, *101*, 370-379.
- Hickok, G., & Poeppel, D. (2007). The cortical organization of speech processing. *Nature reviews neuroscience*, *8*(5), 393-402.
- Jiang, X., Chevillet, M. A., Rauschecker, J. P., & Riesenhuber, M. (2018). Training humans to categorize monkey calls: auditory feature- and category-selective neural tuning changes. *Neuron*, *98*(2), 405-416.
- Kaas, J. H., & Hackett, T. A. (2000). Subdivisions of auditory cortex and processing streams in primates. *Proceedings of the National Academy of Sciences*, *97*(22), 11793-11799.
- Kaufman, M. T., Churchland, M. M., Ryu, S. I., & Shenoy, K. V. (2014). Cortical activity in the null space: permitting preparation without movement. *Nature neuroscience*, *17*(3), 440-448.

- Kobak, D., Brendel, W., Constantinidis, C., Feierstein, C. E., Kepecs, A., Mainen, Z. F., Qi, X. L., Romo, R., Uchida, N. & Machens, C. K. (2016). Demixed principal component analysis of neural population data. *Elife*, 5, e10989.
- Koida, K., & Komatsu, H. (2007). Effects of task demands on the responses of color-selective neurons in the inferior temporal cortex. *Nature neuroscience*, 10(1), 108-116.
- Lara, A. H., Cunningham, J. P., & Churchland, M. M. (2018). Different population dynamics in the supplementary motor area and motor cortex during reaching. *Nature communications*, 9(1), 1-16.
- Lemus, L., Hernández, A., & Romo, R. (2009a). Neural codes for perceptual discrimination of acoustic flutter in the primate auditory cortex. *Proceedings of the National Academy of Sciences*, 106(23), 9471-9476.
- Lemus, L., Hernández, A., & Romo, R. (2009b). Neural encoding of auditory discrimination in ventral premotor cortex. *Proceedings of the National Academy of Sciences*, 106(34), 14640-14645.
- Lemus, L., Hernández, A., Luna, R., Zainos, A., & Romo, R. (2010). Do sensory cortices process more than one sensory modality during perceptual judgments?. *Neuron*, 67(2), 335-348.
- Lemus, L., Hernández, A., Luna, R., Zainos, A., Nácher, V., & Romo, R. (2007). Neural correlates of a postponed decision report. *Proceedings of the National Academy of Sciences*, 104(43), 17174-17179.
- Ley, A., Vroomen, J., & Formisano, E. (2014). How learning to abstract shapes neural sound representations. *Frontiers in neuroscience*, 8, 132.
- Li, S., Ostwald, D., Giese, M., & Kourtzi, Z. (2007). Flexible coding for categorical decisions in the human brain. *Journal of Neuroscience*, 27(45), 12321-12330.
- Libby, A., & Buschman, T. J. (2021). Rotational dynamics reduce interference between sensory and memory representations. *Nature neuroscience*, 24(5), 715-726.
- Lima, C. F., Krishnan, S., & Scott, S. K. (2016). Roles of supplementary motor areas in auditory processing and auditory imagery. *Trends in neurosciences*, 39(8), 527-542.
- Luna, R., Hernández, A., Brody, C. D., & Romo, R. (2005). Neural codes for perceptual discrimination in primary somatosensory cortex. *Nature neuroscience*, 8(9), 1210-1219.
- Luppino, G., Matelli, M., Camarda, R., & Rizzolatti, G. (1993). Corticocortical connections of area F3 (SMA-proper) and area F6 (pre-SMA) in the macaque monkey. *Journal of Comparative Neurology*, 338(1), 114-140.
- Mante, V., Sussillo, D., Shenoy, K. V., & Newsome, W. T. (2013). Context-dependent computation by recurrent dynamics in prefrontal cortex. *nature*, 503(7474), 78-84.
- Mendez, J. C., Prado, L., Mendoza, G., & Merchant, H. (2011). Temporal and spatial categorization in human and non-human primates. *Frontiers in Integrative Neuroscience*, 5, 50.

- Mendoza, G., Méndez, J. C., Pérez, O., Prado, L., & Merchant, H. (2018). Neural basis for categorical boundaries in the primate pre-SMA during relative categorization of time intervals. *Nature communications*, 9(1), 1-17.
- Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., & Fitch, W. T. (2015). Finding the beat: a neural perspective across humans and non-human primates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 370(1664), 20140093.
- Miller, E. K., Freedman, D. J., & Wallis, J. D. (2002). The prefrontal cortex: categories, concepts and cognition. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 357(1424), 1123-1136.
- Min, B., Bliss, D. P., Sarma, A., Freedman, D. J., & Wang, X. J. (2020). A neural circuit mechanism of categorical perception: top-down signaling in the primate cortex. *BioRxiv*.
- Morán, I., Perez-Orive, J., Melchor, J., Figueroa, T., & Lemus, L. (2021). Auditory decisions in the supplementary motor area. *Progress in Neurobiology*, 202, 102053.
- Muhle-Karbe, P. S., & Flesch, T. (2020). Complementary Brain Signals for Categorical Decisions. *Journal of Neuroscience*, 40(30), 5706-5708.
- Nachev, P., Kennard, C., & Husain, M. (2008). Functional role of the supplementary and pre-supplementary motor areas. *Nature Reviews Neuroscience*, 9(11), 856-869.
- Napoli, J. L., Camalier, C. R., Brown, A. L., Jacobs, J., Mishkin, M. M., & Averbeck, B. B. (2021). Correlates of Auditory Decision-Making in Prefrontal, Auditory, and Basal Lateral Amygdala Cortical Areas. *Journal of Neuroscience*, 41(6), 1301-1316.
- Ng, C. W., Plakke, B., & Poremba, A. (2014). Neural correlates of auditory recognition memory in the primate dorsal temporal pole. *Journal of neurophysiology*, 111(3), 455-469.
- Nienborg, H., R. Cohen, M., & Cumming, B. G. (2012). Decision-related activity in sensory neurons: correlations among neurons and with behavior. *Annual review of neuroscience*, 35, 463-483.
- Plakke, B., & Romanski, L. M. (2014). Auditory connections and functions of prefrontal cortex. *Frontiers in neuroscience*, 8, 199.
- Rauschecker, J. P., & Tian, B. (2000). Mechanisms and streams for processing of “what” and “where” in auditory cortex. *Proceedings of the National Academy of Sciences*, 97(22), 11800-11806.
- Rauschecker, J. P. (2018). Where, When, and How: Are they all sensorimotor? Towards a unified view of the dorsal pathway in vision and audition. *Cortex*, 98, 262-268.
- Rauschecker, J. P. (2021). Central Auditory Processing. In *Oxford Research Encyclopedia of Neuroscience*.
- Rauschecker, J. P., Tian, B., & Hauser, M. (1995). Processing of complex sounds in the macaque nonprimary auditory cortex. *Science*, 268(5207), 111-114.

- Romanski, L. M., Tian, B., Fritz, J., Mishkin, M., Goldman-Rakic, P. S., & Rauschecker, J. P. (1999). Dual streams of auditory afferents target multiple domains in the primate prefrontal cortex. *Nature neuroscience*, 2(12), 1131-1136.
- Romanski, L. M., & Averbeck, B. B. (2009). The primate cortical auditory system and neural representation of conspecific vocalizations. *Annual review of neuroscience*, 32, 315-346.
- Romanski, L. M., & Goldman-Rakic, P. S. (2002). An auditory domain in primate prefrontal cortex. *Nature neuroscience*, 5(1), 15-16.
- Romanski, L. M., Averbeck, B. B., & Diltz, M. (2005). Neural representation of vocalizations in the primate ventrolateral prefrontal cortex. *Journal of Neurophysiology*, 93(2), 734-747.
- Romanski, L. M., Bates, J. F., & Goldman-Rakic, P. S. (1999). Auditory belt and parabelt projections to the prefrontal cortex in the rhesus monkey. *Journal of Comparative Neurology*, 403(2), 141-157.
- Romo, R., & Salinas, E. (2001). Touch and go: decision-making mechanisms in somatosensation. *Annual review of neuroscience*, 24(1), 107-137.
- Romo, R., Brody, C. D., Hernández, A., & Lemus, L. (1999). Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature*, 399(6735), 470-473.
- Romo, R., Hernández, A., & Zainos, A. (2004). Neuronal correlates of a perceptual decision in ventral premotor cortex. *Neuron*, 41(1), 165-173.
- Romo, R., Merchant, H., Zainos, A., & Hernández, A. (1997). Categorical perception of somesthetic stimuli: psychophysical measurements correlated with neuronal events in primate medial premotor cortex. *Cerebral cortex (New York, NY: 1991)*, 7(4), 317-326.
- Rossi-Pool, R., Zainos, A., Alvarez, M., Diaz-deLeon, G., & Romo, R. (2021). A continuum of invariant sensory and behavioral-context perceptual coding in secondary somatosensory cortex. *Nature communications*, 12(1), 1-13.
- Rossi-Pool, R., Zainos, A., Alvarez, M., Zizumbo, J., Vergara, J., & Romo, R. (2017). Decoding a decision process in the neuronal population of dorsal premotor cortex. *Neuron*, 96(6), 1432-1446.
- Russ, B. E., Orr, L. E., & Cohen, Y. E. (2008). Prefrontal neurons predict choices during an auditory same-different task. *Current Biology*, 18(19), 1483-1488.
- Salinas, E., & Romo, R. (1998). Conversion of sensory signals into motor commands in primary motor cortex. *Journal of Neuroscience*, 18(1), 499-511.
- Scott, B. H., Mishkin, M., & Yin, P. (2012). Monkeys have a limited form of short-term memory in audition. *Proceedings of the National Academy of Sciences*, 109(30), 12237-12241.
- Seger, C. A., & Miller, E. K. (2010). Category learning in the brain. *Annual review of neuroscience*, 33, 203-219.

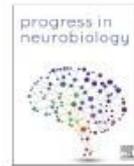
- Shima, K., & Tanji, J. (2000). Neuronal activity in the supplementary and presupplementary motor areas for temporal organization of multiple movements. *Journal of neurophysiology*, *84*(4), 2148-2160.
- Siegel, M., Buschman, T. J., & Miller, E. K. (2015). Cortical information flow during flexible sensorimotor decisions. *Science*, *348*(6241), 1352-1355.
- Sigala, N., & Logothetis, N. K. (2002). Visual categorization shapes feature selectivity in the primate temporal cortex. *Nature*, *415*(6869), 318-320.
- Town, S. M., Wood, K. C., & Bizley, J. K. (2018). Sound identity is represented robustly in auditory cortex during perceptual constancy. *Nature communications*, *9*(1), 1-15.
- Tsunada, J., Cohen, Y., & Gold, J. I. (2019). Post-decision processing in primate prefrontal cortex influences subsequent choices on an auditory decision-making task. *Elife*, *8*, e46770.
- Tsunada, J., Lee, J. H., & Cohen, Y. E. (2011). Representation of speech categories in the primate auditory cortex. *Journal of neurophysiology*, *105*(6), 2634-2646.
- Tsunada, J., Liu, A. S., Gold, J. I., & Cohen, Y. E. (2016). Causal contribution of primate auditory cortex to auditory perceptual decision-making. *Nature neuroscience*, *19*(1), 135-142.
- Vergara, J., Rivera, N., Rossi-Pool, R., & Romo, R. (2016). A neural parametric code for storing information of more than one sensory modality in working memory. *Neuron*, *89*(1), 54-62.
- Wytenbach, R. A., May, M. L., & Hoy, R. R. (1996). Categorical perception of sound frequency by crickets. *Science*, *273*(5281), 1542-1544.
- Xin, Y., Zhong, L., Zhang, Y., Zhou, T., Pan, J., & Xu, N. L. (2019). Sensory-to-category transformation via dynamic reorganization of ensemble structures in mouse auditory cortex. *Neuron*, *103*(5), 909-921.
- Yi, H. G., Leonard, M. K., & Chang, E. F. (2019). The encoding of speech sounds in the superior temporal gyrus. *Neuron*, *102*(6), 1096-1110.
- Yin, P., Strait, D. L., Radtke-Schuller, S., Fritz, J. B., & Shamma, S. A. (2020). Dynamics and hierarchical encoding of non-compact acoustic categories in auditory and frontal cortex. *Current Biology*, *30*(9), 1649-1663.
- Zhou, Y., & Freedman, D. J. (2019). Posterior parietal cortex plays a causal role in perceptual and categorical decisions. *Science*, *365*(6449), 180-185.



Contents lists available at ScienceDirect

Progress in Neurobiology

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



Original Research Article

## Auditory decisions in the supplementary motor area



Isaac Morán<sup>a,1</sup>, Javier Perez-Orive<sup>b,1</sup>, Jonathan Melchor<sup>a</sup>, Tonatiuh Figueroa<sup>a</sup>, Luis Lemus<sup>a,\*</sup>

<sup>a</sup> Department of Cognitive Neuroscience, Institute of Cell Physiology, Universidad Nacional Autónoma de México (UNAM), 04510, Mexico City, Mexico

<sup>b</sup> Instituto Nacional de Rehabilitación "Luis Guillermo Ibarra Ibarra", Mexico City, Mexico

### ARTICLE INFO

#### Keywords:

Word recognition  
Choices  
Premotor cortex  
Macaque  
Psychophysics

### ABSTRACT

In human speech and communication across various species, recognizing and categorizing sounds is fundamental for the selection of appropriate behaviors. However, how does the brain decide which action to perform based on sounds? We explored whether the supplementary motor area (SMA), responsible for linking sensory information to motor programs, also accounts for auditory-driven decision making. To this end, we trained two rhesus monkeys to discriminate between numerous naturalistic sounds and words learned as target (T) or non-target (nT) categories. We found that the SMA at single and population neuronal levels perform decision-related computations that transition from auditory to movement representations in this task. Moreover, we demonstrated that the neural population is organized orthogonally during the auditory and the movement periods, implying that the SMA performs different computations. In conclusion, our results suggest that the SMA integrates acoustic information in order to form categorical signals that drive behavior.

### 1. Introduction

Recognition and categorization of sensory information are essential for deciding which action to take (Freedman et al., 2001). In particular, categorizing dynamic sounds into classes is necessary for social communication in various species (Prather et al., 2009), including monkeys (May et al., 1989) and humans, where grouping of discrete sounds into words is fundamental for speech recognition (Leonard and Chang, 2014). However, it is not clear how the brain links the recognition of complex sounds to behavior.

Studies in primates have shown that the auditory information is carried from the auditory cortex to the prefrontal and premotor cortices via ventral and dorsal auditory streams (Kuśmierk and Rauschecker, 2014), and more recently, that premotor regions may represent an embodiment of acoustic recognitions (Archakov et al., 2020). The premotor supplementary motor area (SMA) has been linked to voluntary action and cognitive control of movement (Nachev et al., 2008; Lara et al., 2018; Shima and Tanji, 2000), working memory and decision making (Romo et al., 1999; Lemus et al., 2007).

The SMA is part of a circuit of several cortical and subcortical areas

that assist acoustic perception (Lima et al., 2016). The SMA receives direct projections from the basal ganglia (Lehéricy et al., 2004; Akkal et al., 2007), the superior temporal sulcus (STS; Luppino et al., 1993; Mayka et al., 2006), the superior temporal gyrus (STG; Reznik et al., 2015), and the inferior-parietal (Luppino et al., 1993) and inferior-frontal cortices (Catani et al., 2012; Vergani et al., 2014); also known as Wernicke's and Broca's speech areas, respectively. Consequently, the SMA is involved in the perception of action-directed sounds and acoustic imagery (Tian et al., 2016), selection of appropriate words (Binder et al., 2008), and expression of sounds and words of social and emotional relevance (Bestelmeyer et al., 2014; McGettigan et al., 2015). It also participates in rhythmic processes such as speech or dancing (Adank et al., 2013; Herrmann et al., 2014).

Nevertheless, few studies have been published on naturalistic sound coding during behavior (Tsunada et al., 2011; Town et al., 2018), and even fewer in non-human primates (Ng et al., 2009; Scott et al., 2012; Melchor et al., 2019). This could be because there is still plenty to learn from experiments using simple, non-naturalistic sounds that do not require behavior (e.g. pure tones, amplitude-modulated noise; Bendor and Wang, 2008; Yin et al., 2020). Additionally, although macaques

Abbreviations: dPC, demixed principal component; dPCA, demixed principal component analysis; nT, non-Target sound; P1, first temporal position of sounds; P2, second temporal position of sounds; P3, third temporal position of sounds; PCA, principal component analysis; SMA, supplementary motor area; T, target sound; VC, green visual cue.

\* Corresponding author.

E-mail address: [lemus@ifc.unam.mx](mailto:lemus@ifc.unam.mx) (L. Lemus).

<sup>1</sup> These authors contributed equally.

<https://doi.org/10.1016/j.pneurobio.2021.102053>

Received 22 December 2020; Received in revised form 6 April 2021; Accepted 20 April 2021

Available online 3 May 2021

0301-0082/© 2021 The Author(s). Published by Elsevier Ltd. This is an open access article under the CC BY license (<http://creativecommons.org/licenses/by/4.0/>).

have brain organizations that are similar to those of humans and although they, too, use vocalizations to communicate (Chang et al., 2010), some consider that non-human primates are challenging for training and studying auditory behavior (Ng et al., 2009; Scott et al., 2012). However, Melchor et al. (2019) recently showed that macaques attend to and memorize numerous naturalistic sounds. Therefore, in macaques, the SMA is amenable for studying the neural correlates of decisions within an auditory categorization task.

Here we present data from extracellular recordings of SMA neurons from two rhesus monkeys trained to discriminate between two groups, or categories, of naturalistic sounds and then choose an action. We also tested acoustic morphings to reveal the underlying responses to linear physical changes rather than discrete categorizations. We found robust categorical responses at both the single neuron and population levels. A small number of neurons were also found to carry information about individual sensory stimuli during the auditory period. By mapping the neural population activity during different periods of the task, we found that it organizes into orthogonal subspaces during the task's auditory and movement periods. Finally, the linear morphings of one stimulus into another indicate that the coding in SMA changes from sensory, during the auditory period, to categorical during the movement period. Taken together, our results support the hypothesis that SMA integrates auditory information necessary for a categorization that provides action signals for behavioral choices.

## 2. Methods

### 2.1. Ethics statement

Animals were handled following the Official Mexican Standard for the Care and Use of Laboratory Animals (NOM-062-ZOO-1999) and the study was approved by the Internal Committee for the Use and Care of Laboratory Animals of the Institute of Cell Physiology, UNAM (CICUAL; LLS80-16).

### 2.2. Animals and experimental setup

Two adult rhesus macaques (*Macaca mulatta*; monkey 1: a 10-year-old male weighing 13 kg; Monkey 2: a 10-year-old female weighing 7 kg) participated in the experiments. Training and recording sessions took place in a soundproof booth. Each monkey sat on a primate chair with its head fixed at a distance of 60 cm from a 21 color LCD monitor (1920 × 1080 resolution, 60 Hz refresh rate). A Yamaha MSP5 speaker (0.05–40 kHz frequency range) was situated fifteen centimeters above and behind the monitor delivering sounds at ~ 65 dB SPL (measured at the monkeys' ear level). Additionally, a Logitech Z120 speaker was placed directly below the Yamaha speaker to generate background white noise at ~ 55 dB SPL. Finally, a metal lever placed at the monkeys' waists captured their responses.

### 2.3. Task

We trained two rhesus macaques to perform an acoustic recognition task. During the task, the monkeys categorized sounds as either Target (T) or non-target (nT). In a trial, a gray circle of 3° aperture appeared at the center of the screen. Then, a sequence of one to three sounds arose. Each 0.5 s sound continued with a 0.5 s delay and then a 0.5 s green visual cue that replaced the initial gray circle. The cue indicated that the monkey should release the lever if the sound consisted of a T. The monkeys received a reward for releasing the lever in a time window of 0.7 s from the start of the T's cue. However, after an nT, the monkey kept the lever pressed, waiting for a T. If the monkey released the lever before a T, the trial was classified as a *false alarm*, and a new one commenced. We required that the monkeys' performance in this task remain above an 80 % hit rate before the electrophysiological recordings were made. The programming task was done in LabVIEW 2014 (64-bit SP1, National

Instruments®).

### 2.4. Animals' training

We implemented numerous strategies to instruct the monkeys to categorize two groups of heterogeneous sounds as T or nT. Nevertheless, the following are some general elements of the monkeys' learning process: First, they learned to press the lever down in response to a gray circle and release it to obtain a liquid reward only after the appearance of a 0.5 s monkey coo vocalization, a 0.5 s delay, and a 0.5 s green cue. After the monkeys recognized the coo vocalization as a T, we introduced an nT, a 0.5 s delay, and a 0.5 s green circle followed by the described T sequence. Eventually, the monkeys recognized several T and nT, so finally, in each trial, we included 0–2 nT randomly before a T.

### 2.5. Auditory stimuli

The sounds were recorded in our laboratory or downloaded from open access online libraries (<https://freesound.org/>). The monkeys learned more than 37 sounds from 4 different categories (Melchor et al., 2019): conspecific vocalizations, heterospecific vocalizations, words, and artificial sounds. However, for statistical repetitions the recordings were made with only 14 sounds (T = 7; nT = 7): conspecific vocalizations (T = 2; nT = 1), interspecific vocalizations (T = 1; nT = 1) and words (T = 2; nT = 3). The sounds' sample rate was 44.1 kHz; their amplitudes were normalized at -10 dB SPL (RMS), and their durations were shortened or lengthened to 0.5 s (Adobe Audition®).

### 2.6. Morphing sounds

We created sets of morphing sounds that change from an nT to a T in a continuum sequence (STRAIGHT software; Kawahara et al., 1999). This procedure consists of a linear interpolation of five spectro-temporal characteristics of a pair of sounds, i.e. fundamental frequencies, formants, duration, spectro-temporal density, and aperiodicity. The resulting sounds contained some percentage of stimulus A and a complementary percentage of B. Thus, each set consisted of seven morphed sounds (i.e., 0 20 40 50 60 80 and 100 % T). The monkeys trained in various sets; however, monkeys 1 and 2 executed 6 and 4 sets during recordings.

### 2.7. Passive condition

In this condition, the sounds appeared in a continuous clip separated by 0.3 s interlapses, and with no other cue or visual stimulus. The monkeys did not perform since the lever was not present. However, the monkeys received a reward randomly every 15–30 s.

### 2.8. Neural recordings

Extracellular recordings of neurons proceeded from an array of five independently movable microelectrodes (1–3 MΩ; Thomas Recording®) inserted at different SMA locations in each session. We positioned 20 mm diameter recording chambers according to the Paxinos brain atlas coordinates (Paxinos et al., 2000) and structural magnetic resonance imaging of both monkeys (interaural 27 mm, left hemisphere: lateral 6 mm). The sampled extracellular membrane potentials were 40 kHz (Plexon®). Isolation of individual neurons was performed manually using Plexon offline sorter software (Plexon®; monkey 1) and automatically using MountainSort (Chung et al., 2017; monkey 2).

### 2.9. Data analysis

#### 2.9.1. Individual neural coding

To evaluate whether SMA neurons coded for T and nT categories or

the identity of particular sounds, we performed F-statistics in a one-way ANOVA of the T vs. nT dependent activity and individual sounds. We aligned the neuronal activity to the start of sounds and calculated the mean firing rate in 200-ms windows sliding in steps of 20 ms. To avoid biases in the F-statistic due to comparing classes of different numbers of trials, we equalized the number of trials using the lowest common number of trials in all classes. We repeated the analysis 1000 times to create an F-statistic distribution. To determine the time bins where the F-statistic was significant, we compared the actual distribution against an F-statistic random distribution obtained by shuffling the stimulus labels 1000 times. A neuron was significant if the 95 % confidence intervals of the actual and random distributions did not overlap in at least one time bin.

### 2.9.2. Population analysis with PCA

We used principal component analysis (PCA) to describe the SMA's neuronal population dynamics during the task. We calculated the neurons' firing rate in blocks of correct trials of 1, 2, or 3 stimuli. We calculated the firing rate using a window of 200 ms moving in steps of 20 ms. Subsequently, we concatenated the firing rate of these three types of trials to obtain a matrix  $M \in \mathbb{R}^{N \times ET}$  of N number of neurons, E types of trials (with 1, 2, or 3 stimuli), and T total time points. Next, we used the demeaned firing rate of matrix M to calculate the principal components (PCs), considering each row as a variable.

### 2.9.3. Population analysis with demixed PCA

We performed demixed principal component analysis (dPCA; Kobak et al., 2016) in order to observe the effects of the task parameters on the population responses. We created a matrix  $M \in \mathbb{R}^{N \times S \times T}$  of the neuronal activity of N neurons (1341 and 524 for monkeys 1 and 2, respectively), responsive to  $S = 2$  types of sound (T and nT) in  $T = 106$  time points. The time points corresponded to 200 ms windows sliding every 20 ms from the beginning of the sounds to the monkeys' responses. During the supervised part, dPCA decomposes the neural activity for each variable using the covariance matrices of different marginalizations. The unsupervised analysis of dPCA was similar to that of the PCA, for each marginalization matrix. The dPCA algorithm separates the matrices from the decoder (D) and the encoder (E) for each parameter of the task by minimizing the loss function:

$$L_{\phi} = \|\bar{X}_{\phi} - F_{\phi} D_{\phi} \bar{X}\|$$

Where  $\phi$  denotes the marginalization for each parameter and  $\bar{X}$  is the mean-centered population matrix (the average of each neuron's activity is 0). Each component can be ordered based on the amount of variance explained. The axes obtained by the decoder and encoder allow reduction of the data in a few components that capture the most significant variance of each parameter of the task.

We implemented a method to determine the time bins with statistically correct decoding (Kobak et al., 2016). First, we separated the data into training and testing sets. We constructed the testing set by randomly separating one trial from each category (T and nT). In contrast, the training set was the average of the rest of the trials in each category. We performed dPCA from the training set to obtain the decoding axes. Then we projected the testing set on these axes and classified them according to the closest category mean (T or nT). We repeated this procedure 1000 times and measured the proportion of correct classification. We used 100 shuffles to calculate the classification accuracies expected by chance. The shuffled distribution was computed by choosing random trials between categories for each neuron. We ran the same method described above and calculated the proportion of correct classifications for each iteration. Then we compared the proportion of correct classifications between the real data vs. the shuffled data. The time bins where the real classification exceeded the shuffled classification distribution were considered statistically different from chance.

### 2.9.4. Identifying different subspaces in auditory and movement periods

To compare the population states during the auditory and movement periods, we performed four procedures to describe the transition between one to other state and a new algorithm developed by Elsayed et al. (2016) to calculate optimally two orthogonal subspaces, where each one captures auditory and movement activity. 1) To compare the population covariation between auditory and movement periods, we computed the correlation between all neurons in each period. We created two matrices for each period:  $A \in \mathbb{R}^{N \times CT}$  and  $M \in \mathbb{R}^{N \times CT}$  where N is the number of neurons, C is the number of stimuli and T is the time. The correlation matrix of each period was obtained computing the Pearson correlation over the rows of each matrix (A and M). 2) To quantify the subspace overlapping between subspaces of both periods, we computed a PCA over  $A \in \mathbb{R}^{N \times CT}$  and  $M \in \mathbb{R}^{N \times CT}$  to obtain the top ten auditory and movement PCs. First, the firing rate of each period was projected onto the auditory PCs and the percentage of explained variance relative to the total variance of the auditory period was quantified. Then, the same procedure was executed for movement PCs. 3) We calculated an alignment index to measure the amount of auditory period variance captured by the top ten movement PCs normalized by the auditory period variance captured by the top ten auditory PCs. The index range varied from 0 to 1, where 0 indicates the orthogonality between both periods and 1 means a perfect alignment. 4) To rule out that two groups of neurons are activated differentially in the auditory and movement periods we use an epoch-preference index, which evaluates the relative weight of each neuron in both periods. We calculated the firing rate range separately in each period and divided it by the mean of the firing rate of the entire trial. Then the range of each neuron was normalized by the mean of the firing rate of all neurons. A neuron with a positive value indicates a preference for the auditory period over the movement period, a negative value indicates the opposite. A bimodal distribution of the index signifies the existence of two neural subpopulations, one that is more selective for auditory period and other for movement period. Finally, we used a method developed by Elsayed et al. (2016) to derive two orthogonal subspaces, each one capturing auditory and movement activity, respectively. This method maximizes the sum of the variance of each period activity in each period subspace and was calculated with the following loss function:

$$\left[ \hat{Q}_{aud}, \hat{Q}_{mov} \right] = \text{argmax}_{\{Q_{aud}, Q_{mov}\}} \frac{1}{2} \left( \frac{\text{Tr}(Q_{aud}^T C_{aud} Q_{aud})}{\sum_{i=1}^{d_{aud}} \sigma_{aud}(i)} + \frac{\text{Tr}(Q_{mov}^T C_{mov} Q_{mov})}{\sum_{i=1}^{d_{mov}} \sigma_{mov}(i)} \right)$$

Where  $C_{aud}$  and  $C_{mov}$  are covariance matrices of the neural activity in auditory and movement period, respectively.  $\sigma_{aud}(i)$  y  $\sigma_{mov}(i)$  are ith singular value of the  $C_{aud}$  and  $C_{mov}$ , respectively.  $Q_{aud}$  y  $Q_{mov}$  are the eigenvectors of auditory and movement subspaces. We selected four dimensions for  $Q_{aud}$  and  $Q_{mov}$  due they capture close to 70 % variance of each period. This method identifies simultaneously the auditory and movement subspaces while constraining them to be orthogonal.

### 2.9.5. Psychometric and neurometric measures

We adjusted hyperbolic tangent functions to the psychophysical and neurometric performance during morphing sets. This function relates the percentage of times in which the monkey (psychometric curve) or the neuron (neurometric curve) recognized a sound like a T as a function of a morphing sound. The function is:

$$f(\%_{morph}) = \beta_0 + \beta_2 \tanh(\beta_1 (\%_{morph} - \theta))$$

$\theta$  and  $\beta_0$  represent the midpoint of the function on the horizontal and vertical axis, respectively,  $\beta_2$  represents the amplitude function from  $\beta_0$ , and  $\beta_1$  defines the morphing percentage dependent of behavior, a high value indicates an acute sigmoid. Additionally, we measured a linear slope at the sigmoid's midpoint to assess the degree of perceptual and neural modulation as a function of the morphing rate. We calculated the maximum and minimum values of the sigmoid's second derivative to

find the central part of the sigmoid. This section corresponds to the behavioral transition in the categorization of morphing stimuli. After obtaining the abscissa's and the ordinate's coordinates, we used the following equation to determine the slope:

$$\text{slope} = \frac{y_2 - y_1}{x_2 - x_1}$$

2.9.6. Neurometric function

A neurometric curve represents the probability that an ideal observer reports the stimulus category (T or nT) based solely on the firing rate. We used a methodology similar to that of De Lafuente and Romo (2005) to calculate the neurometric curves. For each neuron and time bin, the neurometric curve was the percentage of trials in which the firing rate of each trial reached or exceeded a criterion. The criterion was the one that in each bin best differentiated T and nT trials. We used 300 ms windows, sliding every 50 ms. In each window, we fitted a tangent-hyperbolic function to the probability to code a T as a function of morphing levels. Similarly, the psychometric curves were the monkeys' probability of perceiving morphing levels as a T. Psychometric and neurometric sensitivity were compared based on sigmoidal slopes. We used a Spearman's correlation between the raw data from the psychometric and neurometric curves.

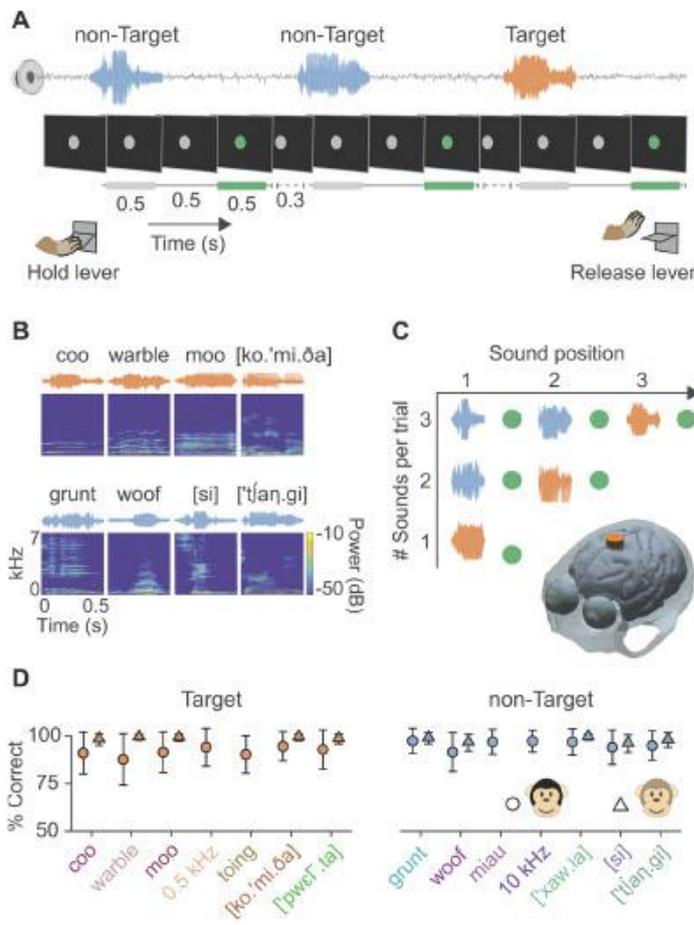
3. Results

3.1. Monkeys discriminate between learned naturalistic sounds

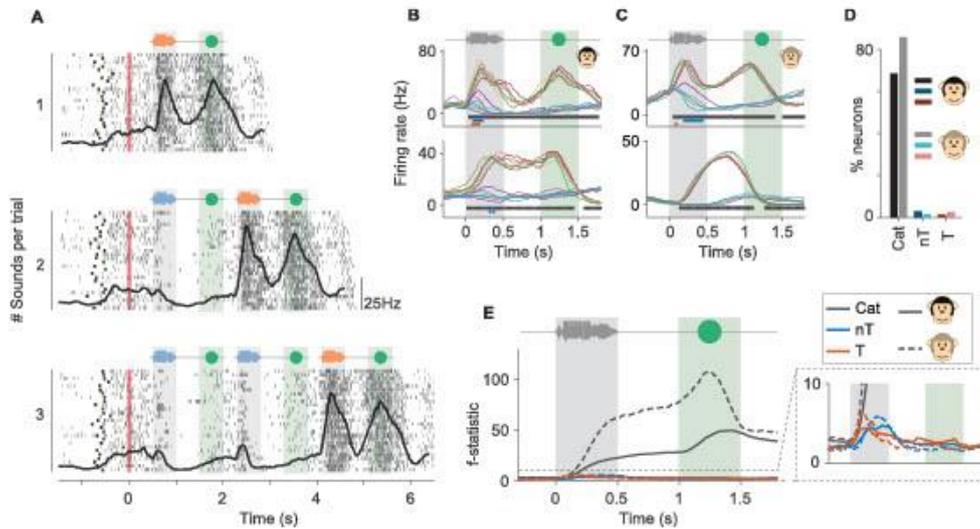
We trained two monkeys (*Macaca mulatta*) to perform an auditory recognition task. During the task, the monkeys obtained a drop of liquid as a reward for releasing the lever only after identifying a target sound (T). A T followed 0, 1, or 2 non-target sounds (nT). In other words, the T could appear in three different temporal positions (P1, P2, and P3; Fig. 1A,C, see Methods). It is important to note that we incorporated a delay and a green visual cue (VC) after each sound to disentangle auditory and movement signals. Thus, after a T, the monkeys waited for a 0.5 s delay before a 0.5 s VC-release period (Fig. 1A,C). The monkeys learned, by trial and error, to recognize up to fourteen sounds (half of them T, half nT) belonging to four different classes: conspecific vocalizations, interspecific vocalizations, words, and artificial sounds (Fig. 1B, D; see Methods). Once the monkeys performed above 85 % correct responses (Fig. 1D), we recorded single neurons (1341 from monkey 1; 524 from monkey 2) from SMA (Fig. 1C inset) during task execution.

3.2. Single units and population activity represent naturalistic sounds

Rasters of a characteristic SMA neuron (Fig. 2A) show a substantial firing rate change during a T regardless of its position (i.e. P1, P2, or P3). In other words, the responses show a clear distinction between T and nT,



**Fig. 1. Acoustic recognition task and behavior.**  
 A) Sequence of events in a trial: The monkey pressed the lever after a gray visual circle indicated the beginning of a trial. Then, a playback of one to three sequential sounds commenced. Each sound lasted 0.5 s followed by a 0.5 s delay and a 0.5 s green visual circle (go cue). The monkey received a reward for releasing the lever within a 0.7 s window after the green cue following a Target sound (T) (red sonograms). False alarms occurred when the monkey released the lever earlier, e.g. during a non-Target sound (nT) (blue sonograms). Finally, not releasing the lever during the T response window counted as a miss trial. B) Examples of sonograms and spectrograms of four T (top) and four nT (bottom) sounds. T and nT included four classes: conspecific vocalizations, interspecific vocalizations, words, and artificial sounds. C) Types of trials regarding the number of sounds. Inset, Brain scheme of MRI of monkey 2 showing the location of the recording chamber over SMA. D) Percentage of correct responses for T (red) and nT (blue) for each monkey (monkey 1: circles; monkey 2: triangles). The color code of stimuli will be referenced in the following figures.



**Fig. 2.** Individual SMA neurons code behavioral categories.

A) Rasters and peristimulus time histograms (PSTHs) of a representative neuron from monkey 1. Each row of ticks is one trial aligned with the lever press; each tick is an action potential. Each plot corresponds to the possible number of sounds in a trial. Thick black ticks indicate the beginning of a trial. Red ticks, lever pressings. B) Example PSTHs of two neurons for monkey 1 aligned with the start of the sounds. The color lines represent the PSTHs for each T and nT stimulus. The lines below indicate bins of significant categorical response for T or nT groups (black) or particular sounds within T (blue) or nT (red) categories. T stimuli are depicted in red and nT in blue palettes, as in Fig. 1D. The top neuron is the same as in (A). C) Same as (B) but for monkey 2. D) Percentage of recorded SMA neurons with statistically significant coding of behavioral categories (Cat) or sensory information (nT and T), and for each monkey. E) Average F-statistic for selective SMA neurons as a function of time for both monkeys (monkeys 1: continuous lines; 2: dashed lines); Inset, a magnification showing small F-statistic values for the few neurons with significant sensory coding for particular T or nT sounds.

but with small or no differences within each category (Fig. 2B and C). The majority of neurons recorded in both monkeys (68.5 % and 85.9 %, respectively) exhibited this type of categorical coding (Fig. 2D), as measured with the F-statistic and bootstrapping (see Methods). Interestingly, a small percentage of neurons displayed significant coding for specific sensory stimuli within the T and nT categories (2.4 % and 1.1 % for T; 1.1 % and 2.1 % for nT in each monkey, respectively; Fig. 2D). As shown in Fig. 2B, C, and E, neurons with this particular sensory coding did this during the auditory periods when sounds were presented (0.0–0.5 s). In contrast, the F-statistic shows that the categorical coding starting at the end of auditory periods rises further during the delays (0.5–1.0 s) and movement periods (1.0–1.5 s; Fig. 2E).

To study the neuronal population dynamics, we used principal component analysis (PCA), a dimensionality reduction technique. Similarly to the observed results of individual neurons, the population dynamics differentiated between T and nT, independently of the position in which T appears (Fig. 3A). The component explaining most of the variance shows a small increase at the beginning of the T stimulus, followed by a larger increase during the movement period, as expected for this premotor area. The second component shows acoustic-anticipatory activity that continues to rise with Ts but decreases for nTs, providing a clear auditory-based signal for premotor activity. The third principal component shows a large increase at the beginning of the trial and an additional increase before the lever release indicating the end of the trial.

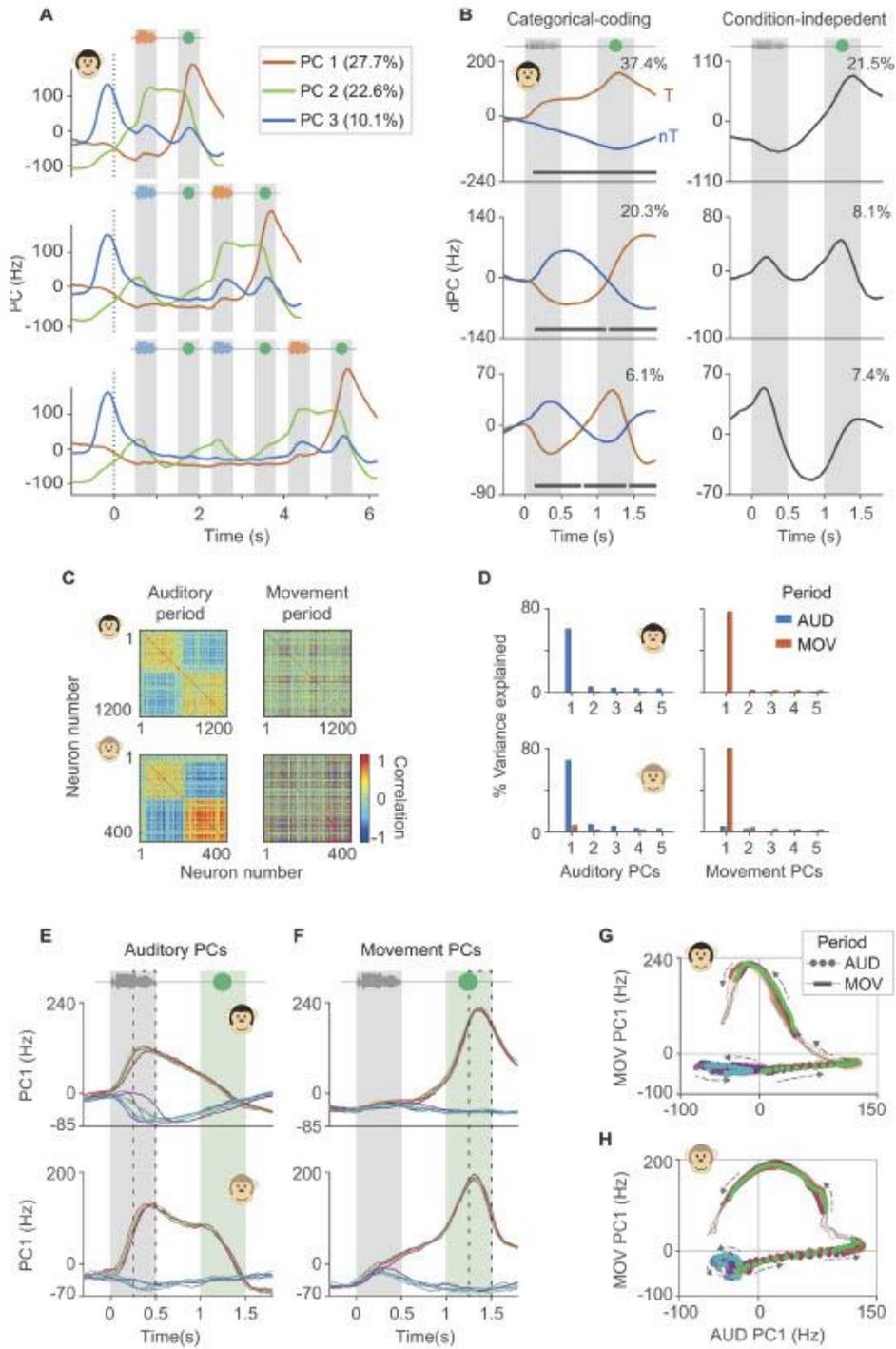
We also used demixed PCA (dPCA), a dimensionality reduction tool that can separate the effects on the population response of different task parameters (Kobak et al., 2016), to separate the effect of the auditory stimulus coding from the condition-independent activity (Fig. 3B). Since the previous results from single units showed strong categorical codification, we computed dPCA for the T and nT categories. The demixed principal components (dPCs) with the most weight explaining the variance of the population activity are those for the categorical coding

(first two categorical coding dPCs explain 57.7 % and 65.3 % of the total variance for each monkey, respectively, while the first two condition-independent dPCs explain 29.6 % and 29.5 % of the variance). For both monkeys, the dPCs separated the T and nT trajectories during the delay and reached its maximum during the movement period (Fig. 3B, top). Here, dPC2 and dPC3 shift their signs in the delay or during the movement period. The neural activity projections over the condition-independent dPCs are associated mainly with the task's temporal dynamics.

### 3.3. Auditory and movement computations occupy different subspaces

After observing the differences between auditory and movement period responses of single neurons and the SMA population, we set out to determine whether such representations proceeded from distinct population organizations, each coming together at different time periods to perform different computations (Elsayed et al., 2016). Fig. 3C highlights the correlation matrices for the full neural population, organized so the structure during the auditory period is highlighted, while keeping this same ordering for the movement period. The correlation matrices show two large groups of neurons for both monkeys, with high correlations during the auditory period, but not during the movement period (Fig. 3C). To evaluate whether this difference was due to two different populations, we calculated a simple epoch-preference index (Elsayed et al., 2016; see Methods). If the preference index distribution for all neurons is bimodal, this implies two pools of neurons. However, we found evidence that both monkeys' data came from a unimodal distribution (Hartigan's dip test  $p = 0.98$  and  $p = 0.96$  for monkeys 1 and 2, respectively), suggesting that the auditory and movement activities did not come from two different populations.

We used PCA to investigate whether these population responses during the auditory and movement epochs occupy partially overlapping



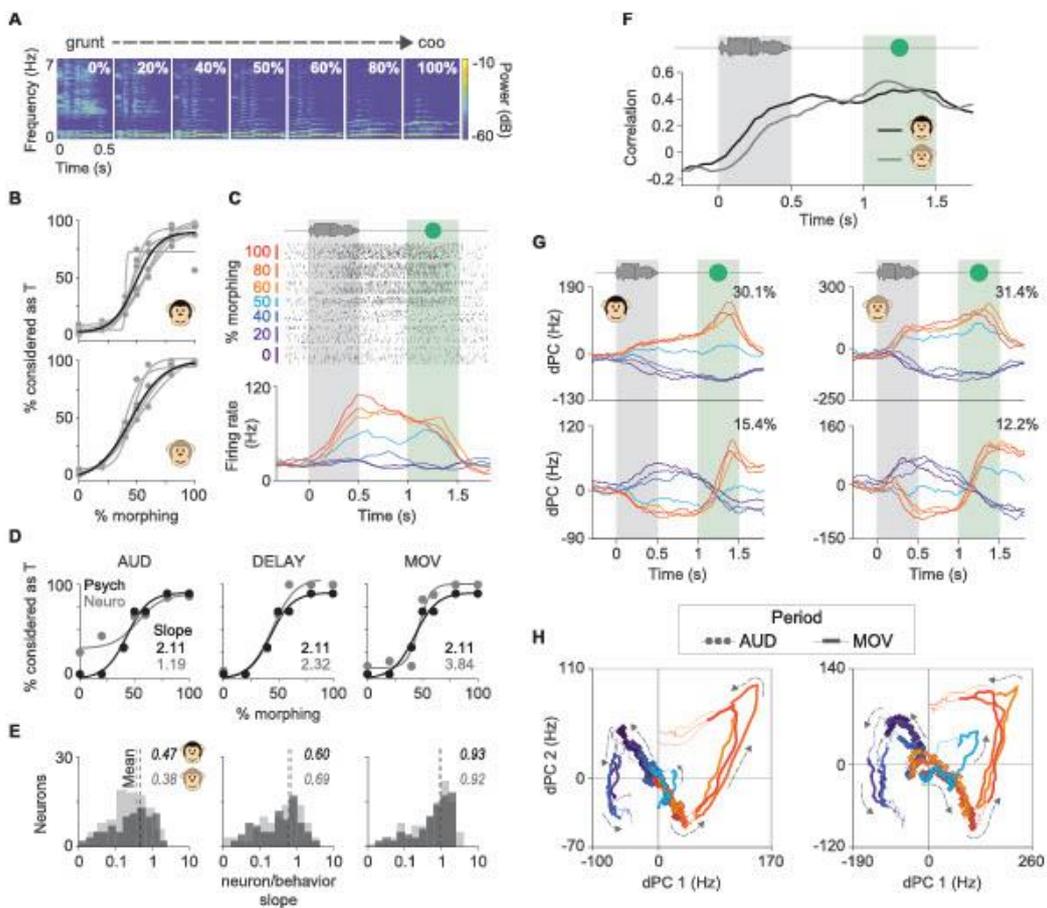
(caption on next page)

**Fig. 3.** Categorical responses of SMA population project in auditory and movement orthogonal subspaces.

A) Projections of the neural population responses in the first three principal components. The plot is grouped by the number of sounds, as in Fig. 2a. Notice PC1 codes for lever release, PC2 for anticipation of sounds together with the likelihood of a T, and PC3 for Ts and movements. B) Left: First three dPCs for categorical coding. Each trace corresponds to T (red) and nT (blue) projections. Right: First three condition-independent dPCs. The numbers on the top right are the percentages of explained variance of all dPCs. The first categorical coding components explain a greater amount of variance compared to the condition-independent components. C) Correlation matrices for auditory (left) and movement (right) periods of all neurons of monkeys 1 (top) and 2 (bottom). Each dot corresponds to the correlation between two neurons. Matrices for both periods sorted as a function of the auditory period. The groups of correlated neurons that exist in the auditory period disappear in the movement period. D) Percentage of explained variance by each of the top five acoustic (left) and movement (right) PCs projected onto the population activity during auditory and movement periods. E) Projections of the neural population activity onto the first PC of the auditory subspace for monkeys 1 (top) and 2 (bottom). F) Same as (E), but for the movement subspace. G) Two-dimensional projection of the auditory and movement orthogonal subspaces for monkey 1. Circles and thick lines show the auditory and movement periods, respectively. Arrows indicate flow of time. Same color code as in Fig. 1D. H) Same as (G), but for monkey 2.

or orthogonal subspaces within the neural activity space. By definition, the top five PCs during the auditory period capture a large amount of variance from this period. However, we found that they captured

minimal variance from the movement period (monkey 1: 76.9 % vs. 3.3 %; monkey 2: 88.8 % vs. 10.9 %, auditory vs. movement period variances, respectively, for PC1 to 5 combined; Fig. 3D). Similarly, the PCs



**Fig. 4.** Neural and behavioral responses to morphing sounds.

A) Spectrogram examples of gradually morphing sounds from grunt (nT) to coo (T). Every morphing set comprises seven morphed sounds. The percentage of T contribution in each sound is at the top of each spectrogram. B) Psychometric functions of monkeys (1: top; 2: bottom) as a function of percentage of morphing. Each line is a sigmoidal fit to the behavioral performance. The black line shows the mean performance for all morphings. C) Raster (top) and PSTHs (bottom) of a representative SMA neuron during the morphing set (grunt-coo) for monkey 2. D) Psychometric (black) and neurometric (gray) comparison in three periods of the neuron in C: auditory, delay, and movement. Each plot shows the percentage of times that behavior or neuronal response considered each morph as T. E) Slope ratios (neurometric/psychometric) of all neurons in three periods: auditory, delay, and movement. Monkeys 1: black; 2: gray. Dashed lines and numbers represent means of the distributions (monkeys 1: black; 2: gray). Note that the neurometric curves' slopes are smaller in the auditory period and become more similar to the psychometric curves during the movement period. F) Time-dependent Spearman correlations between neurometric and psychometric data points for both monkeys. G) Projection of the first two stimulus dPCs in both monkeys (1: left; 2: right). Colors correspond to a particular percentage of morphing (color code in C). H) Two-dimensional projection with the first two dPCs of the population activity for both monkeys (1: left; 2: right)—color code as in G. Circles and thick lines show auditory and movement periods, respectively. Arrows indicate time flow.

calculated for the movement period captured a large proportion of movement period variance, but a minimum of auditory period variance (monkey 1: 4.8 % vs. 86.5 %; monkey 2: 12.0 % vs. 92.8 %, auditory- vs. movement-period variances, respectively, for PC1 to 5; Fig. 3D). The top four PCs of the auditory subspace significantly captured 63.7 % and 75.8 % of the auditory period variance for monkeys 1 and 2, respectively (permutation test,  $p < 0.05$ ), but not significantly for the movement period, capturing only 0.7 % and 4.4 %, respectively. As for the movement subspace, the top four PCs captured 76.7 % and 82.0 % of the variance of the movement period variance (permutation test,  $p < 0.05$ ) for monkeys 1 and 2, respectively; and not significantly (0.7 % and 4.4 %) of the auditory period variance. This result shows that the SMA population's activity during the auditory and movement periods occupied nearly orthogonal subspaces.

We also calculated an alignment index to quantify the subspace overlap between the activities of both signals. Here, an index of zero means no alignment and one full alignment. The alignment index was 0.08 and 0.23 for monkeys 1 and 2, respectively, significantly lower than an alignment index calculated in a random neural sampling distribution ( $p < 0.05$  for both monkeys). This result indicates that the population rearranges its activity for auditory and movement periods in a nearly orthogonal manner. Therefore, since the neuronal population's activity behaves in an almost orthogonal fashion in the auditory and movement periods, it is possible to implement a linear decoder to generate two orthogonal subspaces in order to explain the neuronal variability population in both periods. Thus, we applied an algorithm (Elsayed et al., 2016) to find two sets of bases that identify principal components of each period in an orthogonal manner (Fig. 3E and F). The top four PCs from the auditory subspace captured 63.7 % and 75.8 % of the auditory variance in monkeys 1 and 2, respectively. The top four PCs from the movement subspace captured 76.7 % and 82.0 % of the movement variance for monkeys 1 and 2, respectively. Plotting one against the other revealed that the T and nT categories separate and move in opposite directions during the auditory period. Meanwhile, T elicits a broad trajectory during the movement period while nT remains close to the origin (Fig. 3G and H). There is a striking similarity between the monkeys' results, indicating that the appearance of these different population organizations at different times in the task is not monkey-dependent.

### 3.4. Linear stimulus changes modulate behavior and neural activity

Gradually morphing sounds into one another (Fig. 4A) may reveal the degree to which the neuronal activity changes as a function of the physical dimensions of the stimuli or reflect the behavioral outcome computations. By fitting sigmoidal functions to the monkeys' behavior (Fig. 4B; Table 1) and the neuronal responses to the morphings (Fig. 4C-D) it is possible to observe the relationship between the neuronal activity and behavior. The neuronal responses of an example neuron (Fig. 3C)

identified T at morphing values between 60 % and 100 %, and nT at morphing values between 0% and 40 %. Remarkably, a stimulus at 50 % produced a neuronal response between the two groups. However, there was a gradual change during the auditory period, whereas the responses transition more abruptly in the movement period. Fig. 4D shows this more precisely. The neurometric and psychometric functions are compared (gray and black lines, respectively) during the auditory, delay, and movement periods. During the latter periods of the task, the neurometric curves of all recorded neurons became more similar to the psychometric curve, i.e. with steeper slopes around the 50 % midpoint (Fig. 4E). The overall ratios of the neurometric/psychometric midpoint slopes increased progressively from the auditory, to the delay, to the movement periods: 0.47, 0.60, and 0.93 for monkey 1, and 0.38, 0.69, and 0.92 for monkey 2. Furthermore, Fig. 4F shows increasing Spearman correlations between neurometric and psychometric curves for all recorded neurons as the task progresses. These results imply that SMA neurons have more information about sensory stimuli during the auditory period and more categorical information (thus becoming more similar to the monkeys' behavior) during the movement period.

Demixed PCA revealed that, in response to morphing stimuli, the population activity followed a pattern similar to that of the single neurons, distinguishing between T and nT and with an in-between response at 50 % morphing (Fig. 4G). Divergence started during the auditory period, followed by a separation to two distinct attractors during the movement period (Fig. 4G-H). The two-dimensional dPC projections of the population activity (Fig. 4H) show how neuronal responses to T and nT sounds form two groups and move towards one of two attractors. The transition from one attractor to the other occurred abruptly, further evidencing the SMA's role in the categorical coding of naturalistic sounds.

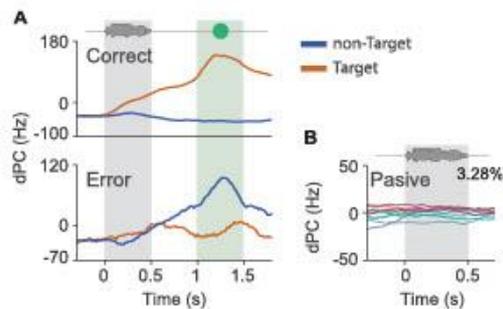
### 3.5. The non-alert condition and errors also point to a role played by SMA in acoustic decisions

To further distinguish between sensory and categorical coding, we projected the neural activity during the errors over the categorical dPCs calculated in correct responses. We found an inversion of activity resembling that of the correct trials (Fig. 5A). That is, in error trials the population activity in SMA reflected the behavioral decision made, rather than the sensory stimulus presented. Moreover, such population activity suggests that when a signal separation during the auditory period does not occur, monkeys make errors. We also recorded the neural activity during sets of passive listening, i.e. when the monkeys listened but did not perform the task because we withdrew the lever and the visual go-cues. Fig. 5B shows that during these sets, SMA ceased coding for auditory or categorical decisions. This result indicates that SMA is only active when the monkeys engage in the task, further supporting the role of SMA involvement during acoustic-driven decision making.

**Table 1**  
Overall fitting coefficients of psychometric curves.

non-Target	Target	Number of sessions	Number of neurons	Slope	$\theta$	$\beta_0$	$\beta_1$	$\beta_2$	$R^2$
<b>Monkey 1</b>									
grunt	coo	7	24	2.0	35.2	34.9	0.6	41.7	0.80
grunt	moo	6	36	1.9	60.3	51.3	0.5	51.0	0.91
grunt	warble	12	57	2.3	45.6	47.3	0.5	51.7	0.96
grunt	['pweɫ.ta]	10	47	1.9	59.6	52.2	0.4	55.9	0.93
grunt	[ko.'mi.ða]	8	38	1.8	60.2	56.2	0.4	52.5	0.93
[si]	moo	2	10	1.6	46.7	46.6	0.4	45.5	0.94
<b>Monkey 2</b>									
grunt	coo	13	91	1.6	52.9	52.0	0.3	59.4	0.96
grunt	moo	3	19	2.6	45.6	52.3	0.6	51.9	0.96
grunt	warble	3	24	1.5	51.6	51.7	0.3	56.6	0.94
['xaw.la]	['pweɫ.ta]	8	52	2.3	36.6	46.2	0.5	52.2	0.95

Grunt, coo, and warble are monkey vocalizations; moo: cow vocalization; ['xaw.la], [ko.'mi.ða], ['pweɫ.ta], and [si] are phonetic transcriptions of Spanish words.



**Fig. 5.** Error trials and the non-alert condition point to a participation of SMA in acoustic decisions.

A) The categorical dPCs calculated from correct trials (top) were used to project the population neural activity in error trials (bottom) in monkey 1. Note that when the dPCs do not separate during the auditory period, the monkeys make errors. B) Projection of first dPC over the neural population during passive trials shows no modulations.

#### 4. Discussion

To study whether the SMA codes for acoustic decisions, we trained two monkeys to categorize naturalistic sounds. Observing the time courses of the neural responses, analyzed in different ways, we present here several lines of evidence supporting the concept that SMA integrates sensory information during the auditory period, in support of categorical representations during the movement period: A) Demixed PCA, in which the main dPC began to separate in the auditory period, remained constant during the delay, and then rose sharply during the movement period. B) The two different organizational states were indicative of two different computations being performed in each, during the auditory and movement periods. C) Single neuron recordings during morphing sets showed coding-shifts from sensory to categorical representations as the task progressed. D) The activity of the neuronal population diverged to two distinct attractors, i.e. T and nT. E) Trials in which the monkeys made errors suggested that a separation in the neural population activity seems to be needed during the auditory period; additionally during the movement period the neuronal population with false alarms, had the same activity as during hits.

In previous work with monkeys trained to solve a delayed match-to-sample acoustic task, Scott et al. (2012) reported that the monkeys performed poorly in such tasks. In our study, although the monkeys' training was more prolonged than in visual or tactile paradigms (Lemus et al., 2009), the monkeys ended up being highly proficient in the categorization of a great diversity of sounds. Regardless of the complexity of natural sounds, in our task the neuronal activity channels them into two behaviors, which in contrast with previous investigations do not result from the physical similarities or direct mathematical comparisons of stimuli (Vergara et al., 2016). Instead, T and nT groups consisted of heterogeneous words and natural sounds we arbitrarily selected. Thus, the monkeys based their decisions on learned associations. In that vein, we propose that the SMA's auditory decisions may also proceed from semantic relationships rather than only phonetics. For example, the words "hear" and "listen" differ phonetically and yet they lead to similar behavior, whereas "hear" and "here," which sound the same, do not. Similarly, the sounds in our study could have appeared to the monkeys as synonyms of either "release-the-lever" or "hold-the-lever-down." Therefore, using natural sounds makes it possible to study realistic speech perception scenarios, such as when a single categorical behavior may emerge from phonetically diverse sounds.

As expected for a premotor area (Nachev et al., 2008; Lara et al., 2018; Shima and Tanji, 2000), during the task's movement period, single neurons and population responses coded for the monkeys'

dichotomous decisions in the task. Furthermore, dPCA showed that the task's categorical component captured most of the variance, as opposed to other commonly observed results with this method, in which the condition-independent components are often those with the most weight (Romo et al., 1999; Kobak et al., 2016). Interestingly, the second principal component (Fig. 3A) rises before each auditory stimulus in a predictable fashion, apparently consistent with models of accumulation of information (Leon and Shadlen, 1999). Despite the fact that such a phenomenon, similar to a hazard function of the probability of appearance of a T, was not previously described for auditory processing in SMA, it is consistent with its role in premotor processing (Shima and Tanji, 2000; Nachev et al., 2008; Zimmnik et al., 2019). Furthermore, this is the first description of how SMA neuronal activity correlates with decisions based on active listening to naturalistic sounds.

Interestingly, a small group of neurons coded for sounds during auditory periods, as predicted by imaging studies that suggested that the premotor cortex is activated by auditory stimuli (Lima et al., 2016; Vergara et al., 2016; Archakov et al., 2020). Nevertheless, to our knowledge, our results constitute the first electrophysiological evidence of SMA coding natural sounds. Previous fMRI studies using acoustic morphings (Jiang et al., 2018) suggested a two-step category-learning process. First, a perceptual stage in the auditory cortex, and later a categorization stage in the lateral prefrontal cortex. Moreover, Elsayed et al. (2016) showed that a population of neurons in the motor cortex acts as two coordinated circuits of preparatory and movement computations. When we applied their analytical method, we, too, found that the activity of a single neural population in SMA occupies orthogonal subspaces during the auditory and movement periods, organizing itself as two separate circuits during the acoustic and the motor periods.

We presented the monkeys with sounds that morphed gradually into one another, in order to determine how SMA responds to linearly changing physical stimuli. Measuring the slope at the midpoint of the sigmoidal curve of the neurometric and psychometric curves, we found that during the auditory period, the neurometric curves have smaller slopes than the psychometric curve, implying that at this moment of the task the neural response is more gradual, and varies more linearly with the changing sensory stimulus. During the movement period, both the neurometric and the psychometric curves had similar slopes, indicating that the neural responses are as categorical as the behavioral choices the animals make. From a population perspective, dPCA shows that sounds are grouped into two clear categories, with the dPC of the 50% response in between these groups. Furthermore, as revealed by the first and second principal components, the population response evolves towards two distinct attractors corresponding to the T and nT categories. The transition from one of these attractors to the other is abrupt rather than gradual, which is a hallmark of categorical coding (Wills et al., 2005; Freedman et al., 2001).

SMA neuronal activity was similar for all morphing sets regardless of heterogeneous stimuli compositions. In other words, response neurometric curves were not significantly different among acoustic categories used in morphing sets (Supplementary Fig. 1). Overall, based on our results, we propose that the SMA scales sensory and motor signals to match proper outcomes, e.g., firing at a higher rate for all Ts and suppressing the activity for nT. Some examples where this operation may be the case include beat synchronization to different intervals (Benoit et al., 2014; Merchant et al., 2015), completing syllables, words (Alario et al., 2006; Blank et al., 2002; Shahin et al., 2009), or sentences, or focusing on acoustic relevant features.

Recent experiments suggest that several cortical and subcortical areas are part of a network involved in decisions and motor control. For example, the anterior dorsal striatum initiates decision-making computations (Ding and Gold, 2013; Lim et al., 2014; Yartsev et al., 2018; Peters et al., 2021), and then projects to the SMA (Johansen-Berg et al., 2004; Lehericy et al., 2004), which is probably involved primarily in updating the decisions and controlling motor responses (Adank et al., 2013; Shahin et al., 2009; Herrmann et al., 2014). In the future,

simultaneous recordings of the SMA and subcortical regions may help to elucidate their role in acoustic perception.

#### Author contributions

IM, JM, TF & LL performed experiments. IM, & JPO analyzed data and prepared figures. LL designed the paradigm. TF programmed the task. JPO & LL wrote the paper.

#### Declaration of Competing Interest

The authors report no declarations of interest.

#### Acknowledgements

We thank Elizabeth Cabrera-Ruiz for valuable comments. We are also grateful for the financial support provided by CONACYT, CB-256767, and Programa de Apoyo a Proyectos de Investigación e Innovación Tecnológica (PAPIIT)IN207919. Special thanks to Gerardo Coello and Ana Escalante of the computing department of the IFC and to Patrick Weill for proofreading the manuscript. Isaac Morán Martínez is a doctoral student from Programa de Doctorado en Ciencias Biomédicas, Universidad Nacional Autónoma de México (UNAM) and received fellowship 288348 from CONACYT, Mexico. Data in this work are part of his doctoral dissertation.

#### Appendix A. The Peer Review Overview and Supplementary data

The Peer Review Overview and Supplementary data associated with this article can be found in the online version: <https://doi.org/10.1016/j.pneurobio.2021.102053>.

#### References

- Adank, P., Stewart, A.J., Connell, L., Wood, J., 2013. Accent imitation positively affects language attitudes. *Front. Psychol.* 4, 1–10. <https://doi.org/10.3389/fpsyg.2013.00280>.
- Akkal, D., Dum, R.P., Strick, P.L., 2007. Supplementary motor area and presupplementary motor area: targets of basal ganglia and cerebellar output. *J. Neurosci.* 27, 10659–10673. <https://doi.org/10.1523/JNEUROSCI.3134-07.2007>.
- Alario, F.X., Chainay, H., Lehericy, S., Cohen, L., 2006. The role of the supplementary motor area (SMA) in word production. *Brain Res.* 1076, 129–143. <https://doi.org/10.1016/j.brainres.2005.11.104>.
- Archakov, D., DeWitt, I., Kušmírek, P., Ortiz-Rios, M., Cameron, D., Cui, D., Morin, E.L., VanMeter, J.W., Sams, M., Jääskeläinen, I.P., Rauschecker, J.P., 2020. Auditory representation of learned sound sequences in motor regions of the macaque brain. *Proc. Natl. Acad. Sci. U. S. A.* 117, 15242–15252. <https://doi.org/10.1073/pnas.1915610117>.
- Bendor, D., Wang, X., 2008. Neural response properties of primary, rostral, and rostrotemporal core fields in the auditory cortex of marmoset monkeys. *J. Neurophysiol.* 100, 888–906. <https://doi.org/10.1152/jn.00884.2007>.
- Benoit, C.E., Dalla Bella, S., Farrugia, N., Obrig, H., Mainka, S., Kotz, S.A., 2014. Musically cued gait-training improves both perceptual and motor timing in Parkinson's disease. *Front. Hum. Neurosci.* 8, 1–11. <https://doi.org/10.3389/fnhum.2014.00494>.
- Bestelmeyer, P.E.G., Maura, P., Rouger, J., Latinus, M., Belin, P., 2014. Adaptation to vocal expressions reveals multistep perception of auditory emotion. *J. Neurosci.* 34, 8098–8105. <https://doi.org/10.1523/JNEUROSCI.4620-13.2014>.
- Binder, J.R., Swanson, S.J., Hammeke, T.A., Sabsevitz, D.S., 2008. A comparison of five fMRI protocols for mapping speech comprehension systems. *Epilepsia* 49, 1980–1997. <https://doi.org/10.1111/j.1528-1167.2008.01683.x>.
- Blank, S.C., Scott, S.K., Murphy, K., Warburton, E., Wise, R.J.S., 2002. Speech production: Wernicke, Broca and beyond. *Brain* 125, 1829–1838. <https://doi.org/10.1093/brain/awf191>.
- Catani, M., Dell'Acqua, F., Vergani, F., Malik, F., Hodge, H., Roy, P., Valabregue, R., Thiebaut de Schotten, M., 2012. Short frontal lobe connections of the human brain. *Cortex* 48, 273–291. <https://doi.org/10.1016/j.cortex.2011.12.001>.
- Chang, E.P., Rieger, J.W., Johnson, K., Berger, M.S., Barbaro, N.M., Knight, R.T., 2010. Categorical speech representation in human superior temporal gyrus. *Nat. Neurosci.* 13, 1428–1432. <https://doi.org/10.1038/nn.2641>.
- Chung, J.E., Magland, J.P., Barnett, A.H., Tolosa, V.M., Tocker, A.C., Lee, K.Y., Shah, K.G., Felix, S.H., Frank, L.M., Greengard, L.P., 2017. A fully automated approach to spike sorting. *Neuron* 95, 1381–1394. <https://doi.org/10.1016/j.neuron.2017.08.030>.
- De Lafuente, V., Romo, R., 2005. Neuronal correlates of subjective sensory experience. *Nat. Neurosci.* 8, 1698–1703. <https://doi.org/10.1038/nn1587>.
- Ding, L., Gold, J.I., 2013. The basal ganglia's contributions to perceptual decision making. *Neuron* 79, 640–649. <https://doi.org/10.1016/j.neuron.2013.07.042>.
- Elsayed, G.F., Lara, A.H., Kaufman, M.T., Churchland, M.M., Cunningham, J.P., 2016. Reorganization between preparatory and movement population responses in motor cortex. *Nat. Commun.* 7, 1–15. <https://doi.org/10.1038/ncomms13239>.
- Freedman, D.J., Riesenhuber, M., Poggio, T., Miller, E.K., 2001. Categorical representation of visual stimuli in the primate prefrontal cortex. *Science* 291, 312–316. <https://doi.org/10.1126/science.291.5502.312>.
- Herrmann, B., Henry, M.J., Scharinger, M., Obleser, J., 2014. Supplementary motor area activations predict individual differences in temporal-change sensitivity and its illusory distortions. *Neuroimage* 101, 370–379. <https://doi.org/10.1016/j.neuroimage.2014.07.026>.
- Jiang, X., Chevillet, M.A., Rauschecker, J.P., Riesenhuber, M., 2018. Training humans to categorize monkey calls: auditory feature- and category-selective neural tuning changes. *Neuron* 98, 405–416. <https://doi.org/10.1016/j.neuron.2018.03.014>.
- Johansen-Berg, H., Behrens, T.E.J., Robson, M.D., Drobnyak, I., Rushworth, M.F.S., Brady, J.M., Smith, S.M., Higham, D.J., Matthews, P.M., 2004. Changes in connectivity profiles define functionally distinct regions in human medial frontal cortex. *Proc. Natl. Acad. Sci. U. S. A.* 101, 13335–13340. <https://doi.org/10.1073/pnas.0403743101>.
- Kawahara, H., Masuda-Katsuse, I., De Cheveigné, A., 1999. Restructuring speech representations using a pitch-adaptive time-frequency smoothing and an instantaneous-frequency-based F0 extraction: possible role of a repetitive structure in sounds. *Speech Commun.* 27, 187–207. [https://doi.org/10.1016/S0167-6399\(98\)00065-5](https://doi.org/10.1016/S0167-6399(98)00065-5).
- Kobak, D., Brendel, W., Constantinidis, C., Feierstein, C.E., Kepecs, A., Mainen, Z.F., Qi, X.L., Romo, R., Uchida, N., Machens, C.K., 2016. Demixed principal component analysis of neural population data. *Elife* 5, 1–36. <https://doi.org/10.7554/eLife.10969>.
- Kuśmírek, P., Rauschecker, J.P., 2014. Selectivity for space and time in early areas of the auditory dorsal stream in the rhesus monkey. *J. Neurophysiol.* 111, 1671–1685. <https://doi.org/10.1152/jn.00436.2013>.
- Lara, A.H., Cunningham, J.P., Churchland, M.M., 2018. Different population dynamics in the supplementary motor area and motor cortex during reaching. *Nat. Commun.* 9, 1–16. <https://doi.org/10.1038/s41467-018-05146-z>.
- Lehéricy, S., Ducros, M., Krainik, A., Francois, C., Van De Moortele, P.F., Ugurbil, K., Kim, D.S., 2004. 3-D diffusion tensor axonal tracking shows distinct SMA and pre-SMA projections to the human striatum. *Cereb. Cortex* 14, 1302–1309. <https://doi.org/10.1093/cercor/bhh091>.
- Lemus, L., Hernández, A., Luna, R., Zainos, A., Nácher, V., Romo, R., 2007. Neural correlates of a postponed decision report. *Proc. Natl. Acad. Sci. U. S. A.* 104, 17174–17179. <https://doi.org/10.1073/pnas.0707961104>.
- Lemus, L., Hernández, A., Romo, R., 2009. Neural codes for perceptual discrimination of acoustic flutter in the primate auditory cortex. *Proc. Natl. Acad. Sci. U. S. A.* 106, 9471–9476. <https://doi.org/10.1073/pnas.0904066106>.
- Leon, M.I., Shadlen, M.N., 1999. Effect of expected reward magnitude on the response of neurons in the dorsolateral prefrontal cortex of the macaque. *Neuron* 24, 415–425. [https://doi.org/10.1016/S0896-6273\(00\)80854-5](https://doi.org/10.1016/S0896-6273(00)80854-5).
- Leonard, M.K., Chang, E.F., 2014. Dynamic speech representations in the human temporal lobe. *Trends Cogn. Sci.* 18, 472–479. <https://doi.org/10.1016/j.tics.2014.05.001>.
- Lim, S.-J., Flex, J.A., Holt, L.L., 2014. How may the basal ganglia contribute to auditory categorization and speech perception? *Front. Neurosci.* 8, 1–18. <https://doi.org/10.3389/fnins.2014.00230>.
- Lima, C.F., Krishnan, S., Scott, S.K., 2016. Roles of supplementary motor areas in auditory processing and auditory imagery. *Trends Neurosci.* 39, 527–542. <https://doi.org/10.1016/j.tins.2016.06.003>.
- Luppino, G., Matelli, M., Camarda, R., Rizzolatti, G., 1999. Corticocortical connections of area F3 (SMA-proper) and area P6 (pre-SMA) in the macaque monkey. *J. Comp. Neurol.* 338, 114–140. <https://doi.org/10.1002/cne.903380109>.
- May, B., Moody, D.B., Stebbins, W.C., 1989. Categorical perception of conspecific communication sounds by Japanese macaques. *Macaca Fuscata J. Acoust. Soc. Am.* 85, 837–847. <https://doi.org/10.1121/1.397555>.
- Mayka, M.A., Corcos, D.M., Leurgans, S.E., Vaillancourt, D.E., 2006. Three-dimensional locations and boundaries of motor and premotor cortices as defined by functional brain imaging: a meta-analysis. *Neuroimage* 31, 1453–1474. <https://doi.org/10.1016/j.neuroimage.2006.02.004>.
- McGettigan, C., Walsh, E., Jessop, R., Agnew, Z.K., Sauter, D.A., Warren, J.E., Scott, S.K., 2015. Individual differences in laughter perception reveal roles for mentalizing and sensorimotor systems in the evaluation of emotional authenticity. *Cereb. Cortex* 25, 246–257. <https://doi.org/10.1093/cercor/bht227>.
- Mejchor, J., Morán, I., Figueroa, T., Lemus, L., 2019. Perceptual Invariance of Words and Other Learned Sounds in Non-human Primates. *5. bioRxiv*, pp. 1–19. <https://doi.org/10.4324/9781315853178>.
- Merchant, H., Grahn, J., Trainor, L., Rohrmeier, M., Fitch, W.T., 2015. Finding the beat: a neural perspective across humans and non-human primates. *Philos. Trans. R. Soc. B Biol. Sci.* 370, 1–15. <https://doi.org/10.1098/rstb.2014.0093>.
- Nachev, P., Kennard, C., Husain, M., 2008. Functional role of the supplementary and pre-supplementary motor areas. *Nat. Rev. Neurosci.* 9, 856–869. <https://doi.org/10.1038/nrn2478>.
- Ng, C.W., Plakke, E., Poremba, A., 2009. Primate auditory recognition memory performance varies with sound type. *Hear. Res.* 256, 64–74. <https://doi.org/10.1016/j.heares.2009.06.014>.

- Paxinos, G., Huang, X.-P., Toga, A.W., 2000. *The Rhesus Monkey Brain in Stereotaxic Coordinates*. Academic Press.
- Peters, A.J., Pabre, J.M.J., Steinmetz, N.A., Harris, K.D., Carandini, M., 2021. Striatal activity topographically reflects cortical activity. *Nature* 591, 420–425. <https://doi.org/10.1038/s41586-020-03166-8>.
- Prather, J.F., Nowicki, S., Anderson, R.C., Peters, S., Mooney, R., 2009. Neural correlates of categorical perception in learned vocal communication. *Nat. Neurosci.* 12, 221–228. <https://doi.org/10.1038/nn.2246>.
- Reznik, D., Ossmy, O., Mukamel, R., 2015. Enhanced auditory evoked activity to self-generated sounds is mediated by primary and supplementary motor cortices. *J. Neurosci.* 35, 2173–2180. <https://doi.org/10.1523/JNEUROSCI.3723-14.2015>.
- Romo, R., Brody, C.D., Hernández, A., Lemus, L., 1999. Neuronal correlates of parametric working memory in the prefrontal cortex. *Nature* 399, 470–473. <https://doi.org/10.1038/20939>.
- Scott, B.H., Mishkin, M., Yin, P., 2012. Monkeys have a limited form of short-term memory in audition. *Proc. Natl. Acad. Sci. U. S. A.* 109, 12237–12241. <https://doi.org/10.1073/pnas.1209685109>.
- Shahin, A.J., Bishop, C.W., Miller, L.M., 2009. Neural mechanisms for illusory filling-in of degraded speech. *Neuroimage* 44, 1133–1143. <https://doi.org/10.1016/j.neuroimage.2008.09.045>.
- Shima, K., Tanji, J., 2000. Neuronal activity in the supplementary and presupplementary motor areas for temporal organization of multiple movements. *J. Neurophysiol.* 84, 2148–2160. <https://doi.org/10.1152/jn.2000.84.4.2148>.
- Tian, X., Zarate, J.M., Poeppel, D., 2016. Mental imagery of speech implicates two mechanisms of perceptual reactivation. *Cortex* 77, 1–12. <https://doi.org/10.1016/j.cortex.2016.01.002>.
- Town, S.M., Wood, K.C., Bizley, J.K., 2018. Sound identity is represented robustly in auditory cortex during perceptual constancy. *Nat. Commun.* 9, 1–15. <https://doi.org/10.1038/s41467-018-07237-3>.
- Tanada, J., Lee, J.H., Cohen, Y.E., 2011. Representation of speech categories in the primate auditory cortex. *J. Neurophysiol.* 105, 2634–2646. <https://doi.org/10.1152/jn.00037.2011>.
- Vergani, P., Lacerda, L., Martino, J., Attems, J., Morris, C., Mitchell, P., De Schotten, M. T., Dell'Acqua, F., 2014. White matter connections of the supplementary motor area in humans. *J. Neurol. Neurosurg. Psychiatry* 85, 1377–1385. <https://doi.org/10.1136/jnnp-2013-307492>.
- Vergara, J., Rivera, N., Rossi-Pool, R., Romo, R., 2016. A neural parametric code for storing information of more than one sensory modality in working memory. *Neuron* 89, 54–62. <https://doi.org/10.1016/j.neuron.2015.11.026>.
- Wills, T.J., Lever, C., Cacucci, F., Burgess, N., O'Keefe, J., 2005. Attractor dynamics in the hippocampal representation of the local environment. *Science* (80) 308, 873–876. <https://doi.org/10.1126/science.1108905>.
- Yartsev, M.M., Hanks, T.D., Yoon, A.M., Brody, C.D., 2018. Causal contribution and dynamical encoding in the striatum during evidence accumulation. *Elife* 7, 1–24. <https://doi.org/10.7554/eLife.34929>.
- Yin, P., Strait, D.L., Radtke-Schuller, S., Pritz, J.B., Shamma, S.A., 2020. Dynamics and hierarchical encoding of non-compact acoustic categories in auditory and frontal cortex. *Curr. Biol.* 30, 1649–1663. <https://doi.org/10.1016/j.cub.2020.02.047>.
- Zimnik, A.J., Lara, A.H., Churchland, M.M., 2019. Perturbation of macaque supplementary motor area produces context-independent changes in the probability of movement initiation. *J. Neurosci.* 39, 3217–3233. <https://doi.org/10.1523/JNEUROSCI.2335-18.2019>.